Notes on Dynamic Programming

Michael N. Katehakis

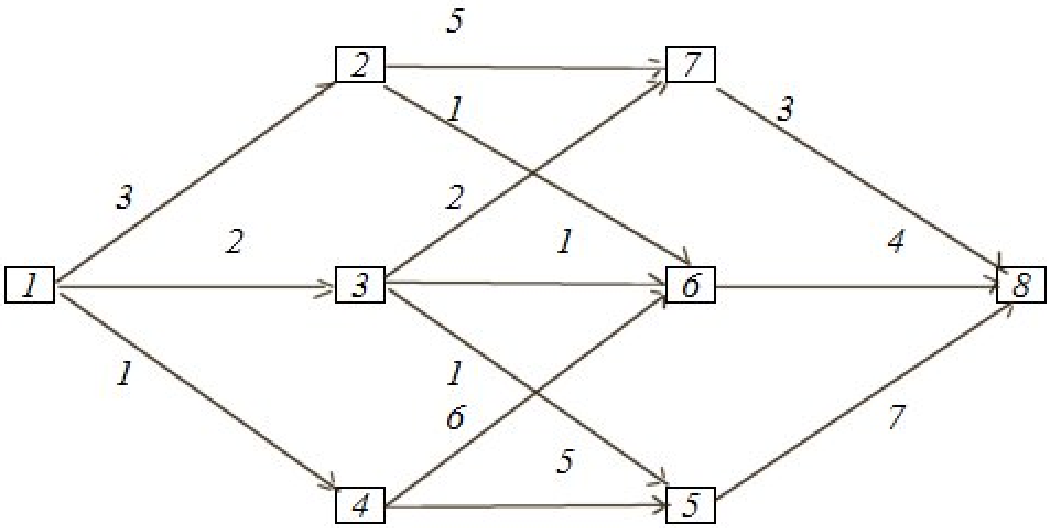
University of Crete
1988

ΚΕΦΑΛΑΙΟ 1.

ΤΟ ΠΡΟΒΛΗΜΑ ΤΟΥ ΠΕΠΕΡΑΣΜΕΝΟΥ ΟΡΙΖΟΝΤΑ ΜΕ ΠΡΟΣΔΙΟΡΙΣΤΙΚΗ ΔΥΝΑΜΙΚΗ

1. Το πρόβλημα της Διαδρομής Ελάχιστου μήκους.

1.1. Εισαγωγικό Παράδειγμα. *Ας θεωρήσουμε ένα δέντρο από πόλεις (κόμβους) που είναι συνδεδεμένες μεταξύ τους μέσω του (οδικού) δικτύου της εικόνας 1 .*



Εικόνα 1.

Οι αριθμοί δίπλα στα βέλη συμβολίζουν τα μήκη των αντίστοιχων δρόμων. Το ζητούμενο είναι να βρούμε τη (τις) διαδρομή (διαδρομές) που συνδέει (συνδέουν) την πόλη 1 με την πόλη 8 που έχει (έχουν) το ελάχιστο δυνατό μήκος.

1.2. Η Εξίσωση Βελτιστοποίησης . *Μια πρώτη ερώτηση, που φαίνεται απλούστερη, είναι: "ποιό είναι το* ***μήκος*** *της διαδρομής ελάχιστου μήκους από την πόλη 1 στην πόλη 8 " ;*

Για να απαντήσουμε στην παραπάνω ερώτηση, ας ορίσουμε για κάθε πόλη (κόμβο) x τη συνάρτηση $v(x)$, με την απαίτηση ότι το $v(x)$ συμβολίζει το μήκος της διαδρομής ελάχιστου μήκους από την πόλη x στην πόλη 8 . Στο παράδειγμα έχουμε οτι $v(8) = 0$, και η απάντηση στην ερώτηση που θέσαμε είναι η αγνωστη προς το παρόν τιμή $v(1)$.

Διαθέτουμε λοιπόν τώρα ένα πρώτο εργαλείο (γλώσσα) που μπορούμε να χρησιμοποιήσουμε στην προσπάθειά μας να απαντήσουμε στην παραπάνω ερώτηση. Τώρα η ερώτηση είναι : "πώς μπορούμε να υπολογίσουμε την άγνωστη συνάρτηση $v(\cdot)$; ". Ο συνήθης τρόπος υπολογισμού αγνώστων ποσοτήτων περιλαμβάνει την εύρεση μιας εξίσωσης (ή ενός συστήματος εξισώσεων) της οποίας (των οποίων) η λύση να μας δίνει την άγνωστη ποσότητα . Για να βρούμε λοιπόν μια τέτοια εξίσωση για τη $v(\cdot)$ εξετάζουμε προσεκτικά το αρχικό πρόβλημα για να δούμε αν υπάρχει κάποια χαρακτηριστική ιδιότητα που μπορούμε να χρησιμοποιήσουμε για το σκοπό αυτό. Εξετάζοντας το πρόβλημα της εικόνας 1, βλέπουμε οτι ακουλουθώντας την πιο σύντομη διαδρομή απο την πόλη x στην πόλη 8 , πρέπει πρώτα να φύγουμε απο την πόλη x και να πάμε σε κάποια πόλη y (είναι δυνατό η πόλη y να είναι η πόλη προορισμός (8)). Τώρα, μια βασική ιδιότητα της διαδρομής ελάχιστου μήκους από την πόλη x στην πόλη 8 είναι η παρακάτω. Αν η ενδιάμεση πόλη y ανήκει στη διαδρομή *ελάχιστου μήκους από την x στην 8, τότε και το κομμάτι της διαδρομής από την y στην 8 είναι επίσης μια διαδρομή ελάχιστου μήκους για το πρόβλημα με αρχική πόλη την y και πόλη προορισμό την 8, όπως φαίνεται με το εξής επιχείρημα.*

Εστω $((x,y),\delta_1)$ μια διαδρομή ελάχιστου μήκους απο την πόλη x στην πόλη 8. Τότε η δ_1 πρέπει να είναι μια διαδρομή ελάχιστου μήκους από την πόλη y στην πόλη 8. Πραγματικά, αν υπήρχε μια διαδρομή δ_2 από την y στην 8 με μήκος μικρότερο από αυτό της δ_1, τότε η $((x,y),\delta_2)$ θα ήταν μια διαδρομή από τη x στην 8 με μήκος μικρότερο από αυτό της $((x,y),\delta_1)$, και αυτό είναι αντίφαση στην αρχική παραδοχή ότι η διαδρομή $((x,y),\delta_1)$ είναι διαδρομή ελάχιστου μήκους.

Η παραπάνω ιδιότητα του προβλήματος συνεπάγεται το εξής σύστημα εξισώσεων για την $v(\cdot)$

$$v(x) = \min_{(x,y) \in D(x)} \Big\{ c(x,y) + v(y) \Big\}, \quad x = 1,\ldots,7, \tag{1}$$

$$v(8) = 0, \tag{2}$$

όπου στην (1) το $D(x)$ συμβολίζει το σύνολο των βελών (δρόμων) που έχουν την πόλη x σαν αρχική πόλη, π.χ. $D(1) = \{(1,2), (1,3), (1,4)\}$, $D(5) = \{(1,8)\}$ κ.λ.π..

Τονίζουμε, λοιπόν, οτι η (1) εκφράζει το εξής φαινόμενο. Το μήκος $v(x)$ μιας διαδρομής ελάχιστου μήκους απο την πόλη x στην πόλη 8, αποτελείται από δύο μέρη: το μήκος από την x σε κάποια γειτονική πόλη y ($c(x,y) \in D(x)$) και το ***μήκος της διαδρομής ελάχιστου μήκους από την πόλη y στην πόλη 8****, που είναι ίσο με την τιμή $v(y)$. Η "βέλτιστη" γειτονική πόλη $y^* = y^*(x)$ πρέπει να επιλεγεί έτσι ώστε το άθροισμα των δύο μερών να είναι ελάχιστο για κάθε πόλη $x = 1,2,\ldots,8$. Η τελευταία αυτή παρατήρηση μας δίδει την απάντηση στην αρχική ερώτηση του προσδιορισμού της (των) διαδρομής (διαδρομών) ελάχιστου μήκους.*

Αν σε κάθε πόλη x διαλέγουμε την επομένη πόλη y^* έτσι ώστε

$$c(x,y^*) + v(y^*) = v(x) = \min_{(x,y) \in D(x)} \Big\{ c(x,y) + v(y) \Big\}, \tag{3}$$

τότε η διαδρομή $(1, y^*(1), y^*(y^*(1)), \ldots, 8)$ θα είναι μια διαδρομή ελάχιστου μήκους.

Ο υπολογισμός της $v(\cdot)$ μέσω της (1) μπορεί να γίνει ως εξής.

$$\begin{aligned} v(8) &= 0 \\ v(7) &= 3 + v(8) = 3, & \qquad y^* &= y^*(7) = 8, \\ v(6) &= 4 + v(8) = 4, & \qquad y^* &= y^*(6) = 8, \\ v(5) &= 7 + v(8) = 7, & \qquad y^* &= y^*(5) = 8, \end{aligned}$$

$$v(4) = \min \begin{cases} 5 + v(5) = 5 + 7 \\ 6 + v(6) = 6 + 4 \end{cases} = 10, \qquad y^* = y^*(4) = 6,$$

$$v(3) = \min \begin{cases} 1 + v(5) = 5 + 7 \\ 1 + v(6) = 1 + 4 \\ 2 + v(7) = 2 + 3 \end{cases} = 5, \qquad y^* = y^*(3) = 6 \text{ ή } 7,$$

$$v(2) = \min \begin{cases} 1 + v(6) = 1 + 4 \\ 5 + v(7) = 5 + 3 \end{cases} = 5, \qquad y^* = y^*(2) = 6,$$

$$v(1) = \min \begin{cases} 1 + v(4) = 1 + 10 \\ 2 + v(3) = 2 + 5 \\ 3 + v(2) = 3 + 5 \end{cases} = 7 \qquad y^* = y^*(1) = 3 .$$

Χρησιμοποιώντας, λοιπόν, την εξίσωση (1) μπορέσαμε να υπολογίσουμε το μήκος της διαδρομής ελάχιστου μήκους από την πόλη 1 στην πόλη 8, και βρήκαμε ότι v(1) = 7.

Παρατηρούμε όμως, ότι έχουμε ταυτόχρονα υπολογίσει και τη διαδρομή (ή διαδρομές) ελάχιστου μήκους απο την πόλη 1 στην πόλη 8. Συγκεκριμένα, οι διαδρομές (1,3,6,8) και (1,3,7,8) είναι οι μόνες διαδρομές ελάχιστου μήκους.

*<u>1.3 Παρατηρήσεις.</u> 1. Σύμφωνα με την "εξίσωση" (1) υπολογίσαμε τη τιμή v(1) ξεκινώντας από το v(8) = 0 (**τέλος** του προορισμού μας) και **πηγαίνοντας προς τα πίσω** μέχρι που φτάσαμε στο να υπολογίσουμε το v(1). Αυτός ο τρόπος σκέψης, που έκφρασή του είναι η εξίσωση (1), είναι γνωστός σαν η **ανάλυση** του προβλήματος. Για να βρούμε τις διαδρομές ελάχιστου μήκους δεν έχουμε παρά να θυμηθούμε τα βήματα μας όταν πήγαμε προς τα πίσω, πράγμα που στο συγκεκριμένο παράδειγμα εκάμαμε όταν γράφαμε στα διάφορα στάδια της προς τα πίσω διαδικασίας τις τιμές των $y^* = y^*(x)$. Αυτή η **προς τα εμπρός** διαδικασία, που γίνεται αφού πρώτα έχουμε τελειώσει την **ανάλυση** (δηλ. την προς τα πίσω διαδικασία), είναι γνωστή σαν η **σύνθεση** του προβλήματος.*

***2.** Υπάρχουν δύο τρόποι για να περιγράψουμε μια διαδρομή ελάχιστου μήκους (λύση του προβλήματος). Ο πρώτος τρόπος είναι το να δώσουμε μια λίστα που να περιέχει όλες τις πόλεις που περιλαμβάνονται στη διαδρομή, π.χ. η διαδρομή (1,3,6,8). Ο δεύτερος τρόπος είναι το να δώσουμε ένα **κανόνα** ή **πολιτική** που να μας λέει ποιό δρόμο να ακολουθήσουμε σε κάθε πόλη (κόμβο) του δικτύου. Δηλαδή αν στο παρόν στάδιο (βήμα) βρισκόμαστε στην πόλη x, στο επόμενο στάδιο πρέπει να ακολουθήσουμε το δρόμο $\alpha^*(x) = (x,y^*(x))$, όπου η $\alpha^*(\cdot)$ (ισοδύναμα η $y^*(\cdot)$) είναι μια γνωστή συνάρτηση, η **πολιτική** μας. Ο πρώτος τρόπος είναι ο συνήθης τρόπος της Μαθηματικής Ανάλυσης, ενώ ο δεύτερος είναι ο συνήθης στη γλώσσα του Δυναμικού Προγραμματισμού.*

***3.** Είναι σημαντικό να προσέξουμε στο παράδειγμα μας ότι, ακολουθώντας την καλύτερη διαδρομή απο την πόλη 1 στην πόλη 8, δεν σημαίνει ότι διαλέγουμε (αναγκαστικά) σε ποιά πόλη θα πάμε στο επόμενο στάδιο, σύμφωνα με το κατά πόσο το μήκος του αμέσως επομένου βήματος είναι το ελάχιστο δυνατό. Πολιτικές πού επιλέγουν την πόλη πού θα πάμε στο επόμενο στάδιο, σύμφωνα με το κατά πόσο το μήκος του αμέσως επομένου βήματος είναι το ελάχιστο δυνατό, λέγονται **μυωπικές** ή πολιτικές που **βλέπουν το επόμενο βήμα μόνο** (myopic ή one step look ahead policies). Στο παραδειγμά μας, στην πόλη 1 η μυωπική πολιτική ορίζει ότι πρέπει να πάμε στην πόλη 4 και αυτή η ενέργεια έχει "**άμεσο κόστος**" (μήκος) 1, ενώ όπως έχουμε ήδη υπολογίσει, για να ακολουθήσουμε μια **βέλτιστη** (ελάχιστου μήκους) διαδρομή απο την πόλη 1, πρέπει να πάμε στην πόλη 3 με άμεσο κόστος 2 > 1.*

***4.** Παρατηρούμε οτι υπάρχει μια μόνο συνάρτηση $v(\cdot)$ που ικανοποιεί τις (1), δηλαδή το σύστημα των εξισώσεων (1) έχει μοναδική λύση. Υπάρχουν όμως 2 διαφορετικές βέλτιστες διαδρομές–πολιτικές.*

*5. Το πρόβλημα της Εικόνας 1. αποτελεί την πιο απλή μορφή μιας γενικότερης κατηγορίας προβλημάτων αποφάσεων, όπου οι αποφάσεις πρέπει να παρθούν **διαδοχικά** (εξακολουθητικά), ενώ υπάρχει ένα **μέτρο απόδοσης** (κόστος ή όφελος) που είναι συνέπεια της όλης ακολουθίας των αποφάσεων (πολιτικής) και όχι μερικών μεμονωμένων βημάτων.*

6. Οι δυνατές αποφάσεις σε κάθε πόλη x δίδονται από το σύνολο D(x) των δυνατών βελών (δρόμων). Ενας ισοδύναμος τρόπος για να περιγράψουμε τις δυνατές αποφάσεις σε κάθε πόλη x, είναι να δώσουμε το σύνολο $\overline{D}(x)$ των πόλεων y για τις οποίες υπάρχει ένα βέλος (δρόμος) $(x,y) \in D(x)$.

Ετσι για το πρόβλημα της εικόνας 1 έχουμε $D(1) = \{(1,2),\ (1,3),\ (1,4)\}$ *και* $\bar{D}(1) = \{2,3,4\}$. *Είναι προφανές ότι τα σύνολα* $D(x)$ *ορίζουν τα σύνολα* $\bar{D}(x)$ *και αντίστροφα. Στη συνέχεια δεν θα κάνουμε διάκριση μεταξύ των* $\bar{D}(x)$ *και* $D(x)$ *εκτός αν υπάρχει κίνδυνος σύγχυσης.*

7. Μια πολιτική είναι ένας πλήρης κανόνας λήψης αποφάσεων, δηλαδή μια συνάρτηση $\alpha(\cdot)$ *απο το σύνολο των κόμβων (πόλεων)* $S = \{1,2,...,8\}$ *στο σύνολο των βελών (δρόμων)* $D = \bigcup_{x\in S} D(x)$ *(ή ισοδύναμα στο σύνολο των "γειτονικών κόμβων"* $\bar{D} = \bigcup_{x\in S} \bar{D}(x)$*), έτσι ώστε* $\alpha(x)$ *συμβολίζει την απόφαση που η πολιτική* $\alpha(\cdot)$ *ορίζει για το κόμβο* x. *Οταν θέλουμε να τονίσουμε την διαφορά της πολιτικής σαν συνάρτησης στο* D *εν αντιθέσει με τη θεώρησή της σαν συνάρτησης στο* $\bar{D}$, *θα γράφουμε* $\alpha(x) = (x, \bar{\alpha}(x))$, $x \in S$. *Το σύνολο* S *ονομάζεται* ***χώρος καταστάσεων*** *ενώ το σύνολο* D *(αντίστοιχα* $\bar{D}$*)* ***χώρος αποφάσεων.***

2. Μέθοδοι Λύσης της Εξίσωσης Βελτιστοποίησης.

Αρχίζουμε με την εισαγωγή πιο ισχυρού συμβολισμού. Εχουμε θεωρήσει την $v(\cdot)$ σαν συνάρτηση της μορφής $v : S \to \mathbb{R}$, όπου $v(8) = 0$. Εστω $\mathfrak{F}(S) := \{u/u : S \to \mathbb{R}, u(8) = 0, |u(x)| < \infty\ \forall x \in S\}$ και έστω $\alpha \in D(x)$. Ορίζουμε τις απεικονίσεις T_α και T ως εξής

$$T_\alpha : \mathfrak{F}(S) \to \mathfrak{F}(S) : (T_\alpha u)(x) := c(x,\alpha) + u(\overline{\alpha}) , \tag{1}$$

$$(T_\alpha u)(8) = 0 , \tag{2}$$

$$T : \mathfrak{F}(S) \to \mathfrak{F}(S) : (Tu)(x) := \min_{(x,y) \in D(x)} \left\{ c(x,y) + u(y) \right\} , \tag{3}$$

$$(Tu)(8) = 0. \tag{4}$$

Με την βοήθεια των απεικονίσεων (τελεστών) T, T_α το σύστημα των βασικών εξισώσεων μπορεί να γραφεί πιο οικονομικά ως εξής

$$v(x) = (Tv)(x) = \min_{\alpha \in D(x)} \left\{ (T_\alpha v)(x) \right\} . \tag{5}$$

ή ισοδύναμα

$$v = Tv . \tag{6}$$

Η συνάρτηση v είναι, λοιπόν, το σταθερό σημείο v τού τελεστή T. Η (6) είναι μια εξίσωση ως πρός την άγνωστη συνάρτηση v (συναρτησιακή εξίσωση). Στη συνέχεια όταν αναφερόμαστε στη εξίσωση βελτιστοποίησης θα εννοούμε μια εξίσωση της μορφής (6) ή, ισοδύναμα, της μορφής (1.1.1). Υποθέτουμε ότι στην (3) το min έχει νόημα, όπως π.χ. συμβαίνει όταν οι χώροι S και $D(x)$ είναι πεπερασμένοι ή συμπαγή σύνολα, και οι u είναι συνεχείς συναρτήσεις.

Αντικείμενα της γενικής θεωρίας του Δυναμικού Προγραμματισμού είναι τα εξής.

1. *Η διατύπωση προβλημάτων αποφάσεων (βελτιστοποίησης) μέσω εξισώσεων της μορφής (6) για κατάλληλο τελεστή T.*
2. *Ερωτήσεις ύπαρξης και μοναδικότητας της λύσης εξισώσεων της μορφής (6)*
3. *Η μελέτη ιδιοτήτων της λύσης εξισώσεων της μορφής (6).*
4. *Εύρεση μεθόδων (αλγορίθμων) υπολογισμού της λύσης.*

Είναι ενδιαφέρον ότι όλες οι μέθοδοι λύσης του γενικού προβλήματος Δυναμικού Προγραμματισμού, για προβλήματα με ***πεπερασμένους*** *χώρους καταστάσεων S και αποφάσεων D, μπορούν να χρησιμοποιηθούν για τη λύση του προβλήματος της παραγράφου 1. Ας δούμε λοιπόν αυτές τις μεθόδους αμέσως.*

2.1. Μέθοδος των Διαδοχικών Προσεγγίσεων στο Χώρο Τιμών.

Η κεντρική ιδέα είναι η ακόλουθη. Εστω μια συνάρτηση $v^{(0)}(\cdot)$ στο $\mathfrak{F}(S)$, η ***αρχική προσέγγιση*** *για τη $v(\cdot)$. Μπορούμε να δημιουργήσουμε μια ακολουθία συναρτήσεων $\{v^{(n)}(\cdot),\ n \geq 1\}$, ως εξής*

$$v^{(n)} = Tv^{(n-1)} \tag{1}$$

ή ισοδύναμα

$$v^{(n)}(x) := \min_{(x,y) \in D(x)} \{ c(x,y) + v^{(n-1)}(y) \} \tag{2}$$

Οι συναρτήσεις $v^{(n)}(\cdot)$ είναι οι διαδοχικές προσεγγίσεις της $v(\cdot)$ και, κάτω απο κατάλληλες συνθήκες που θα δούμε σε λεπτομέρια αργότερα, έχουμε

$$v = \lim_{n \to \infty} v^{(n)} \tag{3}$$

Η μέθοδος των διαδοχικών προσεγγίσεων χρησιμοποιείται σε πολλούς κλάδους των Εφαρμοσμένων Μαθηματικών, π.χ. στη λύση συστημάτων αλγεβρικών ή διαφορικών εξισώσεων.

Για το πρόβλημα της παραγράφου 1.1 ας πάρουμε $v^{(0)}(x) = 0, \ \forall x \in S$. *Τότε*

1η προσέγγιση.

$v^{(1)}(8) = 0$,

$v^{(1)}(7) = 3 + v^{(0)}(8) = 3$, $\bar{\alpha}^{(1)}(7) = 8$,

$v^{(1)}(6) = 4 + v^{(0)}(8) = 4$, $\bar{\alpha}^{(1)}(6) = 8$,

$v^{(1)}(5) = 7 + v^{(0)}(8) = 7$, $\bar{\alpha}^{(1)}(5) = 8$,

$v^{(1)}(4) = \min\left\{5 + v^{(0)}(5),\ 6 + v^{(0)}(6)\right\} = 5$, $\alpha^{(1)}(4) = 5$,

$v^{(1)}(3) = \min\left\{1 + v^{(0)}(5),\ 1 + v^{(0)}(6),\ 2 + v^{(0)}(7)\right\} = 1$, $\alpha^{(1)}(4) = 5$ ή 6,

$v^{(1)}(2) = \min\left\{1 + v^{(0)}(6),\ 5 + v^{(0)}(7)\right\} = 1$, $\alpha^{(1)}(2) = 6$,

$v^{(1)}(1) = \min\left\{1 + v^{(0)}(4),\ 2 + v^{(0)}(3),\ 3 + v^{(0)}(2)\right\} = 1$, $\alpha^{(1)}(1) = 1$.

2η προσέγγιση.

$v^{(2)}(8) = 0$,

$v^{(2)}(7) = 3 + v^{(1)}(8) = 3$, $\bar{\alpha}^{(2)}(7) = 8$,

$v^{(2)}(6) = 4 + v^{(1)}(8) = 4$, $\bar{\alpha}^{(2)}(6) = 8$,

$v^{(2)}(5) = 7 + v^{(1)}(8) = 7$, $\bar{\alpha}^{(2)}(5) = 8$,

$v^{(2)}(4) = \min\left\{5 + v^{(1)}(5),\ 6 + v^{(1)}(6)\right\} = 10$, $\alpha^{(2)}(4) = 6$,

$v^{(2)}(3) = \min\left\{1 + v^{(1)}(5),\ 1 + v^{(1)}(6),\ 2 + v^{(1)}(7)\right\} = 5$, $\alpha^{(2)}(4) = 6$ ή 7,

$v^{(2)}(2) = \min\left\{1 + v^{(1)}(6),\ 5 + v^{(1)}(7)\right\} = 5$, $\alpha^{(2)}(2) = 6$,

$v^{(2)}(1) = \min\left\{1 + v^{(1)}(4),\ 2 + v^{(1)}(3),\ 3 + v^{(1)}(2)\right\} = 3$, $\alpha^{(2)}(1) = 3$.

3η προσέγγιση.

$v^{(3)}(8) = 0$,

$v^{(3)}(7) = 3 + v^{(2)}(8) = 3$, $\bar{\alpha}^{(3)}(7) = 8$,

$v^{(3)}(6) = 4 + v^{(2)}(8) = 4$, $\bar{\alpha}^{(3)}(6) = 8$,

$v^{(3)}(5) = 7 + v^{(2)}(8) = 7$, $\bar{\alpha}^{(3)}(5) = 8$,

$$v^{(3)}(4) = min\left\{5 + v^{(2)}(5)\,,\ \ 6 + v^{(2)}(6)\right\} = 10\,, \qquad \alpha^{(3)}(4) = 6\,,$$

$$v^{(3)}(3) = min\left\{1 + v^{(2)}(5)\,,\ 1 + v^{(2)}(6)\,,\ \ 2 + v^{(2)}(7)\right\} = 5\,, \qquad \alpha^{(3)}(4) = 6 \text{ ή } 7\,,$$

$$v^{(3)}(2) = min\left\{ 1 + v^{(2)}(6)\,,\ \ 5 + v^{(2)}(7)\right\} = 5\,, \qquad \alpha^{(3)}(2) = 6\,,$$

$$v^{(3)}(1) = min\left\{1 + v^{(2)}(4)\,,\ 2 + v^{(2)}(3)\,,\ \ 3 + v^{(2)}(2)\right\} = 7\,, \qquad \alpha^{(3)}(1) = 3\,.$$

!!FOOTNOTE 1:$Tv^{(3)}(x) = Tv^{(2)}(x)$, γιατί $v^{(3)}(x) = v^{(2)}(x)\ \forall\, x$.

Αν συνεχίσουμε με τον ίδιο τρόπο[1] και στην τέταρτη προσέγγιση, θα βρούμε ότι $v^{(4)}(x) = v^{(3)}(x)\ \ \forall\ x = 1,2,\ldots,8$, *άρα* $v^{(5)}(x) = Tv^{(4)}(x) = Tv^{(4)}(x) = v^{(4)}(x)$, *και με τον ίδιο τρόπο παίρνουμε ότι* $v^{(n)} = v^{(3)}\ \forall n \geq 4$. *Αρα,* $\exists \lim_{n\to\infty} v^{(n)} = v^{(3)} = v$.

Οι παραπάνω υπολογισμοί μπορούν να παρουσιαστούν περιληπτικά ως εξής

	$v^{(1)}(\cdot)$	$\overline{\alpha}^{(1)}(\cdot)$	$v^{(2)}(\cdot)$	$\overline{\alpha}^{(2)}(\cdot)$	$v^{(3)}(\cdot)$	$\overline{\alpha}^{(3)}(\cdot)$	$v^{(4)}(\cdot)$	$\overline{\alpha}^{(4)}(\cdot)$
8	0	—	0	—	0	—	0	—
7	3	8	3	8	3	8	3	8
6	4	8	4	8	4	8	4	8
5	7	8	7	8	7	8	7	8
4	5	5	10	6	10	6	10	6
3	1	5,6	5	6,7	5	6,7	5	6,7
2	1	6	5	6	5	6	5	6
1	1	4	3	3	7	3	7	3

Πίνακας 1

Παρατηρήστε ότι η πολιτική $\overline{\alpha}^{(1)}(\cdot)$ *είναι η <u>μυωπική</u> πολιτική που δεν είναι, βέβαια, η βέλτιστη.*

2.2. Μέθοδος των Διαδοχικών Προσεγγίσεων στο Χώρο των Πολιτικών. *Οπως είδαμε στην προηγούμενη παράγραφο, η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο τιμών (Δ.Π.Τ. για συντομία) στηρίζεται στην ιδέα ότι ξεκινώντας από# μια αρχική (και αυθαίρετη) προσέγγιση για την* $v(\cdot)$, *μπορούμε να δημιουργήσουμε, χρησιμοποιώντας την εξίσωση βελτιστοποίησης (3) , μια ακολουθία διαδοχικών προσεγγίσεων* $v^{(n)}(\cdot)$ *που συγκλίνει στη* $v(\cdot)$. *Σε κάθε βήμα καταγράφουμε τις αντίστοιχες προσεγγίσεις* $\alpha^{(n)}(\cdot)$ *για τις βέλτιστες αποφάσεις. Η μέθοδος των* ***διαδοχικών προσεγγίσεων στο χώρο των πολιτικών*** *(Δ.Π.Π. για συντομία) βασίζεται σε ανάλογη ιδέα. Τώρα ξεκινάμε από μια αρχική προσέγγιση (που μπορεί να είναι αυθαίρετη) για τη βέλτιστη πολιτική,* $\alpha^{(0)}(x)$. *Δεδομένης της* $\alpha^{(0)}(x)$, *μπορούμε να υπολογίσουμε τα μήκη των διαδρομών που ορίζονται απο αυτή την πολιτική για όλες τις θέσεις (πόλεις)* $x = 1,2,\ldots,7$. *Εστω* $w^{(0)}(x)$ *οι αντίστοιχες τιμές. Είναι εύκολο να δούμε ότι η βασική εξίσωση βελτιστοποίησης για το πρόβλημα (1.1) έχει μια μόνο λύση (γιατί ;) . Αρα, αν η* $w^{(0)}(\cdot)$ *ικανοποιεί την βασική εξίσωση βελτιστοποίησης , τότε έχουμε ότι* $w^{(0)}(x) = v(x)\ \forall x$, *και επομένως η* $\alpha^{(0)}(x)$ *είναι μια βέλτιστη πολιτική. Διαφορετικά, αν η* $w^{(0)}(\cdot)$ *δεν ικανοποιεί τη βασική εξίσωση για κάποια απόφαση (πόλη), τότε αλλάζουμε την απόφαση για τη συγκεκριμένη πόλη και έτσι παίρνουμε μια νέα πολιτική* $\alpha^{(1)}(\cdot)$. *Υπολογίζουμε τις τιμές* $w^{(1)}(x)$ *που αντιστοιχούν στην* $\alpha^{(1)}(x)$ *και, όπως θα δούμε σε λεπτομέρεια αργότερα ,* $w^{(1)}(x) \leq w^{(0)}(x)\ , \forall x$. *Επαναλαμβάνουμε τα παραπάνω βήματα, και σε κάθε επανάληψη βρίσκουμε μια καλύτερη πολιτική (και την αντίστοιχη συνάρτηση τιμής της). Επειδή υπάρχει ένας πεπερασμένος αριθμός πολιτικών, όπως θα δούμε αργότερα, η μέθοδος Δ.Π.Π. θα μας δώσει τη λύση της βασικής εξίσωσης σε πεπερασμένο αριθμό βημάτων.*

Για το πρόβλημα (1.1) ας ξεκινήσουμε με την αυθαίρετη αρχική πολιτική $\overline{\alpha}^{(0)}(\cdot)$, *όπου:*

$$\overline{\alpha}^{(0)}(x) = 8\ , x = 5,6,7\ , \quad \overline{\alpha}^{(0)}(4) = 5\ , \quad \overline{\alpha}^{(0)}(3) = 5\ , \quad \overline{\alpha}^{(0)}(2) = 7\ , \overline{\alpha}^{(0)}(1) = 2.$$

1

Η $\bar{\alpha}^{(0)}(x)$ *είναι μια* ***νόμιμη*** *πολιτική γιατί* $\bar{\alpha}(x) \in \bar{D}(x)\ \forall\, x$.

Οι διαδοχικές προσεγγίσεις σύμφωνα με τη μέθοδο Δ.Π.Π. δημιουργούνται ως εξής.

Αρχική Προσέγγιση.

x	8	7	6	5	4	3	2	1
$\bar{\alpha}^{(0)}(x)$	—	8	8	8	5	5	7	2

Βήμα 1. *(Υπολογισμός της τιμής* $w^{(0)}(\cdot)$*)*

Οι τιμές $w^{(0)}(x)$ *υπολογίζονται με τη βοήθεια των παρακάτω σχέσεων:*

$$w^{(0)}(x) = T_\alpha\, w^{(0)}(x) = c(x,\bar{\alpha}^{(0)}(x)) + w^{(0)}(\bar{\alpha}^{(0)}(x)) \ ,\ x \in S \tag{1}$$

όπου

$$w^{(0)}(8) = 0, \tag{2}$$

οπότε βρίσκουμε

x	8	7	6	5	4	3	2	1
$w^{(0)}(x)$	0	3	4	7	12	8	8	11

Βήμα 2. *(Ελεγχος για βελτίωση της* $\bar{\alpha}^{0}(\cdot)$*). Η* $\bar{\alpha}^{(0)}$ *είναι βέλτιστη πολιτική αν*

$$w^{(0)}(4) \le 6 + w^{(0)}(6) \quad \Leftrightarrow\ 12 \le 6+4 \quad \Leftrightarrow\ \Delta w^{(0)}(4,6) = 10-12 = -2 < 0$$

$$w^{(0)}(3) \le 1 + w^{(0)}(6) \quad \Leftrightarrow\ 8 \le 1+4 \quad \Leftrightarrow\ \Delta w^{(0)}(3,6) = 5-8 = -3 < 0$$

$$w^{(0)}(3) \le 2 + w^{(0)}(7) \quad \Leftrightarrow\ 8 \le 2+3 \quad \Leftrightarrow\ \Delta w^{(0)}(3,7) = 5-8 = -3 < 0$$

$$w^{(0)}(2) \le 1 + w^{(0)}(6) \quad \Leftrightarrow\ 8 \le 1+4 \quad \Leftrightarrow\ \Delta w^{(0)}(2,6) = 5-8 = -3 < 0$$

$$w^{(0)}(1) \le 1 + w^{(0)}(4) \quad \Leftrightarrow\ 11 \le 1+12 \quad \Leftrightarrow\ \Delta w^{(0)}(1,4) = 13-11 = 2 > 0$$

$$w^{(0)}(1) \le 2 + w^{(0)}(3) \quad \Leftrightarrow\ 11 \le 2+8 \quad \Leftrightarrow\ \Delta w^{(0)}(1,3) = 10-11 = -1 < 0$$

Παρατηρούμε ότι η βασική εξίσωση (1.1) (ή (1.2)) παραβιάζεται σε όλες τις καταστάσεις (πόλεις) εκτός απο τη κατάσταση (πόλη) 1 για την απόφαση (1,4) όπου $\Delta w^{(0)}(1,4) > 0$ *, (δηλ.* $w^{(0)}(1) < 1 + w^{(0)}(4)$*). Αρα η πολιτική* $\bar{\alpha}^{(0)}(\cdot)$ *δεν είναι μια βέλτιστη πολιτική.*

Παίρνουμε λοιπόν σαν νέα προσέγγιση της βέλτιστης πολιτικής την $\bar{\alpha}^{(0)}(x)$ *, διαλέγοντας σαν* $(x,\bar{\alpha}^{(1)}(x)) \in D(x)$ *, για μια τουλάχιστον κατάσταση* x *, εκείνη την απόφαση* (x,y) *για την οποία*

$$\Delta w^{(0)}(x,y) = \min\{\Delta w^{(0)}(x,y) \ :\ (x,y) \in D(x)\ ,\ \Delta w^{(0)}(x,y) < 0\}.$$

Εστω $\bar{\alpha}^{(1)}(x)$ *μια πολιτική που παίρνουμε με το παραπάνω τρόπο. Συγκεκριμένα, διαλέγουμε*

x	8	7	6	5	4	3	2	1
$\bar{\alpha}^{(1)}(x)$	—	8	8	8	6	7	6	3

2η Προσέγγιση.

Βήμα 1. *(Υπολογισμός της τιμής $w^{(1)}(x)$)*

x	8	7	6	5	4	3	2	1
$w^{(1)}(x)$	0	3	4	7	10	5	5	7

Τα στοιχεία του παραπάνω πίνακα έχουν υπολογιστεί σύμφωνα με τις σχέσεις

$$w^{(1)}(x) = c(x, \overline{\alpha}^{(1)}(x)) + w^{(1)}(\overline{\alpha}^{(1)}(x)) \tag{3}$$

$$w^{(1)}(8) = 0. \tag{4}$$

Παρατηρούμε επίσης ότι $w^{(1)}(x) \le w^{(0)}(x)\ \forall x$, και μάλιστα $w^{(1)}(x) < w^{(0)}(x)$ για μερικά x, (π.χ. $x = 1,2,3,4$). Λέμε λοιπόν ότι η $\overline{\alpha}^{(1)}(\cdot)$ είναι μια βελτίωση της πολιτικής $\overline{\alpha}^{(0)}(\cdot)$.

Βήμα 2. *(Ελεγχος για βελτίωση της $\alpha^{(1)}$).*

$$\Delta w^{(1)}(4,5) = 5 + w^{(1)}(5) - w^{(1)}(4) = 12 - 10 = 2 > 0,$$

$$\Delta w^{(1)}(3,5) = 1 + w^{(1)}(5) - w^{(1)}(3) = 8 - 5 = 3 > 0,$$

$$\Delta w^{(1)}(3,6) = 1 + w^{(1)}(6) - w^{(1)}(3) = 5 - 5 = 0,$$

$$\Delta w^{(1)}(2,7) = 5 + w^{(1)}(7) - w^{(1)}(2) = 8 - 5 = 3 > 0,$$

$$\Delta w^{(1)}(1,4) = 1 + w^{(1)}(4) - w^{(1)}(1) = 11 - 7 = 4 > 0,$$

$$\Delta w^{(1)}(1,2) = 3 + w^{(1)}(2) - w^{(1)}(3) = 8 - 7 = 1 > 0.$$

Εχουμε λοιπόν

$$\Delta w^{(1)}(x,\overline{\alpha}^{(1)}(x)) = c(x,\overline{\alpha}^{(1)}(x)) + w^{(1)}(\overline{\alpha}^{(1)}(x)) - w^{(1)}(x) \ge 0 \quad \forall x, \tag{5}$$

άρα η $\overline{\alpha}^{(1)}(\cdot)$ ορίζει μια βέλτιστη πολιτική. Επιπλέον, επειδή $\Delta w^{(1)}(3,6) = 0$ και η πολιτικη $\overline{\alpha}^{(1)}(\cdot)$ είναι βέλτιστη. Η πολιτική $\overline{\alpha}^{(2)}(\cdot)$, που διαφέρει απο την $\overline{\alpha}^{(1)}(\cdot)$ μόνο κατά την απόφαση στην κατάσταση (πόλη) 3, όπου $\overline{\alpha}^{(2)}(3) = 6$, είναι επίσης μια βέλτιστη πολιτική. Πραγματικά, αυτό μπορεί να επιβεβαιωθεί (θα αποδειχθεί αργότερα) αν εκτελέσουμε το βήμα 1 του παραπάνω αλγορίθμου.

Οι παραπάνω υπολογισμοί μπορούν να παρουσιασθούν συνοπτικά όπως στον πίνακα 2. Στον πίνακα 2 χρησιμοποιούμε τον συμβολισμό $\alpha(x) = (x, \overline{\alpha}(x))$, για τις αποφάσεις στις διάφορες πόλεις x. Μπορούσαμε βέβαια να χρησιμοποιήσουμε τον συμβολισμό $\overline{\alpha}(x)$.

Παρατηρούμε ότι ο υπολογισμός των $w^{(n)}(x)$ γίνεται πιο εύκολα αν αρχίσουμε απο το ***τέλος***, *π.χ. υπολογίζουμε πρώτα το $w^{(0)}(8)$ μετά το $w^{(0)}(7)$ κ.λ.π.*

x	$D(x)$	$c(x,y)$	$\alpha^{(0)}(\cdot)$	$w^{(0)}(\cdot)$	$\Delta w^{(0)}(\cdot)$	$\alpha^{(1)}(\cdot)$	$w^{(1)}(\cdot)$	$\Delta w^{(1)}(\cdot)$
1	(1,2)	3	(1,2)	11	0			1
	(1,3)	2			–1	(1,3)	7	0
	(1,4)	1			2			4
2	(2,6)	1			–3	(2,6)	5	0
	(2,7)	5	(2,7)	8	0			3
3	(3,5)	1	(3,5)	8	0			3
	(3,6)	1			–3			0
	(3,7)	2			–3	(3,7)	5	0

4	(4,5)	5	(4,5)	12	0			2
	(4,6)	6			–2	(4,6)	10	0
5	(5,8)	7	(5,8)	7	0	(5,8)	7	0
6	(6,8)	4	(6,8)	4	0	(6,8)	4	0
7	(7,8)	3	(7,8)	3	0	(7,8)	3	0
8	—	—	—	0	0	—	0	0

Πίνακας 2

2.3. Μέθοδος Γραμμικού Προγραμματισμού. *Είναι ενδιαφέρον ότι το πρόβλημα 1.1 μπορεί να λυθεί με τη μέθοδο του γραμμικού προγραμματισμού. Τούτο οφείλεται σε μια βασική ιδιότητα των κυρτών συνόλων όπως θα δούμε αργότερα. Προς το παρόν, απλώς δίνουμε το γραμμικό πρόγραμμα για το πρόβλημα.*

Εστω $\lambda(x)$ *αυθαίρετοι θετικοί αριθμοί* $x = 1,2,\ldots,8$ (*π.χ.* $\lambda(x) = 1\ \forall x$). *Θεωρούμε το παρακάτω γραμμικό πρόγραμμα*

$$(\max)\ z = \sum_{x=1}^{8} \lambda(x)\,\xi(x)$$

με περιορισμούς

$$\begin{array}{llllllllllll}
(1,2) & \xi(1) & -\xi(2) & & & & & & & \leq & 3 \\
(1,3) & \xi(1) & & -\xi(3) & & & & & & \leq & 2 \\
(1,4) & \xi(1) & & & -\xi(4) & & & & & \leq & 1 \\
(2,6) & & \xi(2) & & & & -\xi(6) & & & \leq & 1 \\
(2,7) & & \xi(2) & & & & & -\xi(7) & & \leq & 5 \\
(3,5) & & & \xi(3) & & -\xi(5) & & & & \leq & 1 \\
(3,6) & & & \xi(3) & & & -\xi(6) & & & \leq & 1 \\
(3,7) & & & \xi(3) & & & & -\xi(7) & & \leq & 2 \\
(4,5) & & & & \xi(4) & -\xi(5) & & & & \leq & 5 \\
(4,6) & & & & \xi(4) & & -\xi(6) & & & \leq & 6 \\
(5,8) & & & & & \xi(5) & & & -\xi(8) & \leq & 7 \\
(6,8) & & & & & & \xi(6) & & -\xi(8) & \leq & 4 \\
(7,8) & & & & & & & \xi(7) & -\xi(8) & \leq & 3 \\
(8,9) & & & & & & & & \xi(8) & = & 0
\end{array}$$

$$\xi(x) \geq 0,\ x = 1,\ldots,8.$$

Παρατηρούμε ότι το παραπάνω πρόβλημα γραμμικού προγραμματισμού (Π.Γ.Π.) έχει 8 μεταβλητές (ή 7 μεταβλητές, ανάλογα με το πώς προτιμάει κανείς να βλέπει το $\xi(8)$*), που αντιστοιχούν στις πόλεις του δικτύου. Οι περιορισμοί* (x,y), *για* $(x,y) \in D$, *αντιστοιχούν στους δρόμους του δικτύου και εκφράζουν την απαίτηση της βασικής εξίσωσης*

$$\xi(x) \leq c(x,y) + \xi(y). \tag{1}$$

Επίσης υπάρχει ο περιορισμός (8,9) , $\xi(8) = 0$*, που αντιστοιχεί στην απαίτηση* $v(8) = 0$ *και που εκφράζει το γεγονός ότι η πόλη 8 είναι το τέρμα του δικτύου. Λέμε ότι το Π.Γ.Π. στην παραπάνω του μορφή είναι στην* ***πρωταρχική*** *(primal) μορφή.*

Θα αποδείξουμε παρακάτω (§6) ότι αν η βασική εξίσωση (1.1) έχει μια μόνο λύση, τότε το παραπάνω Π.Γ.Π. έχει, επίσης, μια μόνο βέλτιστη λύση: $\xi^* = (\xi^*(1),\xi^*(2),\ldots,\xi^*(8))$*, και μάλιστα ,* $v(x) = \xi^*(x)$ *,* $x = 1,\ldots, 8$*.*

Αν λύσουμε λοιπόν το Π.Γ.Π., έχουμε αμέσως τη συνάρτηση τιμών $v(\cdot)$*. Η βέλτιστη πολιτική* $\alpha^*(\cdot)$ *βρίσκεται στη συνέχεια από τους περιορισμούς που ισχύουν σαν ισότητες για τη βέλτιστη λύση* ξ^**. Συγκεκριμένα, επιλέγουμε* $\overline{\alpha}^*(x) = y^*$*, αν και μόνο αν* $\xi^*(x) - \xi^*(y^*) = c(x,y^*)$*.*

Ας θεωρήσουμε τώρα το δυικό πρόβλημα που αντιστοιχεί στο πρωτεύον της σελίδας 14. Εστω λοιπόν $\zeta(x,y)$ *η δυική μεταβλητή που αντιστοιχεί στον περιορισμό (x,y) του πρωτεύοντος. Το δυικό πρόβλημα είναι*

$$(\min)\ \zeta = \sum_{(x,y)} c(x,y)\, \zeta(x,y)$$

περιορισμοί

(1)	$\zeta(12)$	$+\zeta(13)$	$+\zeta(14)$												$=\lambda(1)$
(2)	$-\zeta(12)$			$+\zeta(26)$	$+\zeta(27)$										$=\lambda(2)$
(3)		$-\zeta(13)$				$+\zeta(35)$	$+\zeta(36)$	$+\zeta(37)$							$=\lambda(3)$
(4)			$-\zeta(14)$						$+\zeta(45)$	$+\zeta(46)$					$=\lambda(4)$
(5)						$-\zeta(35)$			$-\zeta(45)$		$+\zeta(58)$				$=\lambda(5)$
(6)				$-\zeta(26)$			$-\zeta(36)$			$-\zeta(46)$		$+\zeta(68)$			$=\lambda(6)$
(7)					$-\zeta(27)$			$-\zeta(37)$					$+\zeta(78)$		$=\lambda(8)$
(8)											$-\zeta(58)$	$-\zeta(68)$	$-\zeta(78)$	$+\zeta(89)$	$=\lambda(8)$

$\zeta(x,y) \geq 0, \quad (x,y) \in D$

$\zeta(8,9)$ *μη περιορισμένο ως προς το πρόσημο*

Παρατηρούμε ότι μπορούμε να δώσουμε στο δυικό πρόβλημα ερμηνεία, να βρούμε δηλαδή μια ιστορία της οποίας έκφραση είναι το δυικό πρόβλημα.

Ας υποθέσουμε ότι το δίκτυο των δρόμων της εικόνας 1.1 συμβολίζει ένα δίκτυο σωλήνων (π.χ. σε ένα διυλιστήριο) και οι πόλεις συμβολίζουν τους κόμβους του δικτύου. Υπάρχει ένα κόστος $c(x,y)$ *ανά μονάδα για τη μεταφορά ενός υλικού κατά μήκος του σωλήνα (x,y). Τέλος προσθέτουμε ένα επιπλέον κόμβο, τον κόμβο 9, και το σωλήνα (8,9), με κόστος* $c(8,9) = 0$*. Οι σωλήνες έχουν μεγάλη χωρητικότητα (πρακτικά άπειρη). Σε κάθε κόμβο* $x = 1,2,\ldots,8$ *εισρέει απο το περιβάλλον (έξω από το δίκτυο) μια ποσότητα υλικού ίση με* $\lambda(x) > 0$*. Ο σκοπός μας είναι να βρούμε τον πιο φτηνό τρόπο μεταφοράς του υλικού απο τον αρχικό κόμβο 1 στον τελικό κόμβο 9. Σύμφωνα με αυτή την ερμηνεία οι μεταβλητές του δυικού εκφράζουν ποσότητες μεταφερόμενες απο κόμβο σε κόμβο, η αντικειμενική συνάρτηση εκφράζει συνολικό κόστος μεταφοράς, και οι περιορισμοί το γεγονός ότι όσο υλικό εισέρχεται στον κόμβο x τόσο εξέρχεται, για* $x = 1,\ldots,8$*. Ο κόμβος 9 στον οποίο καταλήγει όλο το υλικό ονομάζεται* ***κόμβος–νεροχύτης (sink)*** *ενώ ο κόμβος 1* ***κόμβος - πηγή*** *(source)***.**

<u>2.4. Σχέση Μεθόδων Γ.Π. – Δ.Π.Π.</u> *Ας δούμε προσεκτικά τη σχέση που υπάρχει μεταξύ των μεθόδων του Γ.Π. και Δ.Π.Π. με ένα παράδειγμα. Για να είναι εύκολος ο έλεγχος των πράξεων, διαλέγουμε το πιο απλό πρόβλημα που είναι δυνατό να διατηρήσει τη βασική δομή του προβλήματος 1.1. Ας θεωρήσουμε λοιπόν το πρόβλημα εύρεσης της διαδρομής ελάχιστου μήκους για το δίκτυο της εικόνας 2.*

4 2

$\boxed{2}$

Εικόνα 2 .

!!FOOTNOTE 2: $\Delta w^{(0)}(x) = c(x,y) + w^{(0)}(y) - w^{(0)}(x)$

Η λύση σύμφωνα με τη μέθοδο Δ.Π.Π. δίδεται στον επόμενο πίνακα (όπου αρχίζουμε με την μη βέλτιστη πολιτική $\bar{\alpha}^0(2) = 3)$[2].

x	D(x)	c(x,y)	$\bar{\alpha}^{(0)}(x)$	$w^{(0)}(x)$	$\Delta w^{(0)}(x)$	$\bar{\alpha}^{(0)}(x)$	$w^{(0)}(x)$	$\Delta w^{(0)}(x)$
1	(1,2)	4	—	—	–2	2	6	0
	(1,3)	8	3	8	0	—	—	2
2	(2,3)	2	3	2	0	3	2	0
3	—	—	—	0	0	—	0	0

Πίνακας 3

Η πολιτική $\bar{\alpha}^{(1)}$ είναι βέλτιστη, δεδομένου ότι $\Delta w^{(1)}(x) \geq 0 \ \forall x$.

Το δυικό πρόγραμμα που αντιστοιχεί στο πρόβλημα της εικόνας 2 είναι:

$$(\min) \quad w = 4\zeta(1,2) + 8\zeta(1,3) + 2\zeta(2,3)$$

περιορισμοί:

$$\begin{array}{llllll} (1) & \zeta(1,2) & +\zeta(1,3) & & & = 1 \\ (2) & -\zeta(1,2) & & +\zeta(2,3) & & = 1 \\ (3) & & -\zeta(1,3) & -\zeta(2,3) & +\zeta(3,4) & = 1 \end{array}$$

$\zeta(1,2)\,,\ \zeta(1,3)\,,\ \zeta(2,3) \geq 0$
$\zeta(3,4)$ χωρίς περιορισμό στο πρόσημο.

Ας λύσουμε το παραπάνω πρόβλημα με τη αναθεωρημένη μέθοδο simplex (revised simplex).

Παρατηρούμε ότι μπορούμε να αποφύγουμε τη διαδικασία εύρεσης μιας αρχικής βασικής εφικτής λύσης . Πραγματικά, έστω ότι διαλέγουμε τις βασικές μεταβλητές έτσι, ώστε να έχουμε για κάθε κόμβο μια βασική μεταβλητή. Π.χ. διαλέγουμε σαν σύνολο βασικών μεταβλητών το $\{\zeta(1,3)\,,\ \zeta(2,3)\,,\ \zeta(3,4)\}$ και όχι το $\{\zeta(1,2)\,,\ \zeta(1,3)\,,\ \zeta(3,4)\}$ (που δεν έχει βασική μεταβλητή αντιστοιχούσα στον κόμβο 2). Ο λόγος για τον περιορισμό αυτό στη επιλογή των βασικών μεταβλητών ειναι ο εξής. Ας παρατηρήσουμε ότι οι πίνακες βάσης που αντιστοιχούν στις παραπάνω δύο επιλογές είναι οι εξής

2

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \qquad \kappa\alpha\iota \qquad B' = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

αντίστοιχα. Εχουμε λοιπόν ότι για τον B υπάρχει ο B^{-1} και όλα τα στοιχεία του B^{-1} είναι θετικά, πράγμα αναμενόμενο δεδομένου ότι τα διαγώνια στοιχεία του B είναι όλα θετικά, ενώ τα εκτός διαγωνίου στοιχεία είναι μη θετικά. Το δεξιό μέρος των περιορισμών είναι ένα διάνυσμα b με θετικά στοιχεία, συνεπώς οι τιμές των βασικών μεταβλητών είναι θετικές: ($x_B = B^{-1}b > 0$) , δηλαδή έχουμε μια βασική εφικτή λύση. Το παραπάνω επιχείρημα δεν ισχύει για τον B' , και έτσι η αντίστοιχη βασική λύση δεν είναι εφικτή. Πραγματικά

$$(B')^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \geq \qquad 0$$

<u>Λύση του δυικού γραμμικού προβλήματος.</u> Εστω η αρχική "βάση" (βασικές μεταβλητές) $\{\zeta(1,3),\zeta(2,3),\zeta(3,4)\}$ με αντίστοιχο αρχικό πίνακα βάσης B_0.

$$B_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \qquad \mu\epsilon \qquad B_0^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \kappa\alpha\iota \ \overline{b} = B^{-1}b = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

<u>Βήμα 0.</u> *Υπολογισμός πολλαπλασιαστών Simplex. Εχουμε*

!!FOOTNOTE 3:Μερικές φορές χρησιμοποιείται ο συμβολισμός $c_j - z_j$ για το $\overline{c}_j$.

!!FOOTNOTE 4:P(1,2) συμβολίζει την (1,2) κολώνα του πίνακα περιορισμών P και $(\overline{P}(i,j))_l$ το l στοιχείο του διανύσματος $\overline{P}(i,j)$.

!!FOOTNOTE 5:$(x_B)_r$ συμβολίζει το r στοιχείο του διανύσματος $x_B = \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_m} \end{pmatrix}$

$$\pi = c_B B^{-1} = [\,8\ ,2\ ,0\,] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = [\,8\ ,2\ ,0\,]$$

<u>Βήμα 1.1</u> *(Κριτήριο Βελτίωσης). Υπολογισμός των νέων συντελεστών της αντικειμενικής συνάρτησης για τις μη–βασικές μεταβλητές.*

$$\overline{c}(1,2) = (c_j - z_i)^3 = c(1,2) - \pi P(1,2) = {}^4$$

$$4 — (\,8\ ,2\ ,0\,) \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 4 - 6 = -2 < 0$$

Επειδή το $\overline{c}(1,2)$ είναι αρνητικό, η B_0 δεν είναι βέλτιστη. Η μεταβλητή $\zeta(1,2)$ εισέρχεται στη βάση (θα είναι βασική μεταβλητή στην επόμενη επανάληψη της Simplex).

3

4

Βήμα 1.2. *(Κριτήριο εφικτότητας). Εστω*

$$\overline{P}(1,2) = B^{-1}P(1,2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} \overline{P}(1,2)_1 \\ \overline{P}(1,2)_2 \\ \overline{P}(1,2)_3 \end{bmatrix}.$$

Η μεταβλητή που εξέρχεται από τη βάση είναι η

$$j_r := \min_{\overline{P}(1,2)_i > 0} \left\{ \frac{\overline{b}_i}{\overline{P}(1,2)_i} \right\} = \frac{1}{1} = \frac{\overline{b}_1}{\overline{P}(1,2)_1},$$

άρα: $(x_B)_r = \zeta(1,3)$[5].

Νέο σύνολο βασικών μεταβλητών: $\{\zeta(1,2), \zeta(2,3), \zeta(3,4)\}$.

Νέος πίνακας βάσης είναι ο

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{με} \quad B_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{και} \quad \overline{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Προχωρούμε λοιπόν στη 2η επανάληψη της Simplex.

Βήμα 1.3 *Υπολογισμός Πολλαπλασιαστών Simplex.*

$$\pi = c_B B^{-1} = [\ 4\ ,\ 2\ ,\ 0\ ] \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = [\ 6\ ,\ 2\ ,\ 0\ ]$$

Βήμα 2.1 *(Κριτήριο Βελτίωσης). Υπολογισμός των νέων συντελεστών της αντικειμενικής συνάρτησης για τις μη βασικές μεταβλητές.*

$$\overline{c}(1,3) = c(1,3) - \pi P(1,3) = 8 - (\ 6\ ,\ 2\ ,\ 0\ ) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 8 - 6 = 2 > 0.$$

Επειδή το $\overline{c}(1,3)$ *είναι μη αρνητικό, η* B_1 *είναι βέλτιστη. Οι αντίστοιχες (βέλτιστες) τιμές των βασικών μεταβλητών είναι οι εξής*

$$(\zeta^*(1,2), \zeta^*(2,3), \zeta^*(3,4)) = (1,2,3)$$

και η βέλτιστη πολιτική αποτελείται απο τα βέλη $(1,2)$, $(2,3)$ *και το, τεχνητό, βέλος* (3.4).

Παρατηρήσεις ***1.*** *Οπως έχουμε ήδη παρατηρήσει, το στάδιο εύρεσης μιας αρχικής βασικής μεταβλητής μπορεί να αποφευχθεί αν διαλέξουμε ένα σύνολο αρχικών βασικών μεταβλητών* $\{\zeta(x,k)\}$ *(ή ισοδύναμα τις βασικές κολώνες* $B(x,k)$*) έτσι ώστε η μεταβλητή* x *να διατρέχει όλες τις καταστάσεις, δηλ.* $x = 1,2,3$ *και για κάθε* x *να υπάρχει ένα μόνο* $k \in \overline{D}(x)$*. Σε μια τέτοια επιλογή βάσης αντιστοιχεί μια μοναδική πολιτική* $\overline{\alpha}$*, με* $\overline{\alpha}(x) = k$*, όπου* k *είναι η μοναδική*

5

απόφαση στο $\overline{D}(x)$ τέτοια ώστε $\zeta(x,k)$ είναι βασική μεταβλητή. Είναι σημαντικό να παρατηρήσουμε ότι όμοια δομή θα έχουν όλα τα επακόλουθα βασικά σύνολα. Συγκεκριμένα αν σε κάποιο στάδιο της Simplex η μεταβλητή $\zeta(x^0,k^0)$ εισέρχεται στη βάση, τότε η μεταβλητή που θα βγει θα είναι η $\zeta(x^0,k)$ που αντιστοιχεί στην ίδια κατάσταση x^0 (γιατί;).

2. Το βήμα υπολογισμού των πολλαπλασιαστών Simplex είναι ισοδύναμο με το βήμα 1 της μεθόδου των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών. Πραγματικά, ο υπολογισμός του διανύσματος $\pi = c_B B^{-1}$ είναι βασικά ισοδύναμος με τη λύση του συστήματος των εξισώσεων

$$\pi B = c_B$$

το οποίο λόγω της εκάστοτε μορφής της βάσης B έχει τη μορφή

$$\pi(x) - \pi(\overline{\alpha}(x)) = c(x,\overline{\alpha}(x)) \qquad x = 1, 2$$

$$\pi(3) = 0$$

όπου $\{ \overline{\alpha}(x) , x = 1,2,3\}$ ορίζουν τη πολιτική που αντιστοιχεί στη βάση B. Αρα $\pi(x) = w(x)$, όπου το $w(x)$ είναι η συνάρτηση τιμών αντιστοιχούσα στην πολιτική $\overline{\alpha}(\cdot)$.

***3.** Το Βήμα 1 της Simplex είναι ισοδύναμο με το Βήμα 2 της μεθόδου των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών. Πραγματικά, για κάθε μια μη–βασική μεταβλητή, $\zeta(x,y)$, έχουμε:*

$$\overline{c}(x,y) = c(x,y) - \pi A(x,y) = c(x,y) - \pi(x) + \pi(y) =$$

$$= c(x,y) + w(y) - w(x) = \Delta w(x,y).$$

Δηλαδή οι νέοι (αναθεωρημένοι) συντελεστές κόστους της Simplex είναι ίσοι με τις ποσότητες ελέγχου $\Delta w(x,y)$ της μεθόδου Δ.Π.Π.

***4.** Οταν ελύσαμε με την μέθοδο των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών το πρόβλημα 1.1 (σελ. 12–15), αλλάξαμε σε ένα βήμα όλες τις αποφάσεις που είχαν $\Delta w^{(0)}(x,y) < 0$. Μπορούσαμε βέβαια να είχαμε αλλάξει μια μόνο απόφαση, οπότε θα χρειαζόμαστε παραπάνω απο ένα βήμα για να φτάσουμε στη βέλτιστη λύση. Σ' αυτή την περίπτωση θα είχαμε τη γνωστή (απο τις παραπάνω παρατηρήσεις) αναλογία με την μέθοδο Simplex. Στη πρώτη περίπτωση έχουμε πάλι αναλογία με μια μορφή της Simplex, στην οποία σε κάθε βήμα επιτρέπουμε να εισέλθουν στη βάση παραπάνω από μία, ή ακόμη και όλες οι μη βασικές μεταβλητές. Αυτή είναι η λεγόμενη Simplex με block-pivoting.*

***5.** Το πρωτεύον πρόβλημα γραμμικού προγραμματισμού που αντιστοιχεί στο πρόβλημα της Εικόνας 2 είναι*

$$(\max)\, z = 1 \cdot \xi(1) + 1 \cdot \xi(2) + 1 \cdot \xi(3)$$

περιορισμοί

$$\begin{array}{rrrcl} \xi(1) & -\ \xi(2) & & \le & 4\,, \\ \xi(1) & & -\ \xi(3) & \le & 8\,, \\ & \xi(2) & -\ \xi(3) & \le & 2\,, \\ & & \xi(3) & = & 0\,. \end{array}$$

Δεδομένου ότι $\xi(3) = 0$, μπορούμε να θεωρήσουμε το πρόβλημα ως προς τις δυο μεταβλητές $\xi(1)$ και $\xi(2)$. Η περιοχή εφικτών λύσεων του δίδεται στην Εικόνα 3 σαν η περιοχή F.

(6,2)

(0,0) 1

Εικόνα 3

Παρατηρούμε λοιπόν ότι i) Το πρωτεύον γραμμικό πρόγραμμα έχει μοναδική λύση $\xi^* = (\xi^*(1),\xi^*(2)) = (6,2)$. *Επιπλέον ,*

$$\xi^*(1) = v(1)\,,\ \xi^*(2) = v(2).$$

ii) Η βέλτιστη πολιτική δίδεται από τους περιορισμούς οι οποίοι ισχύουν σαν ισότητες στο σημείο ξ^*, *και συγκεκριμένα από τους (1,2) και (2,3).*

iii) Το πρωτεύον πρόβλημα έχει μοναδική βέλτιστη λύση το σημείο (6,2) για οποιεσδήποτε τιμές των συντελεστών $\lambda(1)$, $\lambda(2)$, *της αντικειμενικής συνάρτησης z (στο συγκεκριμένο παράδειγμα έχουμε πάρει* $\lambda(1) = \lambda(2) = 1$), *εφόσον* $\lambda(1) > 0$ *και* $\lambda(2) > 0$. *Η μοναδικότητα αυτή δεν είμαστε σίγουροι ότι θα ισχύει αν επιτρέψουμε σε μερικά απο τα* $\lambda(x)$ *να είναι ίσα με το μηδέν. Π.χ., αν πάρουμε* $\lambda(1) = 0$ *και* $\lambda(2) = 1$, *τότε προκύπτει η αντικειμενική συνάρτηση* $z' = \xi(2)$, *έχουμε πολλαπλές βέλτιστες λύσεις, και συγκεκριμένα όλα τα σημεία της μορφής* $(\xi(1),2)$, *για* $\xi(1) \leq 6$. *Εν τούτοις, η λύση της βασικής εξίσωσης* $(v(1),v(2)) = (6,2)$ *περιέχεται στο σύνολο των βέλτιστων λύσεων.*

Ολες οι παραπάνω παρατηρήσεις θα αποδειχθούν για το γενικό πρόβλημα σε επόμενη παράγραφο.

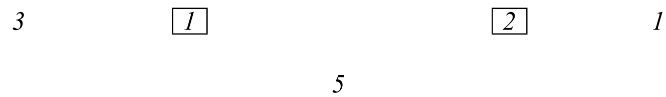
iv) Ενας καλός τρόπος για να γράφουμε σε συντομία τα προβλήματα γραμμικού προγραμματισμού της μορφής που εξετάζουμε είναι ο παρακάτω.

Δ Υ Ι Κ Ο (min)

Π	**Μεταβλητές**	**ζ(1,2)**	**ζ(1,3)**	**ζ(2,4)**	**ζ(3,4)**	**Σχέση**	**Δεξιό μέλος**
Ρ	**ξ(1)**	1	1	0	0	=	λ(1)
Ω							
Τ	**ξ(2)**	− 1	0	1	0	=	λ(2)
Ε							
Υ	**ξ(3)**	0	− 1	− 1	1	=	λ(3)
Ο							
Ν	**Σχέση**	≤	≤	≤	=		
(max)							
	Δ.Μέρος	c(1,2)	c(1,3)	c(2,3)	0		

3. Το Γενικό Πρόβλημα .

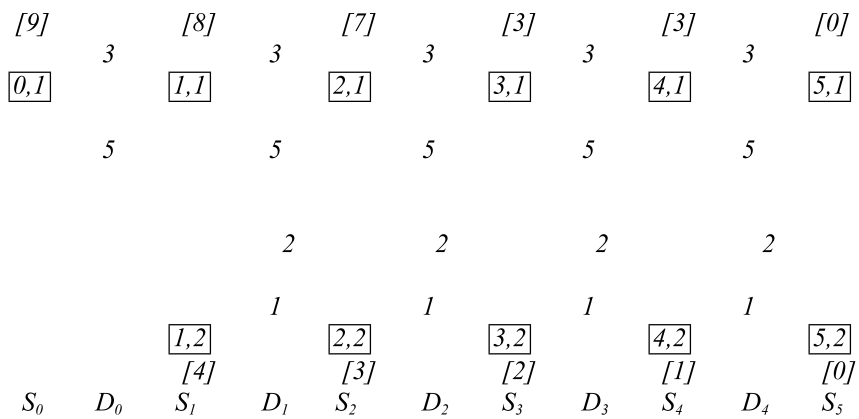
3.1. Εισαγωγικό Παράδειγμα. *Η μέθοδος του δυναμικού προγραμματισμού μπορεί να χρησιμοποιηθεί για την λύση προβλημάτων πολύ πιο γενικών από το πρόβλημα 1.1 που έχουμε ήδη μελετήσει σε λεπτομέρεια. Σαν πρώτο βήμα στη μελέτη πιό γενικών προβλημάτων, ας θεωρήσουμε το πρόβλημα της Εικόνας 4.*



Εικόνα 4

Επιτρέπεται να μετακινηθούμε κατά μήκος των υπαρχόντων βελών, και το ζητούμενο είναι να βρούμε τη διαδρομή (διαδρομές) ελάχιστου μήκους που έχει (έχουν) το σημείο (κόμβο) 1 σαν αρχικό σημείο, και που αποτελούνται από 5 βέλη.

Για να καθορίσουμε τη βέλτιστη διαδρομή στο δίκτυο της εικόνας 4, είναι εύκολο να δούμε ότι σε κάθε βήμα πρέπει να πάρουμε υπ' όψη μας, όχι μόνο τον κόμβο πάνω στον οποίο βρισκόμαστε, αλλά επίσης και το ***πόσα βήματα μας απομένουν.*** *Ετσι οδηγούμαστε στο να θεωρήσουμε το πρόβλημα της διαδρομής ελάχιστου μήκους που περιγράφεται στην εξής εικόνα.*



Εικόνα 5

Στην παραπάνω εικόνα έχουμε τα σύνολα των κόμβων S_0, S_1, S_2, S_3, S_4, S_5, για τις διαφορετικές ***χρονικές περιόδους (ή στάδια)*** *$t = 0,1,2,\ldots,5$. Οι κόμβοι στο σύνολο S_0 είναι οι* ***αρχικοί*** *κόμβοι, ενώ εκείνοι στο σύνολο S_5 οι* ***τερματικοί.*** *Τα σύνολα D_t, $0 \le t \le 4$, συμβολίζουν όλες τις δυνατές αποφάσεις στις διάφορες χρονικές περιόδους. Οι δυνατές αποφάσεις στον κόμβο (t,x) συμβολίζονται με $D(x)$. Ετσι, έχουμε ότι $D(1) = \{(1,1), (1,2)\}$, $D(2) = \{(2,1), (2,2)\}$. Ανάλογα με την περίπτωση, θα χρησιμοποιήσουμε επίσης το συμβολισμό $\bar{D}(x)$ για τις αποφάσεις στον κόμβο x, όπως εκάμαμε στο προηγούμενο πρόβλημα.*

Παρατηρούμε ότι σε κάθε διαδρομή που αποτελείται από 5 βέλη πάνω στο δίκτυο της εικόνας 4, μπορούμε να αντιστοιχίσουμε μια διαδρομή πάνω στο δίκτυο της εικόνας 5. Το πλεονέκτημα στη δεύτερη εικόνα είναι ότι μας παρέχει τη δυνατότητα να περιγράψουμε τις μεταβλητές ενός συστήματος μέσα στο χρόνο. Δηλαδή η δεύτερη περιγράφει τη ***κίνηση (δυναμική*** *του συστήματος) ενώ η πρώτη είναι* ***στατική*** *(περιγράφει τις* ***δυνατότητες*** *για κίνηση). Κάθε διαδρομή έχει τη μορφή $l = x_0\, \alpha(x_0)\, x_1\, \alpha(x_1)\;\; x_2\, \alpha(x_2)\, x_3\, \alpha(x_3)\, x_4\, \alpha(x_4)\, x_5$, όπου $\alpha(x_i) \in D(x_i)$.*

Ορίζουμε την ***αξία*** *της διαδρομής (το μήκος της) ως εξής*

$$w(l) = c(\alpha(x_0)) + c(\alpha(x_1)) + \ldots + c(\alpha(x_4)) \tag{1}$$

ή ισοδύναμα

$$w(l) = \sum_{t=0}^{4} c(\alpha(x_t)) + \hat{c}(x_5), \tag{2}$$

όπου c είναι μια συνάρτηση ορισμένη πάνω στα σύνολα D_t (βέλη) και λέγεται συνάρτηση ***τρέχοντος κόστους,*** *ενώ η $\hat{c}$ είναι συνάρτηση πάνω στο x_5 και λέγεται συνάρτηση* ***τερματικού κόστους.*** *Στο συγκεκριμένο πρόβλημα της εικ. 5 έχομε ότι $\hat{c}(x) = 0$, $x = 1,2$.*

Με τις γνώσεις που έχουμε αποκτήσει λύνοντας το πρόβλημα 1.1 , είναι εύκολο να βρούμε τη λύση για το πρόβλημα της εικόνας 4 . Αν κάποιος επιμένει να δει την ισοδυναμία, δεν έχει παρά να προσθέσει ένα "τεχνητό" κόμβο 3 στο τέλος, για να δημιουργήσει ένα τερματικό σύνολο $x_6 = \{3\}$ *με μια μόνο κατάσταση (σημείο), την 3 , και "τεχνητές" αποφάσεις (1,3) , (2,3) , με μήκη* $c(1,3) = c(2,3) = 0$ *. Το τέχνασμα αυτό όμως αυτή δεν είναι απαραίτητο, αρκεί να ορίσουμε τις τιμές* $v(x) = 0$ *για* $x \in S_5$ *. Χρησιμοποιώντας τα επιχειρήματα που αναπτύξαμε στο εισαγωγικό παράδειγμα 1.1, μπορούμε να γράψουμε την εξίσωση βελτιστοποίησης ως εξής*

$$v(t,x) = \min_{\alpha(x) \in D(x)} \{c(x,\bar{\alpha}(x)) + v(t+1,\bar{\alpha}(x))\}, \qquad x \in S_t, \quad t = 0,1,\ldots,4 \tag{3}$$

$$v(5,x) = 0, \qquad x \in S_5. \tag{4}$$

Στην εικόνα 5 , δίπλα από τους κόμβους, δίνονται οι τιμές $v(t,x)$ *εντός τετραγώνων παρενθέσεων [ ] . Ετσι βρίσκουμε ξεκινώντας απο το τέλος, ότι* $v(0,1) = 9$ *, για* $1 \in S_0$ *, και ότι η μόνη βέλτιστη διαδρομή είναι η διαδρομή*
$l^0 = 1\ (1,2)\ 2\ (2,2)\ 2\ (2,2)\ 2\ (3,2)\ 2\ (4,2)\ 2.$

Ο τρόπος λύσης, λοιπόν, δεν είναι διαφορετικός από αυτόν που χρησιμοποιήσαμε για το πρόβλημα 1.1 . Θα ήταν πάντως διδακτικό για τον αναγνώστη να δοκιμάσει τις διάφορες λύσεις που έχει στη διάθεση του, όπως κάναμε για το πρόβλημα 1.1. Εξετάζουμε το πρόβλημα 3.1 , γιατί τούτο αποτελεί ένα πολύ καλό μοντέλο για το γενικό πρόβλημα του δυναμικού προγραμματισμού με προσδιοριστική δυναμική στη περίπτωση του πεπερασμένου ορίζοντα.

Θα εισάγουμε λοιπόν, πριν προχωρήσουμε στο γενικό πρόβλημα, τον απαραίτητο συμβολισμό, χρησιμοποιώντας το πρόβλημα 3.1 σαν παράδειγμα.

Συμβολισμός ***1.*** *Για τον υπολογισμό της συνάρτησης τιμής* $v(\cdot)$ *για το πρόβλημα πρότυπο, είναι απαραίτητο να θυμόμαστε πόσα βήματα μας απομένουν μέχρι το τέλος του* ***ορίζοντα*** *(5η περίοδος) ή, ισοδύναμα, πόσα βήματα έχουν περάσει απο την αρχή του ορίζοντα. Τούτο επιτυγχάνεται με το να ονομάσουμε (t,x) τους κόμβους στο σύνολο* S_t *, και* (t,α) *τα βέλη στο σύνολο* D_t *,* $t = 1,\ldots,5$*. Ετσι η βασική εξίσωση γράφεται*

$$v\,(t,x) = \min_{\bar{\alpha} \in \bar{D}(x)} \{c((t,x),\bar{\alpha}) + v(t+1,\bar{\alpha}\,(x))\}, \tag{5}$$

$$v(5,x) = 0. \tag{6}$$

2. *Εστω P το αρχικό πρόβλημα πάνω στους χώρους καταστάσεων* S_t *και αποφάσεων* D_t *,* $t = 0,\ldots,5$ *. Συμβολίζουμε με* $P(t,x)$ *το πρόβλημα εύρεσης της διαδρομής ελάχιστου μήκους απο τον κόμβο* $x \in S_t$ *(ισοδύναμα τον κόμβο (t,x)) στο τερματικό σύνολο* S_5 *. Το πρόβλημα* $P(t,x)$ *είναι της ίδιας μορφής με το P , αλλά* ***μικρότερο,*** *δεδομένου ότι προέρχεται απο το P με τη διαγραφή των συνόλων* $S_0,D_0,S_1,D_1,\ldots,S_{t-1},D_{t-1}$ *, και τη συρρίκνωση των* S_t *,* D_t *στα* $\{x\}$ *,* $D(x)$ *. Η τιμή του υποπροβλήματος* $P(t,x)$ *δεν είναι τίποτα άλλο απο την τιμή* $v(t,x)$ *. Παρατηρούμε λοιπόν ότι όταν βρίσκουμε την λύση της βασικής εξίσωσης πηγαίνοντας προς τα πίσω, δεν κάνουμε τίποτα άλλο από το να λύνομε διαδοχικά τα υποπροβλήματα:* $P(4,x)\ x \in S_4$ *,* $P(3,x)\ x \in S_3$ *, ... ,* $P(1,x)\ x \in S_1$ *,* $P\ (0,1)$ *. Δηλαδή* ***βάζουμε*** *το P* ***μέσα*** *σε μια ολόκληρη οικογένεια συναφών αλλά μικρότερων προβλημάτων.*

3. *Η συνάρτηση τιμών* $v(\cdot)$ *είναι ορισμένη πάνω στο σύνολο των κόμβων*

$$S = S_0 \cup S_1 \cup \ldots \cup S_5.$$

Σε αναλογία μπορούμε αν ορίσουμε μια συνάρτηση τιμών $u(\cdot,\cdot)$ *πάνω στο σύνολο των βελών (αποφάσεων)*

$$D = D_0 \cup D_1 \cup \ldots \cup D_5$$

ως εξής

$$u(x,\alpha) = c(x,\alpha) + v(\alpha) \qquad \forall\, x \in S - S_5,\ \alpha \in D(x) \tag{7}$$

Χρησιμοποιώντας τη συνάρτηση $u(\cdot)$ *μπορούμε να σπάσουμε τη βασική εξίσωση σε δυο εξισώσεις ως εξής*

$$v(x) = 0 \qquad x \in S_5 \tag{8}$$
$$u(x,\alpha) = c(x,\alpha) + v(\overline{\alpha}) \qquad \forall x,\ \alpha \in D(x) \tag{9}$$
$$v(x) = \min_{\alpha \in D(x)} u(x,\alpha) \qquad x \in S - S_5 \tag{10}$$

3.2 Ορισμός του Γενικού Προβλήματος. *Είμαστε τώρα σε θέση να ορίσουμε το γενικό πρόβλημα ελέγχου δυναμικών συστημάτων με προσδιοριστική δυναμική κίνηση, πάνω σε ένα πεπερασμένο ορίζοντα. Τούτο ορίζεται από τα παρακάτω στοιχεία.*

*1ον Τους **χώρους καταστάσεων** $S_0, S_1, \ldots, S_{n+1}$, όπου $n+1$ είναι ένας γνωστός ακέραιος, το μήκος του ορίζοντα. S_0 (αντίστοιχα S_{n+1}) είναι οι χώροι των **αρχικών** (αντίστοιχα **τερματικών**) καταστάσεων.*

*2ον Την **Δυναμική του συστήματος,** δηλαδή το μηχανισμό που ορίζει τη μεταβολή της κατάστασης του συστήματος (νόμο κίνησης) σαν αποτέλεσμα των αποφάσεων που παίρνουμε. Συγκεκριμένα, για κάθε κατάσταση $x \in S_t$, $0 \le t \le n$, υπάρχουν γνωστά σύνολα $D(x)$ που περιέχουν όλες τις δυνατές (επιτρεπτές) αποφάσεις στην κατάσταση x. Το σύνολο $D(x)$ ονομάζεται **δέσμη αποφάσεων** στο x, ενώ το $D_t = \bigcup_{x \in S_t} D(x)$ είναι ο **χώρος αποφάσεων** για τη χρονική περίοδο t, $0 \le t \le n$.*

*Κάθε απόφαση (βέλος) (x,α), $\alpha \in D(x) \subset D_t$, $0 \le t \le n$, ορίζει την **μοναδική** κατάσταση στην οποία θα ευρίσκεται το σύστημα στην αρχή της επόμενης χρονικής περιόδου. Το γεγονός ότι η απόφαση (x,α) προσδιορίζει την κατάσταση στην επόμενη χρονική περίοδο, ονομάζεται **προσδιοριστική δυναμική** κίνηση του συστήματος.*

*3ον Τη συνάρτηση **τρέχοντος κόστους** $c_t(x,\overline{\alpha})$, $x \in S_t$, $\overline{\alpha} \in \overline{D}(x)$, $0 \le t \le n$, και την συνάρτηση **τερματικού κόστους** $\hat{c}(x)$, $x \in S_{n+1}$.*

***Παρατηρήσεις 1.** Οι χώροι καταστάσεων S_t και αποφάσεων D_t ήταν πεπερασμένα σύνολα στα προβλήματα 1.1 και 1.2 που έχουμε ήδη μελετήσει. Στο γενικό πρόβλημα επιτρέπουμε πιο γενικούς χώρους, π.χ. τα S_t, D_t είναι δυνατόν να είναι υποσύνολα του $\mathbb{R}^n$ ή πιο γενικών ακόμη χώρων (π.χ. χώρων συναρτήσεων).*

***2.** Στο γενικό πρόβλημα επιτρέπουμε στη συνάρτηση τρέχοντος κόστους να εξαρτάται απο το χρόνο t, γι αυτό γράψαμε $c_t(x,\overline{\alpha})$.*

***3.** Στα προβλήματα 1.1, 1.2 η απόφαση (x,α) ήταν στην κυριολεξία βέλος, και έτσι η δυναμική του συστήματος καθοριζόταν ως εξής: αν $x_t = x$, τότε $x_{t+1} = \overline{\alpha}(x)$. Στο γενικό πρόβλημα απλώς ζητάμε η κατάσταση x_{t+1} να **προσδιορίζεται με μοναδικό τρόπο** από τη γνώση των $x_t = x$ και της απόφασης α που παίρνουμε στην κατάσταση x. Ετσι μπορούμε να γράψουμε*

$$x_{t+1} = g_t(x,\alpha),$$

για γνωστές συναρτήσεις $g_t(\cdot)$, ορισμένες στους χώρους $S_t \times D_t$, $0 \le t \le n$. Εχουμε δηλαδή δύο ισοδύναμους τρόπους για να περιγράψουμε τη δυναμική του συστήματος:

$$(\alpha)\quad x_{t+1} = \overline{\alpha}_t, \qquad \alpha_t \in D(x_t),$$

και

$$(\beta)\quad x_{t+1} = g_t(x_t,\alpha_t), \qquad \alpha_t \in D_t(x_t).$$

***4.** Στο πρόβλημα πρότυπο έχουμε $S_o = \{1\}$. Εχει όμως νόημα, και είναι εύκολο να το δούμε, ότι η ίδια μεθοδολογία μπορεί να εφαρμοστεί για τη λύση του προβλήματος αν πάρουμε σαν χώρο αρχικών καταστάσεων το σύνολο $\{1,2\}$. Σ' αυτή την περίπτωση το ζητούμενο είναι να βρούμε τις λύσεις (διαδρομές που αποτελούνται απο 5 βέλη) που έχουν σαν αρχική κατάσταση τον κόμβο 1,* *καθώς και εκείνες που έχουν σαν αρχική κατάσταση τον κόμβο 2. Επίσης, έχει νόημα να θεωρήσουμε προβλήματα στα οποία η αρχική κατάσταση δεν είναι γνωστή, αλλά ορίζεται από μια γνωστή αρχική κατανομή πάνω στο* S_0 .

Πολιτικές και Τροχιές. *Εστω δύο σύνολα* D,E *. Μια απεικόνιση* Φ *, που σε κάθε σημείο* x *του* D *αντιστοιχεί ένα υποσύνολο* $\Phi(x)$ *του* E *ονομάζεται αντιστοιχία του* D *στο* E *. Μια συνάρτηση* ϕ *από το* D *στο* E *, με την ιδιότητα* $\phi(x) \in \Phi(x)\ \forall x \in D$ *, ονομάζεται* ***επιλογή*** *της* Φ.

Στο γενικό πρόβλημα του δυναμικού προγραμματισμού έχουμε την αντιστοιχία που ορίζεται απο τα σύνολα–δέσμες αποφάσεων $D(x)$, $x \in S - S_{n+1}$ *, δηλαδή από το* $S - S_{n+1}$ *στο* $D = D_0 \cup \ldots D_n$.

Μια επιλογή $\alpha(\cdot)$ *της παραπάνω αντιστοιχίας ονομάζεται* ***απλή πολιτική*** *(ή απλή στρατηγική).*

Χρησιμοποιώντας μια απλή πολιτική $\alpha(\cdot)$ *παίρνουμε μια* ***τροχιά*** *του συστήματος* $l = x_0 , \alpha_0 , x_1 , \alpha_1 , \ldots , x_n , \alpha_n , x_{n+1}$ *, όπου* $x_{t+1} = g_t(x_t,\alpha_t)$, $x_t \in S_t$, $\alpha_t \in D(x_t)$.

Σε κάθε χρονική περίοδο t *,* x_t *συμβολίζει την παρούσα κατάσταση ,* α_t *την παρούσα απόφαση ,* $h_t = x_0 , \alpha_0 , \ldots , x_{t-1} , \alpha_{t-1} , x_t$ *την* ***ιστορία*** *του συστήματος και* $f_t = x_{t+1} , \alpha_{t+1} , \ldots , x_n , \alpha_n , x_{n+1}$ *το* ***μέλλον*** *του συστήματος.*

Μια πολιτική είναι ένας κανόνας για να παίρνουμε αποφάσεις σε κάθε χρονική περίοδο και σε κάθε κατάσταση (κόμβο).

Υπάρχουν πολλοί τρόποι να ορίσουμε πολιτικές. Ο πιο γενικός τρόπος είναι να θεωρήσουμε ότι σε κάθε χρονική περίοδο t *επιλέγουμε την απόφαση* α_t *, λαμβάνοντας υπ' όψη όχι μόνο την κατάσταση του συστήματος* x_t *, αλλά όλη την ιστορία* h_t *. Ετσι μια πολιτική* π *είναι μια συνάρτηση που σε κάθε ιστορία* $h_t = x_0 , \alpha_0 , \ldots , x_{t-1} , \alpha_{t-1} , x_t$ *αντιστοιχεί μια μοναδική απόφαση* $\alpha_t = \pi(ht) \in D(x_t)$*. Μια πολιτική ονομάζεται* ***Μαρκοβιανή*** *αν* $\pi(h_t) = \pi(t,x_t)$*, δηλαδή ο κανόνας σύμφωνα με τον οποίο λαμβάνονται οι αποφάσεις δεν εξαρτάται από όλη την ιστορία παρά μόνο από την παρούσα κατάσταση* x_t *και την χρονική περίοδο* t *. Αν επιπλέον ισχύει ότι* $\pi(h_t) = \pi(x_t)$ *, δηλαδή η εκάστοτε απόφαση καθορίζεται απο την παρούσα κατάσταση, μόνο τότε η* $\pi(\cdot)$ *λέγεται* ***στάσιμη.*** *Για συστήματα με προσδιοριστική δυναμική* ***στάσιμες*** *πολιτικές είναι οι απλές πολιτικές.*

Εστω Π *το σύνολο των πολιτικών ,* Π_M *το σύνολο των Μαρκοβιανών και* Π_Σ *το σύνολο των στάσιμων πολιτικών. Είναι προφανές ότι:* $\Pi_\Sigma \subset \Pi_M \subset \Pi$.

3.3 . Η Εξίσωση Βελτιστοποίησης . *Ας θεωρήσουμε το γενικό πρόβλημα που εισαγάγαμε στην παράγραφο (3.2)*

(min) $$w = \sum_{t=0}^{n} c_t(x_t, \alpha_t) + \hat{c}(x_{n+1})$$

δυναμική κίνηση $x_{t+1} = \bar{\alpha}_t,\quad x_t \in S_t ,\ \alpha_t \in D_t(x_t) ,\quad t = 0,1,\ldots, n$

αρχική θέση $X_0 = \{x_0\}$.

Το πρόβλημα αυτό το συμβολίζουμε με $P(x_0)$.

Η κεντρική ιδέα για τη λύση του $P(x_0)$ *με τη μέθοδο του Δυναμικού Προγραμματισμού βασίζεται στο να θεωρήσουμε το* $P(xo)$ *σαν ένα πρόβλημα της εξής οικογένειας προβλημάτων* $P(k,x)$.

(min) $$w = \sum_{t=k}^{n} c_t(x_t,\alpha_t) + \hat{c}(x_{n+1})$$

δυναμική κίνηση $x_{t+1} = \bar{\alpha}_t,\qquad t = k,k+1,\ldots, n$

$x_t \in S_t ,\quad \alpha_t \in D_t(x_t),\qquad t = k,k+1,\ldots, n.$

αρχική θέση $x_k = x.$

Στους παραπάνω ορισμούς έχουμε επιλέξει να εκφράζουμε τη δυναμική κίνηση του συστήματος μέσω των σχέσεων: $x_{t+1} = \overline{\alpha}_t$, *όπου* $\overline{\alpha}_t$ *είναι το τελικό σημείο του βέλους* α_t. *Ενας ισοδύναμος τρόπος που θα χρησιμοποιήσουμε παρακάτω, είναι το να θεωρήσουμε ότι υπάρχουν συναρτήσεις* $g_t = g_t(x_t,\alpha_t)$ *έτσι που* $x_{t+1} = g_t(x_t,\alpha_t)$.

Ας υποθέσουμε ότι τα προβλήματα $P(k,x)$ *έχουν βέλτιστες λύσεις για κάθε* $x \in S_t$ *και για κάθε* k, $0 \le k \le n$. *Το παρακάτω αποτέλεσμα είναι στοιχειώδες, αλλά βασικό για τη λύση του* $P(x_0)$ *με τη μέθοδο του Δυναμικού Προγραμματισμού.*

Λήμμα 1. *Εστω* $\{\alpha_t^*, t = k,\ldots, n\}$ *μια βέλτιστη ακολουθία αποφάσεων για το πρόβλημα* $P(k,x)$, *και έστω* $\{x_t^*, t = k,\ldots, n+1\}$ *η αντίστοιχη βέλτιστη τροχιά. Τότε για κάθε* $m \in \{k,\ldots,n\}$ *η ακολουθία των αποφάσεων* $\{\alpha_t^*, t = m,m+1,\ldots,n\}$ *είναι μια βέλτιστη πολιτική για το πρόβλημα* $P(m,x_m^*)$.

Απόδειξη. *Ας υποθέσουμε το αντίθετο, ότι δηλαδή υπάρχει κάποιο* $m>k$ *τέτοιο, ώστε για το πρόβλημα* $P(m,x_m^*)$ *μπορούμε να βρούμε μια πολιτική* $\{\tilde{\alpha}_t, m \le t \le n\}$, *με αντίστοιχη τροχιά* $\{\tilde{x}_t, m \le t \le n+1\}$ *έτσι ώστε:*

$$\sum_{t=m}^{n} c_t(\tilde{x}_t,\tilde{\alpha}_t) + \hat{c}(\tilde{x}_{n+1}) < \sum_{t=m}^{n} c_t(x_t^*,\alpha_t^*) + \hat{c}(x_{n+1}^*) \tag{1}$$

Ας θεωρήσουμε τώρα την πολιτική $\{\hat{\alpha}_t, k \le t \le n\}$, *με αντίστοιχη τροχιά* $\{\hat{x}_t, k \le t \le n+1\}$, *για το πρόβλημα* $P(k,x_k^*)$, *που ορίζεται ωσεξής*

$$\hat{\alpha}_t = \begin{cases} \alpha_t^*, & t = k,k+1,\ldots,m-1 \\ \tilde{\alpha}_t, & t = m,m+1,\ldots, n \end{cases} \tag{2}$$

Παρατηρούμε ότι ο παραπάνω ορισμός της $\{\hat{\alpha}_t, k \le t \le n\}$ *συνεπάγεται ότι*

$$\hat{x}_t = \begin{cases} x_t^*, & t = k,k+1,\ldots,m-1 \\ \tilde{x}_t, & t = m,m+1,\ldots,n+1 \end{cases} \tag{3}$$

Εστω $\hat{w}$ *η τιμή της αντικειμενικής συνάρτησης του προβλήματος* $P(k,x_k^*)$, *που αντιστοιχεί στην πολιτική* $\{\hat{\alpha}_t, k \le t \le n\}$. *Τότε:*

$$\begin{aligned} \hat{w} &= \sum_{t=k}^{n} c_t(\hat{x}_t,\hat{\alpha}_t) + \hat{c}(\hat{x}_{n+1}) = \\ &\sum_{t=k}^{m-1} c_t(\hat{x}_t,\hat{\alpha}_t) + \sum_{t=m}^{n} c_t(\hat{x}_t,\hat{\alpha}_t) + \hat{c}(\hat{x}_{n+1}) = && ((2),(3)) \\ &= \sum_{t=k}^{m-1} c_t(x_t^*,\alpha_t^*) + \sum_{t=m}^{n} c_t(\tilde{x}_t,\tilde{\alpha}_t) + \hat{c}(\tilde{x}_{n+1}) < && ((1)) \\ &< \sum_{t=k}^{m-1} c_t(x_t^*,\alpha_t^*) + \sum_{t=m}^{n} c_t(x_t^*,\alpha_t^*) + \hat{c}(x_{n+1}^*) \quad \Rightarrow \end{aligned}$$

$$\hat{w} < \sum_{t=k}^{n} c_t(x_t^*,\alpha_t^*) + \hat{c}(x_{n+1}^*) \tag{4}$$

Η (4) βέβαια σημαίνει ότι η πολιτική $\{\alpha_t^*, k \le t \le n\}$ *δεν είναι βέλτιστη για το πρόβλημα* $P(k,x)$, *πράγμα άτοπο. Αρα το λήμμα ισχύει.*

Παρατήρηση *Σχηματικά, η απόδειξη του λήμματος μπορεί να περιγραφεί ως εξής*

$$\begin{array}{lll} & & \delta_1 \\ P(k,x) \;:\; x = x^*_\kappa\, \alpha^*_\kappa\, x^*_{\kappa+1} \ldots\; \alpha^*_{m-1} & & x^*_m\, \alpha^*_m\, x^*_{m+1} \ldots\; x^*_n\, \alpha^*_n\, x^*_{n+1} \\ \qquad\qquad\qquad \delta_0 & & \\ P(m,x^*_m) \;:\; \tilde{x}_m \;= & & x^*_m\, \tilde{\alpha}_m\, \tilde{x}_{m+1} \ldots\; \tilde{x}_n\, \tilde{\alpha}_n\, \tilde{x}_{n+1} \\ & & \delta_2 \end{array}$$

Αν ο δρόμος $l = (\delta_0,\delta_1)$ *είναι βέλτιστος για το* $P(k,x)$, *τότε ο δρόμος* δ_1 *είναι βέλτιστος για το* $P(m,x^*_m)$, *διαφορετικά ο δρόμος* $l' = (\delta_0,\delta_2)$ *θα ήτανε καλύτερος από τον δρόμο* l *για το πρόβλημα* $P(k,x)$, *πράγμα άτοπο.*

Συγκρίνετε το παραπάνω επιχείρημα με εκείνο του 1.2 .

Θεώρημα 1. *Ας υποθέσουμε ότι τα προβλήματα* $P(k,x)$ *έχουν βέλτιστες λύσεις για κάθε* $x \in S_k$, $k = 0,1,\ldots, n$. *Δηλαδή υποθέτουμε ότι υπάρχει η συνάρτηση τιμών* $v(k,x)$ *του* $P(k,x)$ *που ορίζεται ως εξής*

$$v(x,k) \;=\; min \left\{ \sum_{t=k}^{n} c_t\,(x_t,\alpha_t) + \hat{c}(x_{n+1}) \;\;/\;\; x_{t+1} \;=\; \overline{\alpha}_t\,,\;\; \alpha_t \;\in\; D_t(x_t)\,,\;\; x_t \;\in\; S_t \right\}$$

Τότε

(1) η συνάρτηση $v(\cdot\,,\cdot)$ *είναι η μόνη λύση της εξίσωσης*

$$v(k,x) \;=\; \min_{\alpha \in D_k(x)} \{c_k(x,\alpha) + v(k+1, \overline{\alpha})\} \qquad\qquad 0 \le k \le n\,,\; x \in S_k, \tag{5}$$

όπου

$$v(n+1,x) \;=\; \hat{c}\;(x) \qquad\qquad x \in S_{n+1}$$

2) υπάρχει (τουλάχιστον) μια βέλτιστη πολιτική π^0 , *που είναι Μαρκοβιανή, και που ορίζεται μέσω της εξίσωσης*

$$c_k(x,\overline{\pi}^0(x)) + v(k + 1, \overline{\pi}^0(x)) \;=\; \min_{\alpha \in D(x_t)} \{c_k(x,\overline{\alpha}) + v(k+1,\overline{\alpha})\} \tag{6}$$

Απόδειξη. *Εστω* $l^* = (x,\alpha^*_k,x^*_{k+1},\ldots,x^*_n,\alpha^*_{n+1},x^*_{n+1},x_T)$ *μια βέλτιστη τροχιά για το πρόβλημα* $P(k,x)$, *όπου έχουμε προσθέσει ένα τεχνητό (φανταστικό) τερματικό κόμβο* x_T . *Εστω* l' *μια άλλη τροχιά ,* $l' = (x,\alpha_k{}',x_{k+1}{}',\ldots, x_{n+1}{}',x_T)$. *Τα* $w(l^*)$, $w(l')$ *συμβολίζουν τις αξίες των τροχιών. Εχουμε λοιπόν ότι*

$$w(l') \;\ge\; w(l^*) \;=\; v(k,x) \tag{7}$$

Απο το λήμμα 1. έχουμε ότι

$$w(l^*) \;=\; c_k(x,\overline{\alpha}^*_k) + v(k+1,\overline{\alpha}^*_k) \tag{8}$$

ενώ

$$w\,(l') \;\ge\; c_k(x,\overline{\alpha}_k{}') + v(k+1,\overline{\alpha}_k{}'), \tag{9}$$

και επί πλέον η (1.21) ισχύει σαν ισότητα αν η τροχιά $(x_{k+1},\alpha_{k+1},\ldots,x_{n+1},x_\tau)$ *είναι βέλτιστη για το πρόβλημα* $P(k+1,x_{k+1})$. *Κάνοντας λοιπόν αυτή την παραδοχή, παίρνουμε απο τις (1.19) , (1.20) , (1.21) ότι*

$$v(k,x) = c_k(x,\overline{\alpha}_k^*) + v(k+1,\overline{\alpha}_k^*) \le c_k(x,\overline{\alpha}_k) + v(k+1,\overline{\alpha}_k) \tag{10}$$

για κάθε $\alpha_k \in D(x)$. *Ετσι η απόδειξη γι α το (1) είναι πλήρης.*

2) *Αρκεί να παρατηρήσουμε ότι μια πολιτική είναι βέλτιστη, αν και μόνο αν* $c_k(x,\overline{\pi}(x)) = v(k,x) - v(k+1,\overline{\pi}(x))$.

Παρατηρήσεις. **1.** *Στον ορισμό του προβλήματος (3.3) θεωρήσαμε σαν* ***μέτρο απόδοσης*** $w = w(l)$ *το άθροισμα* $\sum_0^n c_t(x_t,\overline{\alpha}_t) + \hat{c}(x_{n+1})$. *Είναι σημαντικό να παρατηρήσουμε, και εύκολο να αποδειχθεί, ότι το βασικό θεώρημα 1 ισχύει για πιο γενικά μέτρα απόδοσης,* $w = w(l)$. *Συγκεκριμένα, το θεώρημα εξακολουθεί να ισχύει για όλα τα μέτρα απόδοσης* w , *τα οποία πληρούν την εξής απαίτηση: αν οι πρώτες* s *αποφάσεις οδηγούν στην κατάσταση* x_m , *τότε η επίδραση των εναπομενόντων* $n+1-s$ *αποφάσεων πάνω στο* $w = w(l)$ *εξαρτάται μόνο απο την κατάσταση* x_m *και τις τελευταίες* $n+1-s$ *αποφάσεις. Η απαίτηση αυτή ονομάζεται* ***ιδιότητα Markov*** *και συμβολικά μπορεί να εκφραστεί ως εξής αν*
$l = (x_0, \alpha_0, x_1, \ldots, x_s, \ldots, x_n, \alpha_n, x_{n+1})$, $l' = (x_0, \alpha_0', x_1', \ldots, x_s', \ldots, x_n', \alpha_n', x_{n+1}')$
είναι δυο τροχιές για τις οποίες ισχύει

1) $x_t = x_t,\qquad t = 0,1,\ldots,s$
2) $\alpha_t = \alpha_t,\qquad t = 0,1,\ldots,s$
3) $w(x_s,\alpha_s,\ldots,x_{n+1}) = w(x_s,\alpha_s',\ldots,x_{n+1}')$,

τότε $w(l) = w(l')$

Παραδείγματα μέτρων απόδοσης που πληρούν την ιδιότητα Markov είναι τα

i) $w = c(x_{n+1})$

ii) $w = \max\{c_t(x_t,\overline{\alpha}_t),\ t = 0,1,\ldots,n\}$

iii) $w = \sum_{t=0}^{n} b^t c_t(x_t,\overline{\alpha}_t)$

iv) $w = \prod_{t=0}^{n} c_t(x_t,\overline{\alpha}_t),\qquad c_t \ge 0$

v) $w = \sum_{t=0}^{N_Y} c(x_t,\overline{\alpha}_t)$, *όπου* $Y \subset S$, N_Y *είναι η "πρώτη φορά" που* $x_t \in Y$.

Παράδειγμα μέτρου απόδοσης που δεν πληροί την ιδιότητα Markov είναι το εξής

$$w = c_0(x_0,\overline{\alpha}_0)\, c_1(x_1, \overline{\alpha}_1) + (c_1(x_1,\overline{\alpha}_1))^2 .$$

όπου οι $c_t(\cdot\,,\cdot)$ *δεν είναι αναγκαστικά θετικές συναρτήσεις.*

2. *Τονίζουμε ότι ο νόμος κίνησης, όπως ορίζεται στο πρόβλημα (1.4), πληροί την παρακάτω ιδιότητα (αρχή του αιτιατού) για όλες τις πολιτικές* π *που δεν εξαρτώνται από το μέλλον:*

Αν $x_{t+s} = x_{t+s}(\pi)$ *είναι η κατάσταση του συστήματος μετά από* $t+s$ *βήματα, όταν η αρχική κατάσταση είναι* x *κάτω απο την πολιτική* π , *και αν* $x_s = x_s(\pi)$ *είναι η θέση μετά από* s *βήματα, τότε το* x_{t+s} *μπορεί να θεωρηθεί σαν η κατάσταση του συστήματος μετά από* t *βήματα, όταν η αρχική κατάσταση είναι η* x_s.

3. *Η Βασική παραδοχή στο θεώρημα 1 είναι η υπόθεση ότι όλα τα προβλήματα* $P(k,x_k)$ *έχουν λύση. Τούτη η παραδοχή είναι αναγκαία, όπως φαίνεται από τα παρακάτω παραδείγματα.*

Παράδειγμα 1. *Ας υποθέσουμε ότι* $S_0 = \{1\}$, $S_1 = \{2\}$, $D(1) = \{\alpha_1, \alpha_2, \alpha_3, \ldots\}$, $c(1,\alpha_i) = -i$, $\hat{c}(2) = 0$. *Τότε* $v(1) = -\infty$, $v(2) = 0$, *αλλά δεν υπάρχει απόφαση* $\alpha^* \in D(1)$, *έτσι ώστε* $v(1) = c(1,\alpha^*) + v(2)$.

Παράδειγμα 2. *Ας υποθέσουμε ότι* $S_0 = \{1\}$, $S_1 = \{2\}$, $D(1) = \{\alpha_1, \alpha_2, \alpha_3, \ldots\}$, $c(1,\alpha_i) = -1/i$, $\hat{c}(2) = 0$. *Τότε η συνάρτηση τιμών είναι πεπερασμένη παντού:* $v(1) = 0$, $v(2) = 0$, *αλλά πάλι δεν υπάρχει απόφαση* $\alpha^* \in D(1)$ *έτσι ώστε*
$v(1) = c(1,\alpha^*) + v(2)$.

*Σε αντίθεση με το παράδειγμα 1 στο 2 μπορούμε να ορίσουμε κατά προσέγγιση βέλτιστες πολιτικές, όπου η ακρίβεια της προσέγγισης είναι όσο καλή θέλουμε (**ϵ–βέλτιστες** πολιτικές), ως εξής*

$$\forall\ \epsilon>0 \quad \exists\ \alpha_i = \alpha_i(\epsilon), \ \text{έτσι ώστε}$$

$$c(1,\alpha_i^*) + v(2) \in (v(1), v(1) + \epsilon)\ .$$

Πράγματι, αρκεί να πάρουμε την απόφαση α_i, *για την οποία* $i>1/\epsilon$.

4. *Οι υπολογισμοί που χρειάζονται για να λυθεί αριθμητικά ένα πρόβλημα δυναμικού προγραμματισμού συνήθως απαιτούν πολύ χρόνο. Για παράδειγμα, έστω* $n+1 = 9$, *και ότι κάθε σύνολο* S_t $(t = 0,\ldots,10)$ *περιέχει 20 κόμβους. Τότοε πρέπει να υπολογίσουμε τις τιμές της* $v(\cdot,.)$ *για* 10×20 *σημεία. Αν όμως* $S = \mathbb{R}^k$, *και προσεγγίσουμε κάθε διάσταση του* $\mathbb{R}^k$ *με 20 σημεία, τότε χρειάζεται να υπολογίσουμε την* $v(\cdot,.)$ *για* $10 \times (20)^k$ *τιμές.*
Για $k = 3$, *αυτός ο αριθμός είναι* 80.000, *για* $k = 5$ *είναι* $32.000.000$.

4. Εφαρμογές

4.1 Πρόβλημα Σχεδιασμού Παραγωγής–Αποθεμάτων. *Μια εταιρεία σκοπεύει να αναπτύξει μία πολιτική για την παραγωγή κάποιου προιόντος για τις επόμενες n χρονικές περιόδους. Η ζήτηση για το προιόν είναι διαφορετική απο περίοδο σε περίοδο, και είναι πιο οικονομικό για την εταιρεία να παράγει το προιόν σε μαγάλες ποσότητες. Γι αυτό θα ήταν πιο αποδοτικό να παράγει παραπάνω απο τη ζήτηση σε ορισμένες περιόδους και να αποθηκεύσει το πλεόνασμα για να καλύψει τη ζήτηση αργότερα. Η διατήρηση αποθεμάτων όμως κοστίζει, πράγμα που οφείλεται σε παράγοντες, όπως η συντήρηση των αποθηκών, δαπάνες ασφάλισης, αλλά κυρίως η χρηματική αξία του δεσμευμένου κεφαλαίου. Ο στόχος της εταιρείας είναι να καθορίσει μια πολιτική που ελαχιστοποιεί το συνολικό κόστος παραγωγής και αποθεμάτων, και που ικανοποιεί τη ζήτηση για όλες τις χρονικές περιόδους. Τα δεδομένα του προβλήματος είναι:*

1) $n = 4$ *(Χειμώνας, Ανοιξη, Καλοκαίρι, Φθινόπωρο)*
2) *Η ζήτηση είναι σταθερή* $d_t = 3$, $t = 1,2,3,4$.
3) *Το κόστος παραγωγής* α *μονάδων προιόντος δίνεται από τον παρακάτω πίνακα,*

α	0	1	2	3	4	5
$c(\alpha)$	0	15	17	19	21	23

όπου 5 μονάδες είναι η μέγιστη ποσότητα που είναι δυνατόν η εταιρεία να παράγει σε μια περίοδο.

4) *Το κόστος διατήρησης αποθεμάτων είναι 1 ανά μονάδα αποθεματικού.*
5) *Η χωρητικότητα της αποθήκης της εταιρείας είναι 4 μονάδες.*

Διατύπωση του Προβλήματος.

1. Καταστάσεις – Αποφάσεις. *Εστω*

$\alpha_t :=$ *ποσότητα παραγωγής στην ααρχή της περιόδου* t, $t = 1,\ldots,4$, *και*
$x_t :=$ *απόθεμα κατά το τέλος της περιόδου* t.

2. Η δυναμική του συστήματος *περιγράφεται ως εξής*

1) $x_{t+1} = x_t + \alpha_t - d_t$.

2) $5 \geq \alpha_t \geq 0$, $\quad x_t + \alpha_t - d_t \leq 4$,

Δηλαδή : $D(x_t) := [0, \min\{5, 4+d_t - x_t\}]$

3. Δομή Κόστους. *i) Η συνάρτηση τρέχοντος κόστους ορίζεται ως εξής*

$$c_t(x_t,\alpha_t) = 1 \cdot x_t + c(\alpha_t),$$

όπου η $c(\cdot)$ *δίνεται από τον παραπάνω πίνακα.*

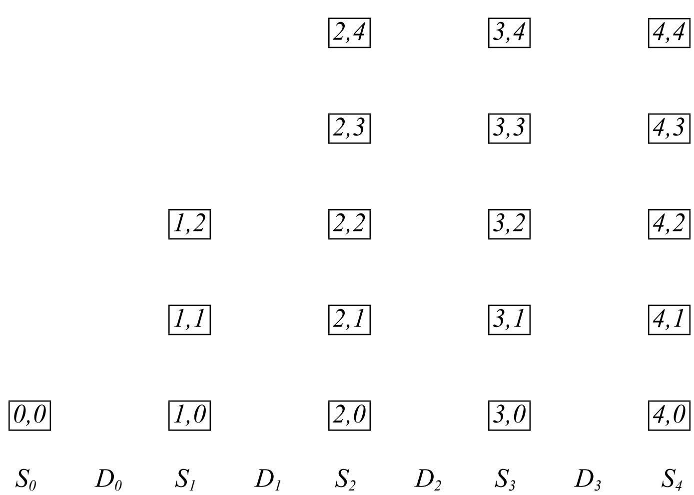
ii) Δεχόμαστε ότι η ποσότητα αποθεμάτων στο τέλος του ορίζοντα προγραμματισμού πρέπει να είναι 0. *Γι' αυτό θεωρούμε τη συνάρτηση τερματικού κόστους*

$$\hat{c}(x) = \begin{cases} M, & x > 0 \\ 0, & x = 0 \end{cases},$$

όπου M *είναι έναςμεγάλος αριθμός.*

4. *Υποθέτουμε ότι το αρχικό απόθεμα είναι* 0, *δηλ.* $S_0 = \{0\}$.

Βλέπουμε λοιπόν, ότι το πρόβλημα σχεδιασμού της παραγωγής μπορεί να περιγραφεί μέσω του παρακάτω δικτύου.



Εικόνα 6

Στο παραπάνω δίκτυο αναφέρονται ενδεικτικά οι αποφάσεις (επίπεδο παραγωγής) που αντιστοιχούν στις καταστάσεις $(0,0)$ *και* $(1,2)$, *καθώς και το κόστος που αντιστοιχεί σε κάθε απόφαση. Ετσι π.χ. έχουμε:*

$$c_1((1, 2),(2,3)) = c(4) + 1 \cdot 2 = 21 + 2 = 23$$

Οι περιορισμοί για το μέγεθος μέγιστης παραγωγής και αποθήκευσης δίδονται απο τα σύνολα $D(t,x)$. *Το τερματικό κόστος για τους κόμβους* $(4,x)$ *είναι* M *αν* $x \neq 0$, *και* 0 *αν* $x = 0$.

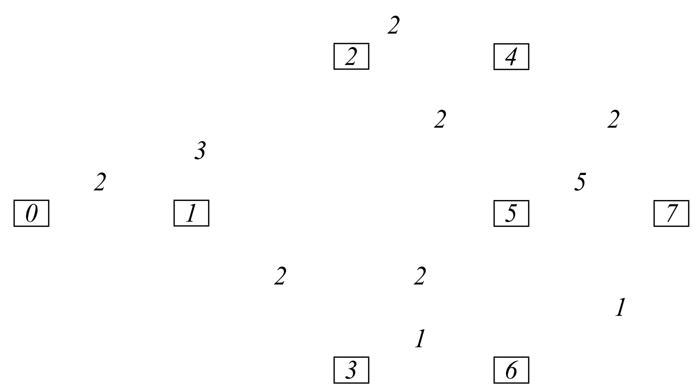
Ασκηση *Να συμπληρωθεί το δίκτυο του παραπάνω προβλήματος και να βρεθεί η βέλτιστη πολιτική.*

Παρατήρηση. *Η ερμηνεία που σύμφωνα με τον ορισμό γ(i) δίδεται στο κόστος αποθήκευσης, είναι ότι το απόθεμα που δημιουργείται κατά την t περίοδο συνεισφέρει στο κόστος αποθεμάτων για την περίοδο t+1 . Αν θέλαμε τούτο το κόστος να εμφανίζεται στο κόστος της περιόδου t , τότε η συνάρτηση κόστους θα έπρεπε να τροποποιηθεί σε*

$$c_t(x_t,\alpha_t) = (x_t+\alpha_t - 3)^+ + c(\alpha_t),$$

όπου $\alpha^+ = 0$, *αν* $\alpha \leq 0$, *και* $\alpha^+ = \alpha$, *αν* $\alpha > 0$.

4.2. Η Μέθοδος PERT/CPM. *Η μέθοδος PERT/CPM είναι συστηματικός τρόπος για το σχεδιασμό της εκτέλεσης, και την παρακολούθηση της προόδου των εργασιών έργων. Για παράδειγμα, ας θεωρήσουμε το σχεδιασμό της κατασκευής μιας μονοκατοικίας, και ας υποθέσουμε ότι η κατασκευή τελειώνει όταν μια σειρά επι μέρους κατασκευών (π.χ. θεμέλια, σκελετός, τοίχοι, ηλεκτρικά, υδραυλικά κ.λ.π.) τελειώσουν. Μερικές απο αυτές τις κατασκευές είναι δυνατόν να γίνουν ταυτόχρονα, ενω άλλες όχι. Γνωρίζουμε τους χρόνους που απαιτούνται για να τελειώσουν όλες οι επιμέρους κατασκευές, και θέλουμε να υπολογίσουμε τον χρόνο που θα απαιτηθεί για την κατασκευή του όλου έργου. Η παραπάνω διαδικασία εκτέλεσης έργου μπορεί να περιγραφεί με ένα δίκτυο, π.χ. αυτό της εικόνας 7 , όπου κάθε βέλος συμβολίζει μια επιμέρους κατασκευή.*



Εικόνα 7

Ο αριθμός που αντιστοιχεί σε κάθε βέλος είναι ο χρόνος που απαιτείται για να τελειώσει η αντίστοιχη επιμέρους κατασκευή. Επίσης οι κόμβοι συμβολίζουν το γεγονός ότι επιμέρους εργασίες δεν μπορούν να αρχίσουν πριν τελειώνουν ορισμένες άλλες εργασίες. Π.χ. οι εργασίες (2,4) , (2,5) δεν μπορούν να αρχίσουν πριν οι εργασίες (0,1) και (1,2) να έχουν τελειώσει. Ο χρόνος που απαιτείται λοιπόν για την όλη κατασκευή, είναι ο χρόνος κατά τον οποίο η πιο καθυστερημένη εργασία θα τελειώσει, και είναι ίσος με το μήκοστης διαδρομής μεγίστου μήκους από τον κόμβο 0 στον κόμβο 7 . Η διαδρομή αυτή είναι η (0,1), (1,2) , (2,5) , (5,7) στο παράδειγμά μας. Η διαδρομή μέγιστου μήκους ονομάζεται ***κρίσιμος δρόμος*** *(critical path) και αποτελείται από όλες τις επιμέρους εργασίες, που αν μια απο αυτές καθυστερήσει, τότε η συνολική κατασκευή θα καθυστερήσει.*

Είναι εύκολο να τροποποιήσουμε τις μεθόδους που έχουμε μάθει, έτσι ώστε να βρούμε τη (τις) διαδρομή (ες) μέγιστου μήκους σε δίκτυα.

Παρατηρούμε λοιπόν, ότι οι μέθοδοι του Δυναμικού Προγραμματισμού είναι χρήσιμες για τη λύση προβλημάτων που δεν είναι προβλήματα αποφάσεων.

4.3. Πρόβλημα Αντικατάστασης–Συντήρησης Μηχανημάτων Ι. *Μία απλή μορφή του προβλήματος έχει ως εξής. Υποθέτουμε ότι έχουμε ένα μηχάνημα, του οποίου η αξία ελαττώνεται όσο περνά ο χρόνος, και πρέπει να αποφασίσουμε πότε να το αντικαταστήσουμε. Το μηχάνημα μας χρειάζεται για τις επόμενες N χρονικές περιόδους, και μας δίδονται τα εξής δεδομένα.*

(1) $c(x) :=$ *κόστος λειτουργίας ενός μηχανήματος που έχει ηλικία x , για μια χρονική περίοδο.*

(2) $\mu(x) :=$ *τιμή του μεταχειρισμένου μηχανήματος ηλικίας* x, *αν πουληθεί σαν μεταχειρισμένο.*

(3) $T :=$ *τιμή ενός καινούργιου μηχανήματος.*

(4) $H :=$ *ανώτατο όριο ηλικίας του μηχανήματος, όταν το μηχάνημα έχει ηλικία* H, *πρέπει να αντικατασταθεί υποχρεωτικά.*

(5) $x_o := 0$, *υποθέτουμε ότι ξεκινάμε με ένα καινούργιο μηχάνημα.*

Διατύπωση του Προβλήματος.

1. Καταστάσεις–Αποφάσεις. *i) Είναι εύκολο να δούμε ότι η κατάσταση μπορεί να περιγραφεί απο ένα ζεύγος* (t,x), *όπου* t *συμβολίζει την παρούσα χρονική περίοδο (δηλαδή απομένουν* $N-t$ *περίοδοι που μας χρειάζεται το μηχάνημα) και* x *είναι η ηλικία του μηχανήματος.*

ii) Το σύνολο των αποφάσεων που διαθέτουμε σε κάθε κατάσταση (t,x), $t = 0,1,\ldots,N-1$, *συμβολίζεται με* $D = \{0, 1\}$, *όπου* $\alpha_t = 0$ *σημαίνει αντικατάσταση του μηχανήματος, και* $\alpha_t = 1$ *σημαίνει συντήρηση. Η απόφαση* $\alpha = 1$ *είναι επιτρεπτή μόνο αν* $x_t \leq H-1$. *Δηλαδή* $D(x) = D$ *αν* $x_t \leq H-1$, $D(x) = \{0\}$ *αν* $x_t = H$.

2. Η Δυναμική του συστήματος *περιγράφεται ως εξής*

$$x_{t+1} = g(x_t,\alpha_t) = \begin{cases} 1, & \text{αν } \alpha_t=0 \\ x_t+1, & \text{αν } \alpha_t=1. \end{cases}$$

3. Δομή Κόστους. *i) Συνάρτηση τρέχοντος κόστους*

$$c_t(x_t,\alpha_t) = \begin{cases} T + c(o) - \mu(x), & \text{αν } \alpha_t = 0 \\ c(x), & \text{αν } \alpha_t = 1 \end{cases}$$

ii) Συνάρτηση Τερματικού κόστους

$$\hat{c}(x_N) = -\mu(x_N)$$

Μπορούμε τώρα να γράψουμε την εξίσωση για τη συνάρτηση τιμής του προβλήματος.

$$v(t,x) = \min\{T + c(0) - \mu(x) + v(t+1,1),\ c(x) + v(t+1,x+1)\}, \qquad \text{αν } x \leq H-1,\ t \leq N-1,$$

$$v(t,x) = T + c(0) - \mu(x), \qquad \text{αν } x = H, \quad t \leq N-1,$$

$$v(N,x) = -\mu(x), \qquad x = 1,2,\ldots,H.$$

4.4. Πρόβλημα Αντικατάστασης–Συντήρησης ΙΙ. *Σε τούτη την παράγραφο αυξάνουμε την πολυπλοκότητα (και τη ρεαλιστικότητα) του προβλήματος Αντικατάστασης–Συντήρησης Ι. Συγκεκριμένα, "μοντελοποιούμε" την περίπτωση, κατά την οποία σε κάθε χρονική περίοδο υπάρχει μια επιπλέον απόφαση, αυτή της γενικής επισκευής. Επιπλέον, δεχόμαστε ότι το κόστος λειτουργίας του μηχανήματος εξαρτάται απο την ηλικία του, και απο το πόσος χρόνος έχει περάσει απο την τελευταία γενική επισκευή, και όχι απο τον αριθμό των γενικών επισκευών που έχει υποστεί. Τα δεδομένα του προβλήματος είναι τα εξής*

1. $T(t,x,y) :=$ *κόστος ανταλλαγής ενός μηχανήματος ηλικίας* x, *στο οποίο έγινε γενική επισκευή πριν από* y *μονάδες χρόνου, με ένα καινούργιο, κατά την περίοδο* t.

2. $c(x,y) :=$ *κόστος λειτουργίας ενός μηχανήματος ηλικίας* x, *στο οποίο έγινε (η τελευταία) γενική επισκευή πριν από* y *μονάδες χρόνου.*

3. $\epsilon(x) :=$ *κόστος γενικής επισκευής ενός μηχανήματος ηλικίας* x.

4. $s(x,y) :=$ *τιμή πώλησης ενός μηχανήματος ηλικίας* x, *που του έγινε γενική επισκευή πριν από* y *μονάδες χρόνου, στο τέλος της περιόδου προγραμματισμού, ή ορίζοντα (salvage value).*

5) $H :=$ *μέγιστη επιτρεπτή ηλικία του μηχανήματος.*

Διατύπωση του Προβλήματος.

1. Καταστάσεις - Αποφάσεις. *i) Η κατάσταση του συστήματος περιγράφεται απο την τριάδα* (t,x,y), *όπου* t *συμβολίζει την χρονική περίοδο,* x *την ηλικία του μηχανήματος,* y *την ηλικία που είχε το μηχάνημα όταν του έγινε η τελευταία γενική επισκευή.*

ii) Σε κάθε κατάσταση (t,x,y), $(t=0,\ldots, N-1)$, *διαθέτουμε τις αποφάσεις* $\alpha = 0$: *ανταλαγή με καινούργιο,* $\alpha = 1$: *συντήρηση, και* $\alpha = 2$: *γενική επισκευή,*
όπου οι αποφάσεις $\alpha=1$ *και* $\alpha=2$ *επιτρέπονται μόνο όταν* $x \le H-1$.

2. Η Δυναμική του συστήματος *περιγράφεται σχηματικά ως εξής:*

$$\boxed{(t, x, y)} \begin{cases} \xrightarrow{\alpha=0} \boxed{(t+1, 1, 0)} \\ \xrightarrow{\alpha=1} \boxed{(t+1, x+1, y+\delta(y))}, \qquad \text{αν } H > x \\ \xrightarrow{\alpha=2} \boxed{(t+1, x+1, 1)} \end{cases}$$

όπου $\delta(y) = 1$, *αν* $y \neq 0$, *και* $\delta(y) = 0$, *αν* $y = 0$.

3. Δομή Κόστους. *i) Συνάρτηση Τρέχοντος Κόστους:*

$$c_t(x_t,y_t,\alpha_t) = \begin{cases} T(t,x_t, y_t) + c(0,0), & \alpha = 0 \\ c(x_t, y_t), & \alpha = 1 \\ \epsilon(x_t), & \alpha = 2 \end{cases}$$

ii) Συνάρτηση Τερματικού Κόστους:

$$\hat{c}(x_N, y_N) = s(x_N, y_N)$$

Η εξίσωση του δυναμικού προγραμματισμού για την συνάρτηση τιμής του προβλήματος είναι η εξής:

$$v(t,x,y) = \min\Big\{T(t,x,y)+c(0,0)+v(t+1, 1, 0),$$
$$c(x,y)+v(t+1,x+1,y+\delta(y)),$$
$$\epsilon(x)+c(x,0)+v(t+1,x+1,1)\Big\}, \qquad \text{αν } x \le H-1,\ t \le N-1,$$

$v(t,x,y) = T(t,x,y) + c(0,0)+v(t+1,1,0)$, αν $x = H$, $t \le N-1$,

$v(N,x,y) = -s(x,y)$.

Παρατήρηση. *Μπορούμε, βέβαια, να υποθέσουμε ότι αρχικά το μηχάνημα είναι καινούργιο,* $S_0 = \{(0,0,0)\}$ *. Η δομή των παραπάνω εξισώσεων δεν θα αλλάξει αν θεωρήσουμε ότι ξεκινάμε με ένα μηχάνημα στη κατάσταση* (x_0,y_0) *, οπότε* $S_0 = \{(0, x_0,y_0)\}$.

4.5 Πρόβλημα Κατανομής Πόρων (Σύγκριση Μεθόδων Δυναμικού Προγραμματισμού – Διαφορικού Λογισμού). *Μία απλή μορφή του προβλήματος έχει ως εξής. Υπάρχουν B μονάδες κάποιου πλουτοπαραγωγικού πόρου, π.χ. κεφάλαιο, ώρες εργασίας, καλιεργήσιμος χώρος κ.λ.π., οι οποίες πρέπει να διατεθούν (κατανεμηθούν) σε N δραστηριότητες. Εάν διατεθούν y μονάδες του πόρου στην δραστηριότητα* k *,* $k = 0,1,\dots,N-1$ *, παίρνουμε την απόδοση (π.χ. κέρδος, ωφέλεια, κ.λ.π.)* $R_k(x)$ *, όπου* $R_k(\cdot)$ *είναι γνωστές, αύξουσες συναρτήσεις. Το ζητούμενο είναι να βρούμε την κατανομή των B μονάδων στις διάφορες δραστηριότητες, που μεγιστοποιεί την συνολική απόδοση. Συμβολικά το πρόβλημα μπορεί να γραφτεί ως εξής*

$$(max) \qquad \sum_{k=0}^{N-1} R_k(y_k), \tag{1}$$

$$\text{όπου} \qquad \sum_{k=0}^{N-1} y_k \le B, \tag{2}$$

$$y_k \ge 0\,,\ k = 0, 1,\dots, N-1\,. \tag{3}$$

Το παραπάνω πρόβλημα σε πρώτη ανάγνωση φαίνεται να είναι πρόβλημα μιας μόνο απόφασης (της $\vec{y} = (y_0,\dots,y_{N-1})$*), για τη λύση του οποίου μπορούν να χρησιμοποιηθούν γνωστές μέθοδοι βελτιστοποίησης. Για παράδειγμα, όταν οι συναρτήσεις* $R_k(\cdot)$ *είναι γραμμικές, τότε έχουμε ένα απλό πρόβλημα γραμμικού προγραμματισμού, ενώ άν οι συναρτήσεις* $R_k(\cdot)$ *είναι μη γραμμικές, τότε μπορούμε να το θεωρήσουμε σαν ένα πρόβλημα μη γραμμικού προγραμματισμού. Εαν όμως οι μεταβλητές του προβλήματος,* y_k *, περιορίζονται σε ακέραιες μόνο τιμές, τότε το πρόβλημα γίνεται πρόβλημα ακέραιου γραμμικού (ή μη γραμμικού) προγραμματισμού για το οποίο ικανοποιητικές μέθοδοι λύσης δεν έχουν βρεθεί ακόμη. Εν τούτοις, η μεθοδολογία του Δυναμικού Προγραμματισμού μπορεί να χρησιμοποιηθεί για την λύση του παραπάνω προβλήματος.*

Διατύπωση του Προβλήματος.

1. Καταστάσεις - αποφάσεις. *i) Το πρόβλημα μπορεί να διατυπωθεί στη γλώσσα του Δυναμικού Προγραμματισμού, αν θεωρήσουμε ότι κάνουμε την κατανομή* ***διαδοχικά****, πρώτα στην δραστηριότητα 1 , μετά στην 2 , κ.λ.π.. Εστω λοιπόν* x_t *η ποσότητα που* ***απομένει*** *για κατανομή στις δραστηριότητες* $t, t+1, \dots, N-1$ *. Η μεταβλητή* x_t *περιγράφει την κατάσταση του συστήματος, και* $x_t \in [0,B]$ *.*

ii) Η ***απόφαση*** α_t *είναι η ποσότητα που διαθέτουμε στη δραστηριότητα* t *και πρέπει* $\alpha_t \le x_t$ *, δηλ.* $D(x_t) = [0, x_t]$.

2. Δυναμική του συστήματος. *Είναι εύκολο να δούμε ότι*

$$x_{t+1} = x_t - \alpha_t\,,\quad t = 0, 1, \dots, N-1\,. \tag{4}$$

Η αρχική κατάσταση είναι η $x_0 = B$.

3. Δομή Κόστους. *i) Η συνάρτηση τρέχοντος κέρδους ορίζεται ως εξής*

$$c_t(x_t,\alpha_t) = R_t(\alpha_t)\,. \tag{5}$$

ii) Η συνάρτηση τερματικού κέρδους είναι η

$$\hat{c}(x_n) = R(x_n)\,. \tag{6}$$

Η εξίσωση του δυναμικού προγραμματισμού για την συνάρτηση τιμής είναι η εξής

$$v(t,x) = \max_{0 \le \alpha \le x} \{R_t(\alpha) + v(t+1, x-\alpha)\}\,, \quad t = 0, 1, \ldots, N-1\,.\,. \tag{7}$$

$$v(N,x) = 0\,.\,. \tag{8}$$

Η λύση που ζητάμε είναι η $v(0,B)$.

Ειδικές Περιπτώσεις. *Εαν οι ποσότητες που κατανέμονται στις διάφορες δραστηριότητες πρέπει να είναι ακέραιοι αριθμοί, τότε* $D(x) = \{0,1,2,...,x\}$ *, και λύνουμε αριθμητικά την εξίσωση βελτιστοποίησης με τις μεθόδους που έχουμε ήδη μάθει.*

Οταν οι ποσότητες που κατανέμονται στις δραστηριότητες επιτρέπεται να παίρνουν και μη ακέραιες τιμές, τότε $D(x) = [0,x]$ *. Πάλι μπορούμε να λύσουμε αριθμητικά την εξίσωση βελτιστοποίησης, αν προσεγγίσουμε τα σύνολα* $D(x)$ *με πεπερασμένα σύνολα της μορφής* $\hat{D}(x) = \{0, z_1, z_2, \ldots, z_n = x\}$.

Υπάρχουν δύο ειδικές περιπτώσεις, στις οποίες μπορούμε να λύσουμε την εξίσωση βελτιστοποίησης αναλυτικά. Τούτες οι περιπτώσεις είναι σημαντικές, γιατί ανάλογο φαινόμενο παρουσιάζεται συχνά σε προβλήματα Δυναμικού Προγραμματισμού. Παρακάτω λοιπόν εξετάζουμε σε λεπτομέρεια τις περιπτώσεις κατά τις οποίες οι συναρτήσεις απόδοσης $R_k(x)$ *είναι όλες* ***κοίλες*** *ή όλες* ***κυρτές.***

Εστω, λοιπόν, $D(x) = [0,x]$ *. Πριν προχωρήσουμε ας θεωρήσουμε την εξής ερώτηση. Υπάρχει βέλτιστη λύση για το πρόβλημα; Η απάντηση ειναι καταφατική αν υποθέσουμε ότι οι συναρτήσεις* $R_k(\cdot)$ *είναι συνεχείς* $\forall k = 1, \ldots, N-1$ *. Ο λόγος είναι ότι τότε η συνάρτηση* $\sum_0^{N-1} R_k(y_k)$ *είναι συνεχής, και μια συνεχής συνάρτηση πάνω σε ένα συμπαγές σύνολο φτάνει τις μέγιστες και ελάχιστες τιμές της. Το σύνολο δε των περιορισμών* $\sum_0^{N-1} y_k \le B$, $y_k \ge 0$ *, είναι συμπαγές.*

4.6. Κοίλες Συναρτήσεις Κέρδους. *Ας υποθέσουμε ότι* $R_k(y) = R(y)$*, όπου η συνάρτηση* R *είναι κοίλη και αύξουσα. Ας υποθέσουμε επιπλέον ότι η συνάρτηση* R *έχει δεύτερη παράγωγο παντού. Ετσι, οι απαιτήσεις ότι η* R *είναι αύξουσα και κοίλη είναι ισοδύναμες με τις παρακάτω απαιτήσεις*

$$\frac{\partial R(y)}{\partial y} \ge 0 \;\; \text{και} \;\; \frac{\partial^2 R(y)}{\partial y^2} \le 0 \quad \forall y\,.$$

Λύση της εξίσωσης βελτιστοποίησης

1. Τερματική συνθήκη

$$v(N,x) = 0\,. \tag{1}$$

2. Για $t = N-1$ *έχουμε*

$$v(N-1,x) = \max_{\alpha \in [0,x]} \{R(\alpha) + v(N, x-\alpha)\} = \max_{\alpha \in [0,x]} \{R(\alpha)\} = R(x) \tag{2}$$

όπου η τελευταία ισότητα οφείλεται στην υπόθεση ότι η R *είναι αύξουσα. Ετσι, έχουμε*

$$\alpha^* = \alpha^*(N-1,x) = x \in [0,x]. \tag{3}$$

Τώρα από την (2) παίρνουμε

$$v(N-2,x) = \max_{\alpha \in [0,x]} \{R(\alpha)+v(n,x-\alpha)\} = \text{(λόγω της (1))}$$

$$= \max_{\alpha \in [0,x]} \{ R(\alpha) + R(x-\alpha) \} . \tag{4}$$

Ας θεωρήσουμε τη συνάρτηση

$$h(\alpha) = R(x) + R(x-\alpha), \ \alpha \in [0,x] . \tag{5}$$

Θέλουμε να μεγιστοποιήσουμε την συνάρτηση $h(\alpha)$ *πάνω στο σύνολο* $[0,x]$. *Εχουμε*

$$h'(\alpha) = R'(\alpha) - R'(x-\alpha) \tag{6}$$

$$h''(\alpha) = R''(\alpha) + R''(x-\alpha) \tag{7}$$

Τώρα $h''(\alpha) \le 0 \ \forall\alpha$, *λόγω του ότι* $R''(\alpha) \le 0 \ \forall\alpha$, $R''(x-\alpha) \le 0 \ \forall\alpha$, *επειδή η* R *είναι κοίλη. Αρα, το σημείο* α^* *στο οποίο* $h'(\alpha^*) = 0$, *είναι σημείο στο οποίο η* $h(\alpha)$ *φτάνει τη μέγιστη τιμή της στο* $[0,x]$. *Εχουμε λοιπόν*

$$h'(\alpha^*) = 0 \quad \Leftrightarrow \quad R'(\alpha^*) = R'(x-\alpha^*) \quad \Leftrightarrow \ \alpha^* = x-\alpha^* \quad \Leftrightarrow$$

$$\alpha^* = \alpha^*(N-2,x) = x/2 \in [0,x]. \tag{8}$$

Αρα

$$v(N-2,x) = h(\alpha^*) = R(x/2) + R(x- x/2) \quad \Rightarrow$$

$$v(N-2,x) = 2R(x/2) \tag{9}$$

Ανάλογα

$$v(N-3,x) = \max_{\alpha \in [0,x]} \{R(\alpha) + v(N-2,x-\alpha)\} = \text{(λόγω της (9))}$$

$$= \max_{\alpha \in [0,x]} \left\{R(\alpha) + 2R(\tfrac{x-\alpha}{2})\right\} \tag{10}$$

Πάλι θέτουμε

$$h(\alpha) = R(\alpha) + 2R(\tfrac{x-\alpha}{2}), \tag{11}$$

και έχουμε

$$h'(\alpha) = R'(\alpha) - 2R'(\tfrac{x-\alpha}{2}) \cdot \tfrac{1}{2} \tag{12}$$

$$h''(\alpha) = R''(\alpha) + R''(\tfrac{x-\alpha}{2}) \cdot \tfrac{1}{2} \le 0, \ \forall\alpha, \tag{13}$$

άρα η h *είναι κοίλη, και το σημείο* $\alpha^*: \ h'(\alpha^*) = 0$, *είναι σημείο στο οποίο η* h *φτάνει τη μέγιστη τιμή της. Τώρα,*

$$h'(\alpha^*) = 0 \quad \Leftrightarrow \quad R'(\alpha^*) = R'(\tfrac{x-\alpha^*}{2}) \quad \Rightarrow$$

$$\alpha^* = \alpha^*(N-3,x) = \tfrac{x}{3} . \tag{14}$$

Αρα

$$v(N-3,x) = R(\tfrac{x}{3}) + 2R(\tfrac{x-x/3}{2}) = R(\tfrac{x}{3}) + 2R(\tfrac{x}{3}) = 3R(\tfrac{x}{3}) . \tag{15}$$

Ας κάνουμε την επαγωγική υπόθεση ότι

$$v(N-t,x) \;=\; t\,R(\tfrac{x}{t}), \qquad (16)$$

Τότε

$$v(N-t-1,x) \;=\; \max_{\alpha\in[0,x]} \{R(\alpha)+v(N-t,x)\} \;=$$

$$= \max_{\alpha\in[0,x]} \left\{R(\alpha)+t\,R(\tfrac{x}{t})\right\}. \qquad (17)$$

Ετσι παίρνουμε, ακολουθώντας ανάλογα βήματα με τα παραπάνω

$$v(N-t-1,x) \;=\; (t+1)\,R(\tfrac{x}{t+1}), \quad \text{και}\;\; \alpha^* \;=\; \alpha^*(N-t-1,x) \;=\; \tfrac{x}{t+1}. \qquad (18)$$

Εχουμε αποδείξει επαγωγικά ότι

$$v(N-t,x) \;=\; t\;R(\tfrac{x}{t}), \qquad \alpha^* \;=\; \alpha^*(N-t,x) \;=\; \tfrac{x}{t}. \qquad (19)$$

Ετσι, η λύση που ζητούμε είναι

$$v(0,B) \;=\; v(N-N,B) \;=\; N\,R(\tfrac{B}{N}), \qquad (20)$$

και

$$\alpha^*(0,B) \;=\; \alpha^*(N-N,B) \;=\; \tfrac{B}{N}. \qquad (21)$$

Τώρα, πηγαίνοντας μπροστά, μπορούμε να βρούμε τη βέλτιστη πολιτική και την αντίστοιχη τροχιά ως εξής:

$$x^*(0) \;=\; B, \qquad \alpha^*(0,B) \;=\; \tfrac{B}{N}$$

$$x^*(1) \;=\; x^*(0)-\alpha^*(0,B) \;=\; B-\tfrac{B}{N} \;=\; \tfrac{N-1}{N}B, \qquad \alpha^*(1,\tfrac{N-1}{N}B) = \frac{\frac{N-1}{N}B}{N-1} \;=\; \tfrac{B}{N}$$

$$x^*(2) \;=\; x^*(1)-\alpha^*(1,x^*(1)) = \tfrac{N-1}{N}B-\tfrac{B}{N} = \tfrac{N-2}{N}B, \qquad \alpha^*(2,\tfrac{N-2}{N}B) = \tfrac{B}{N}$$

$$\vdots \qquad\qquad \vdots$$

$$x^*(t) \;=\; x^*(t-1)-\alpha^*(t-1,x^*(t-1)) = \tfrac{N-t+1}{N}B, \qquad \alpha^*(t,x^*(t)) \;=\; \tfrac{B}{N}.$$

Παρατηρούμε , λοιπόν, ότι όταν οι συναρτήσεις $R(\cdot)$ είναι κοίλες, τότε η βέλτιστη κατανομή αποδίδει ίσες ποσότητες πόρων στις διάφορες δραστηριότητες.

4.7. Κυρτές Συναρτήσεις Κέρδους. *Υποθέτουμε ότι $R_k(y) = R(y)$, όπου η συνάρτηση R είναι κυρτή και αύξουσα. Υποθέτουμε επιπλέον ότι η συνάρτηση R έχει δεύτερη παράγωγο παντού. Ετσι, οι απαιτήσεις ότι η R είναι αύξουσα και κυρτή είναι ισοδύναμες με τις παρακάτω απαιτήσεις*

$$\frac{\partial R(y)}{\partial y} \;\geq\; 0 \;\;\text{και}\;\; \frac{\partial^2 R(y)}{\partial y^2} \;\geq\; 0 \quad \forall\, y.$$

Λύση της εξίσωσης βελτιστοποίησης

1) Τερματική συνθήκη

$$v(N,x) \;=\; 0. \qquad (1)$$

$$2)\;\; v(N-1,x) \;=\; \max_{\alpha\in[0,x]}\{R(\alpha)\} \;=\; R(x), \qquad \alpha^* = \alpha^*(N-1,x) = x. \qquad (2)$$

όπου η δεύτερη ισότητα παραπάνω οφείλεται στην υπόθεση ότι η R είναι αύξουσα. Συνεχίζοντας έχουμε

$$v(N-2,x) \;=\; \max_{\alpha \in [0,x]}\{R(\alpha) + v(N\text{-}1,x-\alpha)\} \overset{(2)}{=} \max_{\alpha \in [0,x]}\{R(\alpha) + R(x-\alpha)\}\,. \tag{3}$$

Ας ορίσουμε πάλι τη βοηθητική συνάρτηση $h(\alpha) \;=\; R(\alpha) + R(x-\alpha)$. Εχουμε

$$h'(\alpha) \;=\; R'(\alpha) \;-\; R'(x-\alpha), \tag{4}$$

και

$$h''(\alpha) \;=\; R''(\alpha) + R''(x-\alpha)\,. \tag{5}$$

Τώρα, λόγω του ότι η συνάρτηση R είναι κυρτή, δηλ. $R''(x) \;\geq\; 0$, έχουμε ότι $h''(\alpha) \;\geq\; 0 \;\forall\alpha$, και συνεπώς το σημείο στο οποίο ισχύει $h'(\alpha^*) \;=\; 0$, είναι σημείο στο οποίο η $h(\cdot)$ παίρνει την ελάχιστη τιμή της στο διάστημα $[0,x]$. Αρα η μέγιστη τιμής της $h(\cdot)$ επιτυγχάνεται στα άκρα του διαστήματος $[0,x]$, δηλαδή

$$\begin{aligned} \max_{\alpha \in [0,x]} h(\alpha) \;&=\; \max\,\{h(0)\,,\, h(x)\} \;= \\ &=\; \max\,\{R(0) + R(x)\,,\, R(x) + R(0)\} \;=\; R(0) + R(x)\,. \end{aligned} \tag{6}$$

Εχουμε λοιπόν

$$v(N-2,x) \;=\; R(0) + R(x), \qquad \text{και } \;\alpha^* = \alpha^*(N-2,x) = \;0 \text{ ή } x\,. \tag{7}$$

Επαγωγικά λοιπόν, είναι εύκολο να δούμε ότι:

$$v(N-t,x) \;=\; (t-1)\,R(0) + R(x), \qquad \text{και } \;\alpha^* \;=\; \alpha^*(N-t,x) \;=\; 0 \text{ ή } x. \tag{8}$$

Ετσι παίρνουμε

$$v(0,B) \;=\; (N-1)\,R(0) + R(x) \qquad \text{και } \;\alpha^* \;=\; \alpha^*(0,B) \;=\; 0 \text{ ή } B. \tag{9}$$

Παρατηρούμε λοιπόν πως, όταν οι συναρτήσεις R είναι κυρτές, τότε η βέλτιστη πολιτική κατανέμει όλη την ποσότητα του πλουτοπαραγωγικού πόρου σε μια μόνο δραστηριότητα.

Παρατήρηση. *Είναι δυνατόν, το πρόβλημα της κατανομής πόρων της προηγούμενης παραγράφου, να λυθεί με τη λεγόμενη μέθοδο των πολλαπλασιαστών Lagrange, την οποία εισάγουμε αμέσως με ένα παράδειγμα.*

4.8 Η Μέθοδος τών Πολλαπλασιαστών Lagrange. *Ας θεωρήσουμε το πρόβλημα*

$$(\max) \qquad f(x,y) \;=\; 5x - 3x^2 + 6y - 2y^2$$

περιορισμός: $\qquad g(x,y) \;=\; x + y \;=\; 20.$

Ο περιορισμός $g(x,y) \;=\; 20$ είναι ισοδύναμος με τον $g(x,y) - 20 \;=\; 0$, και εάν πολλαπλασιάσουμε τον περιορισμό $g(x,y) - 20 \;=\; 0$ με κάποια άγνωστη παράμετρο λ και προσθέσουμε το αποτέλεσμα στην αντικειμενική συνάρτηση, παίρνουμε μια καινούργια αντικειμενική συνάρτηση (τη λεγόμενη Lagrangian του προβλήματος), που είναι η εξής.

$$L(x,y,\lambda) \;=\; 5x - 3x^2 + 6y - 2y^2 + \lambda(x + y - 20).$$

Παρατηρούμε ότι σε όλα τα σημεία (x,y), για τα οποία ο περιορισμός $g(x,y) - 20 = 0$ δεν παραβιάζεται, η τιμή της Lagrangian είναι ίση με την τιμή της αντικειμενικής συνάρτησης $f(x,y)$. Μπορούμε λοιπόν, ισοδύναμα, να θεωρήσουμε το πρόβλημα μεγιστοποίησης της $L(x,y,\lambda)$.

Για την συγκεκριμένη $L(x,y,\lambda)$ θέτουμε τις πρώτες παραγώγους ίσες με το 0 και παίρνουμε τις εξισώσεις

$$\frac{\partial L}{\partial x} = 5 - 6x + \lambda = 0$$

$$\frac{\partial L}{\partial x} = 5 - 4y + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = x + y - 20 = 0 .$$

Παρατηρούμε ότι αν (x^0,y^0,λ^0) είναι μια λύση στο παραπάνω σύστημα των εξισώσεων η οποία μεγιστοποιεί την Lagrangian, τότε το ζεύγος (x^0,y^0) έχει τις εξής ιδιότητες

1) *ικανοποιεί τον περιορισμό $x + y - 20 = 0$*
2) *μεγιστοποιεί την συνάρτηση $f(x,y)$.*

Για να συμπληρώσουμε την λύση του παραδείγματος, λύνουμε το παραπάνω σύστημα ως προς x,y,λ και βρίσκουμε:

$$x^0 = \frac{237}{30}, \quad y^0 = \frac{242}{20}, \quad \lambda^0 = \frac{212}{5} .$$

Επιπλέον, το σημείο (x_0,y_0) πληροί την ***αναγκαία*** *συνθήκη για μεγιστοποίηση της f: $h' \bigtriangledown^2 f(x^0,y^0) h < 0, \ \forall h \in \mathbb{R}^2$. Πράγματι,*

$$\bigtriangledown^2 f(x,y) = \begin{bmatrix} -6 & 0 \\ 0 & -4 \end{bmatrix}$$

και $$[\, h_1 \ \ h_2 \,] \cdot \begin{bmatrix} -6 & 0 \\ 0 & -4 \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -6h_1^2 - 4h_2^2 < 0 , \ \forall h \in \mathbb{R}^2 , \ h \neq (0,0).$$

4.9 Λύση του Προβλήματος Κατανομής Πόρων με την Μέθοδο των Πολλαπλασι-αστών Lagrange.

Εξετάζουμε την ειδική περίπτωση 1 . Η Lagrangian του προβλήματος είναι

$$L(x_1,y_2,\ldots,y_N,\lambda) = \sum_1^N R(y_k) + \lambda(\sum_1^N y_k - B)$$

Εχουμε λοιπόν

$$\frac{\partial L}{\partial y_k} = R'(y_k) + \lambda = 0 \qquad (1)$$

$$\frac{\partial L}{\partial \lambda} = \sum_1^N y_k - B = 0 , \qquad (2)$$

και η λύση των (1),(2) είναι $y_\kappa^* = \frac{B}{N}$, $\lambda^0 = -R'(y_k^*)$, Μπορούμε να δούμε ότι η αναγκαία συνθήκη βελτιστοποίησης της $f(y1,\ldots,y_n) = \sum_1^N R(y_k)$ ικανοποιείται,
επειδή οι συναρτήσεις $R(y_k)$ είναι κοίλες. Πράγματι

$$\nabla^2 f(x_1,\ldots,y_n) = \begin{bmatrix} -R^{(2)}(y_1) & 0 & \cdots & 0 \\ 0 & -R^{(2)}(y_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -R^{(2)}(y_n) \end{bmatrix},$$

!!FOOTNOTE 6:Με h' συμβλίζουμε γενικά το ανάστροφο διάνυσμα του h (διάνυσμα γραμμής) .
και έτσι $h' \cdot \nabla^2 f \cdot h < 0$[6] $\forall h \in \mathbb{R}^n$, $h \neq 0$, *επειδή* $-R^{(2)}(y_k) \leq 0$.

Παρατήρηση. *Οταν λύνουμε το πρόβλημα βέλτιστης κατανομής πόρων με την μέθοδο του Δυναμικού Προγραμματισμού, ουσιαστικά ανάγουμε την λύση ενός προβλήματος N–μεταβλητών στην διαδοχική λύση N προβλημάτων μιας μεταβλητής. Τούτο το φαινόμενο είναι τυπικό της μεθόδου του Δυναμικού Προγραμματισμού, και οδηγεί στα εξής πλεονεκτήματα της μεθόδου αυτής, σε σχέση με τις μεθόδους του Διαφορικού Λογισμού, όπως είναι η μέθοδος των πολλαπλασιαστών Lagrange.*

1) Σε πολύπλοκα προβλήματα η Lagrangian τυπικά οδηγεί σε συστήματα που είναι δύσκολο να λυθούν. Επιπλέον, για κάθε λύση του συστήματος

$$\frac{\partial L}{\partial y_k} = 0\,,\ \frac{\partial L}{\partial \lambda} = 0,$$

έχουμε την υποχρέωση να εξακριβώσουμε κατά πόσον η αναγκαία συνθήκη για μέγιστο ή ελάχιστο ισχύει. Η πολυπλοκότητα του προβλήματος αυξάνει όταν αυξηθεί ο αριθμός των περιορισμών.

2) Οι μέθοδοι του Διαφορικού λογισμού δεν επαρκούν στην περίπτωση κατά την οποία βέλτιστα σημεία βρίσκονται στο σύνορο της εφικτής περιοχής όπως. π.χ. στην ειδική περίπτωση 2 παραπάνω.

3) Οι μέθοδοι του Διαφορικού λογισμού δεν επαρκούν στην περίπτωση κατά την οποία η αντικειμενική συνάρτηση και (ή) οι περιορισμοί δεν είναι διαφορίσιμες συναρτήσεις. Σε αντίθεση, βασικά χαρακτηριστικά της μεθόδου του Δυναμικού Προγραμματισμού είναι τα παρακάτω.

1) Δεν απαιτούνται υποθέσεις για την ύπαρξη παραγώγων της αντικειμενικής συνάρτησης, ή/και των περιορισμών.
2) Οι περιορισμοί μπορούν να αντιμετωπισθούν εύκολα, αρκεί να είναι δυνατόν να εκφρασθούν με απαιτήσεις της μορφής

$$x_t \in S_t\,,\ \alpha_t \in D_t(x_t)\,.$$

3) Το μόνο πράγμα που είναι απαραίτητο να εξακριβώσουμε σε κάθε βήμα, είναι η ύπαρξη λύσης του προβλήματος $P(t,x)$, δηλαδή η ύπαρξη της τιμής $v(t,x)$.

4) Με την μέθοδο του Δυναμικού Προγραμματισμού βρίσκουμε πάντα απόλυτα βέλτιστα (μέγιστα ή ελάχιστα) σημεία, και όχι τοπικά βέλτιστα ή, ακόμη, χειρότερα σημεία "σέλλας" ("sadle points").

5) Παίρνουμε σαν λύση μια βέλτιστη πολιτική $\alpha^* = \alpha^*(t,x)$.

Εν τούτοις, η μέθοδος του Δυναμικού Προγραμματισμού έχει σαν κύριο μειονέκτημα τον πολύ χρόνο που απαιτείται για την αριθμητική λύση, ειδικά όταν οι χώροι καταστάσεων , S_t , είναι πολυδιάστατοι ("curse of dimensionality" κατά τον Bellman). Γι αυτό μερικές φορές, όταν τα δεδομένα του προβλήματος το επιτρέπουν, καταφεύγουμε σε συνδυασμό των δυο μεθόδων, όπως θα δούμε παρακάτω.

6

4.10. Γενίκευση του Προβλήματος Κατανομής Πόρων. *Εστω w_k, B, w_k^i, B_1, B_2 γνωστοί αριθμοί, $w_k \geq 0$, $w_k^i \geq 0$. Ας θεωρήσουε τα εξής προβλήματα.*

Πρόβλημα 1

$$(\max) \sum_{k=1}^{N} R_k(y_k)$$

περιορισμοί

$$\sum_{k=1}^{N} w_k y_k \leq B,$$

$$y_k \geq 0,$$

Πρόβλημα 2.

$$(\max) \sum_{k=1}^{N} R_k(y_k)$$

περιορισμοί

$$\sum_{k=1}^{N} w_k^1 y_k \leq B_1,$$

$$\sum_{k=1}^{N} w_k^2 y_k \leq B_2,$$

$$y_k \geq 0.$$

Τα παραπάνω προβλήματα ονομάζονται προβλήματα εφοδιασμού σάκκου (Knapsack–problem), και εμφανίζονται συχνά σε εφαρμογές της επιχειρησιακής έρευνας. Η ονομασία αυτή δικαιολογείται ως εξής: Υποθέτουμε ότι έχουμε ένα σάκκο (knapsack), τον οποίο θέλουμε να εφοδιάσουμε με ποσότητες $y_1, \ldots, y_n$, από N εφόδια. Ο σάκκος επιβάλει περιορισμούς βάρους (π.χ. B_1 είναι το μέγιστο επιτρεπτό βάρος των εφοδίων) και όγκου (B_2 συμβολίζει τον μέγιστο επιτρεπτό όγκο των εφοδίων). Οι συναρτήσεις $R_k(y_k)$ δίδουν την αξία της ποσότητας y_k του εφοδίου k για τον ιδιοκτήτη του σάκκου. Το πρόβλημα γίνεται πιο ρεαλιστικό αν φανταστούμε ότι ο σάκκος είναι ένα μεταφορικό μέσο, αλλά όπως θα δούμε αργότερα, υπάρχουν πολλά προβλήματα για τα οποία είναι κατάλληλη η παραπάνω διατύπωση, αν και εκ πρώτης όψεως δεν φαίνεται ότι έχουν σχέση με το πρόβλημα εφοδιασμού σάκκου.

Λύση του Προβλήματος 1.

Μέθοδος 1. *Ας θεωρήσουμε τις νέες μεταβλητές: $z_k = w_k y_k$ και τη νέα συνάρτηση αξίας του εφοδίου k: $\widetilde{R}_k(z_k) = R_k(z_k / w_k)$. Τότε το πρόβλημα 1 είναι ισοδύναμο με το εξής πρόβλημα*

$$(\max) \sum_{1}^{N} \widetilde{R}_k(z_k)$$

περιορισμοί

!!FOOTNOTE 7:Γράφουμε την τερματική συνθήκη σ' αυτή τη μορφή, σε αντιδιαστολή μ' εκείνη της σελίδας 48, όταν δεν απαιτούμε οι συναρτήσεις $R_k(\alpha)$ να είναι αύξουσες.

$$\sum_{k=1}^{N} z_k \leq B,$$

$$z_k \ \geq \ 0 \, .$$

Το παραπάνω πρόβλημα έχει την μορφή εκείνου που εξετάσαμε στην προηγούμενη παράγραφο.

Μέθοδος 2. *Μία νέα διατύπωση του προβλήματος είναι η εξής.*

*1. **Καταστάσεις - Αποφάσεις.** Οι μεταβλητές που τις περιγράφουν ορίζονται όπως και στο συνηθισμένο πρόβλημα κατανομής πόρων που ορίσαμε στη σελίδα 47. Τα σύνολα όμως των εφικτών αποφάσεων πρέπει να τροποποιηθούν έτσι, που να εκφράζουν τον περιορισμό του προβλήματος 1. Ετσι έχουμε*

$$D(x_t) \ = \ \{\alpha \ : \ w_t\alpha \ \leq \ x_t\} \ = \ \{\alpha \ : \ \alpha \ \leq \ x_t/w_t\} \, .$$

*2. **Δυναμική του συστήματος.***

$x_{t+1} \ = \ x_t - w_t\alpha_t$, *ενώ αρχική κατάσταση είναι η* $x_1 \ = \ B$.

*3. Οι **συναρτήσεις τρέχοντος και τερματικού κόστους** παραμένουν οι ίδιες.*

Η εξίσωση βελτιστοποίησης είναι η εξής.

$$v(t,x) = \max_{\alpha \in D(x_t)} \{R_t(\alpha) + v(t+1, x - w_t\alpha)\},$$

με τερματική συνθήκη

$$v(N,x) \ = \max_{\alpha \in D_N(x)} R_N(\alpha)^7 \, ,$$

και λύση :

$$v(1,B) \, .$$

5. Η Τεχνική των Γενικευμένων Πολλαπλασιαστών Lagrange. *Ας θεωρήσουμε το πρόβλημα βελτιστοποίησης*

$$(\max) \ f(y_1,y_2,\ldots,y_n)$$

περιορισμοί:

$$g_k(y_1,\ldots,y_n) \ \leq \ B_k \, , \qquad\qquad k \ = \ 1,2,\ldots,m$$
$$(y_1,\ldots,y_n) \ \in \ S \, ,$$

όπου s είναι κάποιο γνωστό υποσύνολο του $\mathbb{R}^n$, *π.χ. στα προβλήματα γραμμικού προγραμματισμού, συνήθως έχουμε* $S \ = \ \mathbb{R}^n_+ \ = \ \{y_i \, , \ y_i \geq 0 \, , \ i = 1,\ldots,n\}$. *Ορίζουμε την συνάρτηση Lagrangian ως εξής.*

$$L(y_1,\ldots,y_n \, ; \, \lambda_1,\ldots,\lambda_m) \ = \ f(y_1,\ldots, y_n) + \sum_{k=1}^{m} \lambda_k g(y_1,\ldots, y_n) \, ,$$

όπου υποθέτουμε ότι $\lambda_i \ \leq \ 0$, *είναι γνωστοί αριθμοί. Οι επόμενες ιδιότητες της Lagrangian είναι μερικές φορές χρήσιμες στη λύση των προβλημάτων που εξετάζουμε.*

Πρόταση 1. *Αν* $y^* \ = \ (y_1^*,\ldots, \ y_n^*) \ \in \ S$ *μεγιστοποιεί την* $L(y_1,\ldots, \ y_n \, , \, \lambda)$ *για κάποιο σταθερό* $\lambda = (\lambda_1,\ldots,\lambda_m) \ \leq \ (0,\ldots,0)$, *τότε το σημείο* y^* *είναι μια βέλτιστη λύση του προβλήματος*

7

$$(max)\ f(y_1,\ldots,y_n)$$

περιορισμοί:

$$g_k(y_1,\ldots,y_n) \ \le\ g(y_1^*,\ldots,y_n^*)$$
$$(y_1,\ldots,y_n) \ \in\ S\,.$$

Απόδειξη.

$$L(y^*,\lambda) \ \ge\ L(y,\lambda) \quad \forall y \in S\,, \qquad g_k(y) \ \le\ g_k(y^*) \quad \forall k\,,$$

δηλαδή

$$f(y_1^*,\ldots,y_n^*) + \sum_{k=1}^{m}\lambda_k g_k(y_1^*,\ldots,y_n^*) \ \ge\ f(y_1,\ldots,y_n) + \sum_{k=1}^{m}\lambda_k g(y_1,\ldots,y_n) \quad \Rightarrow$$

$$f(y_1^*,\ldots,y_n^*) - f(y_1,\ldots,y_n) \ \ge\ \sum_{k=1}^{m}\lambda_k\Big(g_k(y_1,\ldots,y_n) - g_k(y_1^*,\ldots,y_n^*)\Big),$$

αλλά $\lambda_k \le 0$, *και έχουμε δεχτεί ότι* $g_k(y) - g_k(y^*) \le 0$.
Αρα, η τελευταία ανισότητα για την $f(\cdot)$ *συνεπάγεται ότι* $f(y^*) \ge f(y)$.

Πρόταση 2. *Ας θεωρήσουμε την Lagrangian για το πρόβλημα που έχει ένα μόνο περιορισμό*

$$L(y_1,\ldots,y_n\,,\,\lambda) \ =\ f(y_1,\ldots,y_n) + \lambda g(y_1,\ldots,y_n).$$

Ας συμβολίσουμε με $P(\lambda)$ *το πρόβλημα μεγιστοποίησης της Lagrangian πάνω στο σύνολο* S, *και ας υποθέσουμε ότι για* $\lambda_1\,,\,\lambda_2 \ \le\ 0$, *όπου* $\lambda_1 < \lambda_2$, $y^1\,,y^2$ *είναι αντίστοιχες βέλτιστες λύσεις των προβλημάτων* $P(\lambda_1)\,,\,P(\lambda_2)$. *Τότε*

$$g(y^1) \ \le\ g(y^2)\,.$$

Απόδειξη. *Λόγω του ότι η* y^1 *είναι βέλτιστη για το* $P(\lambda_1)$, *έχουμε*

$$L(y^1,\lambda_1) \ \ge\ L(y^2,\lambda_1) \tag{1}$$

Ανάλογα παίρνουμε

$$L\,(y^2,\lambda_2) \ \ge\ L(y^1,\lambda_2) \tag{2}$$

Δηλαδή

$$f(y_1^1,\ldots,y_n^1) + \lambda_1 g(y_1^1,\ldots,y_n^1) \ \ge\ f(y_1^2,\ldots,y_n^2) + \lambda_1 g(y_1^2,\ldots,y_n^2)\,, \tag{3}$$

και

$$f(y_1^2,\ldots,y_n^2) + \lambda_2 g(y_1^2,\ldots,y_n^2) \ \ge\ f(y_1^1,\ldots,y_n^1) + \lambda_2 g(y_1^1,\ldots,y_n^1)\,. \tag{4}$$

Προσθέτοντας τις (3) και (4) παίρνουμε:

$$\lambda_1\Big(g(y_1^1,\ldots,y_n^1) - g(y_1^2,\ldots,y_n^2)\Big) \ \ge\ \lambda_2\Big(g(y_1^1,\ldots,y_n^1) - \ g(y_1^2,\ldots,y_n^2)\Big)\,. \tag{5}$$

Τώρα, επειδή $\lambda_1 < \lambda_2 \le 0$, *η (5) συνεπάγεται ότι :*

$$g(y^1) \ \le\ g(y^2).$$

Παρατήρηση. *Ας υποθέσουμε ότι θέλουμε να λύσουμε το πρόβλημα*

$$\max\ f(y_1,\ldots,y_n)$$

$$g(y_1,\ldots,y_n)\ \le\ B,$$

$$(y_1,\ldots,y_n)\ \in S.$$

Βασιζόμενοι στις προτάσεις 1 και 2 μπορούμε να κατασκευάσουμε την εξής εξακολουθητική μέθοδο. Εστω $\lambda_1 \le 0$, και έστω y^1 η βέλτιστη λύση του προβλήματος $P(\lambda_1)$. Τότε, σύμφωνα με την πρόταση 1 , η y^1 είναι βέλτιστη για το πρόβλημα

$$\max f(y_1,\ldots,y_n)\ ,\ g(y_1,\ldots,y_n)\ \le\ g(y_1^1,\ldots,y_n^1)\quad y\in S\,.$$

Αν $g(y_1^1,\ldots,y_n^1)\ <\ B$, τότε διαλέγουμε ένα $\lambda_2 \le 0$, μεγαλύτερο του λ_1 , λύνουμε πάλι το $P(\lambda_2)$ γνωρίζοντας από την πρόταση 2 ότι $g(y_1^2,\ldots,y_n^2)\ \ge\ g(y_1^1,\ldots,y_n^1)$.

Αν $g(y_1^1,\ldots,y_n^1)\ >\ B$, διαλέγουμε ένα λ_2 μικρότερο του λ_1 , και λύνουμε το $P(\lambda_2)$. Θα σταματήσουμε τις επαναλήψεις (iterations) όταν βρούμε κάποιο σημείο y^k για το οποίο έχουμε

$$B-\epsilon\ \le\ g(y_1^k,\ldots,y_n^k)\ \le\ B.$$

Αν και δεν υπάρχει βεβαιότητα ότι η παραπάνω μέθοδος θα βρει κάποιο σημείο y^k που πληροί την παραπάνω ιδιότητα, στην πράξη έχει αποδειχθεί χρήσιμη. Αν έχουμε ένα πρόβλημα με παραπάνω απο ένα περιορισμούς, τότε μπορούμε να κατασκευάσουμε μια ανάλογη επαναληπτική μέθοδο, αλλά τώρα η πρόταση 2 δεν ισχύει. Υπάρχει όμως μια ανάλογη αλλά πιο ασθενής μορφή της πρότασης 2 (άσκηση) που είναι δυνατόν να μας βοηθήσει.

Λύση του Προβλήματος 2. *Ας θεωρήσουμε πάλι το πρόβλημα*

$$(\max)\ \sum_{k=1}^{N} R_k(y_k)$$

περιορισμοί

$$\sum_{k=1}^{N} w_k^1 y_k\ \le\ B_1\ ,$$

$$\sum_{k=1}^{N} w_k^2 y_k\ \le\ B_2\ ,$$

$$y_k\ \ge\ 0\ ,$$

όπου $w_t\ \ge\ 0, B_1, B_2$ είναι γνωστοί αριθμοί.

Μέθοδος 1 .

1. Καταστάσεις–Αποφάσεις . *i) Εστω (x_t^1,x_t^1) το ζεύγος των μεταβλητών, όπου x_t^1 (x_t^2) συμβολίζει την ποσότητα του πόρου 1 (2) που απομένει για κατανομή στίς δραστηριότητες $t,t+1,\ldots,N$ δηλαδή μετά τις $t-1$ πρώτες κατανομές.*

ii) Η ποσότητα που διαθέτουμε στην δραστηριότητα t είναι η απόφαση α_t που αντιστοιχεί στην "περίοδο" t , και πρέπει : $w_t^1\alpha_t\ \le\ x_t^1$, και $w_t^2\alpha_t\ \le\ x_t^2$. Το σύνολο $D(x_t)$ είναι λοιπόν το : $D(x_t)\ =\ [0\ ,\ \min\{x_t^i/w_t^i\ ,\ i=1,2\}]$.

2. Δυναμική του Συστήματος

$$(x_{t+1}^1,x_{t+1}^2)\ =\ (x_t^1-w_t^1\alpha_t\ ,\ x_t^2-w_t^2\alpha_t)$$

3. *Η συνάρτηση **τρέχοντος κόστους** είναι*

$$c((x_t^1, x_t^2), \alpha_t) = R_t(\alpha_t),$$

και η συνάρτηση ***τερματικού κόστους***

$$\hat{c}(x_N^1, x_N^2) = \max_{\alpha_N \in D(x_N^1, x_N^2)} \{R_N(\alpha_N)\} .$$

Η εξίσωση βελτιστοποίησης είναι η

$$v(t,(x_t^1, x_t^2)) = \max_{\alpha \in D(x_t^1, x_t^2)} \left\{ R_t(\alpha) + v(t+1, x_t^1 - w_t^1\alpha, x_t^2 - w_t^2\alpha) \right\}$$

με τερματική συνθήκη

$$v(N,(x_N^1, x_N^2)) = \hat{c}(x_N^1, x_N^2)$$

και λύση : $v(1, B_1, B_2)$.

Μέθοδος 2. *Εστω* $\lambda_1 \le 0$. *Θεωρούμε το εξής πρόβλημα*

$$\max \sum_{k=1}^{N} R_k(y_k) + \lambda_1 \sum_{1}^{N} w_k^2 y_k$$

περιορισμοί :

$$\sum_{k=1}^{N} w_k^1 y_k \le B_1 ,$$

$$y_k \ge 0 .$$

Παρατηρούμε ότι η αντικειμενική συνάρτηση έχει την μορφή

$$\sum_{k=1}^{N} \tilde{R}_k(y_k) , \text{ όπου } \tilde{R}_k(y_k) = R_k(y_k) + \lambda_1 w_k^2 y_k .$$

Ετσι, το παραπάνω πρόβλημα είναι ένα μονοδιάστατο πρόβλημα εφοδιασμού, και μπορεί να λυθεί όπως το πρόβλημα 1 . Εστω y^* *η βέλτιστη λύση. Αν* $\sum_{1}^{N} w_k^2 y_k^* = g_2(y^*) = B_2$, *τότε σύμφωνα με την πρόταση 1 έχουμε βρει την λύση στο αρχικό πρόβλημα, εαν όχι συνεχίζουμε με μια νέα τιμή για τον πολλαπλασιστή Lagrange* λ , *σύμφωνα με τα όσα είδαμε στην παρατήρηση που εκάναμε μετά την πρόταση 2.*

Παρατηρήσεις. *Ας θεωρήσουμε το πρόβλημα κατανομής* m*–πόρων*

$$\max \sum_{k=1}^{N} R_k(y_k)$$

περιορισμοί

$$\sum_{k=1}^{N} w_k^i y_k \le B_i , \quad i = 1, \ldots, m ,$$

$$y_k \ge 0 .$$

1. *Οι καταστάσεις του συστήματος, τώρα περιγράφονται απο διανύσματα της μορφής* $(x_t^1, x_t^2, \ldots, x_t^m)$ *όπου* x_t^i *είναι* $t, t+1, \ldots, N$. *Οι χώροι καταστάσεων είναι της μορφής* $S_t = [0,B_1] \times [0,B_2] \times \cdots \times$

$[0,B_m]$. *Αν επιπλέον θεωρήσουμε ότι οι κατανομές πρέπει να είναι ακέραιες, τότε ο αριθμός των κόμβων (=καταστάσεων) θα είναι ίσος με* $(B_1+1)(B_2+1)\cdots(B_m+1)(1+2+\ldots+N) = \prod_1^m (B_i+1)$ $\frac{N(N+1)}{2}$. *Αν επιπλέον* $B_i = B \ \forall i$, *τότε ο παραπάνω αριθμός είναι ίσος με* $K = (1+B)^m \frac{N(N+1)}{2}$. *Ο αριθμός αυτός τείνει να γίνει μεγάλος πολύ γρήγορα, π.χ. αν* $B = 9$, $N = 5$, $m = 10$, *τότε* $K = 10^{10} \times 15$, *ενώ αν :* $B = 9$, $N = 5$, $m = 20$, *τότε* $K = 10^{20} \times 15$. *Η λύση, λοιπόν, με την βοήθεια Η/Υ θα απαιτήσει πολύ χρόνο (ώρες) όταν είναι εφικτή. Η μέθοδος των Πολλαπλασιαστών Lagrange είναι ένα εργαλείο για να ελαττώσουμε την διάσταση του προβλήματος m.*

2. *Παρατηρούμε ότι στη περίπτωση που οι συναρτήσεις* $R_k(y_k)$ *είναι γραμμικές* $R_k(y_k) = r_k y_k$, *τότε το παραπάνω πρόβλημα είναι ένα πρόβλημα γραμμικούπρογραμματισμού. Ο δυναμικός προγραμματισμός λοιπόν μας δίδει μια άλλη μέθοδο λύσης προβλημάτων γραμμικού προγραμματισμού, που είναι ιδιαίτερχρήσιμη στη περίπτωση που οι μεταβλητές* y_k *πρέπει να είναι ακέραιες.*

6 . Εφαρμογές του Προβλήματος Κατανομής Πόρων.

6.1. Πρόβλημα Σχεδιασμού Συστημάτων. *Ας θεωρήσουεμ ένα σύστημα που αποτελείται απο N στοιχεία (components), και ας υποθέσουμε ότι το σύστημα δουλεύει, όταν όλα του τα στοιχεία δουλεύουν. Σχηματικά το σύστημα περιγράφεται ως εξής.*

$\boxed{c_1}$ $\boxed{c_2}$ $\boxed{c_N}$

Εστω P_k *η πιθανότητα ότι το στοιχείο* c_k *δουλεύει, και έστω* w_k *το "βάρος" (ή κόστος) του στοιχείου* c_k , $\kappa = 1,\ldots,N$. *Δεχόμαστε ότι τα στοιχεία είναι ανεξάρτητα το ένα απο το άλλο. Τότε η* ***αξιοπιστία*** *(reliability) του συστήματος είναι ίση μέ*

$$R = \prod_{k=1}^{N} P_k \, .$$

Ενας τρόπος για να αυξήσουμε την αξιοπιστία του συστήματος είναπροσθέσουμε σε παράλληλη διάταξη στοιχεία. Εστω y_k *ο αριθμός των επιπλέον στοιχείων που προσθέτουμε στο στοιχείο* c_k.

Σχηματικά, το σύστημα τώρα περιγράφεται ως εξής

$\boxed{c_1}$ $\boxed{c_2}$ $\boxed{c_N}$

$\boxed{c_1}$ $\boxed{c_2}$ $\boxed{c_N}$

$\vdots$ $\vdots$ $\vdots$

$\boxed{c_1}$ $\boxed{c_2}$ $\boxed{c_N}$

Το k–υποσύστημα δουλεύει όταν τουλάχιστον ένα απο τα στοιχεία του δουλεύουν. Η αξιοπιστία του k–υποσυστήματος, στο οποίο έχουμε προσθέσει y_k *στοιχεία, είναι ίση με*

$$\begin{aligned} R_k(y_k) &= 1 - P(\text{και τα } y_k+1 \text{ στοιχεία δεν δουλεύουν}) \\ &= 1 - (1-P_k)^{1+y_k}. \end{aligned}$$

Η αξιοπιστία του όλου συστήματος είναι ίση με

$$R(y_1,\ldots,y_N) = \prod_{k=1}^{N} R_k(y_k)\,.$$

Ο σκοπός μας είναι να βρούμε τις τιμές των $y_1,\ldots,y_N$, που μεγιστοποιούν την αξιοπιστία του συστήματος κάτω απο τον περιορισμό ότι το συνολικό βάρος δεν πρέπει να ξεπερνά μια ανώτατη τιμή w. Συμβολικά έχουμε

$$max\;\; R(y_1,\ldots,y_N) = \prod_{k=1}^{N} R_k(y_k)\,,$$

περιορισμός

$$\sum_{i=1}^{N} w_i(y_i + 1) \leq w\,,$$

$$y_i \geq 0\,.$$

Είναι εύκολο να δούμε ότι το παραπάνω πρόβλημα είναι ισοδύναμο με το

$$(max)\;\; log\, R(y_1,\ldots,y_N) = \sum_{k=1}^{N} log\, R_k(y_k)\,,$$

περιορισμός

$$\sum_{i=1}^{N} w_i y_i \leq w^1 = (w - \sum_{i=1}^{N} w_i)\,,$$

$$y_i \geq 0\,.$$

Το επόμενο πρόβλημα είναι ένα πρόβλημα κατανομής πόρων στη συνήθη του μορφή, όπου οι μεταβλητές y_i πρέπει να είναι ακέραιοι αριθμοί.

6.2. Πρόβλημα Σχεδιασμού Συστημάτων Πρώτων Βοηθειών *Ας υποθέσουμε ότι μια πόλη χωρίζεται σε N περιοχές Σε κάθε περιοχή k κατανέμουμε ένα αριθμό y_k μονάδων πρώτων βοηθειών (π.χ. νοσοκομειακά αυτοκίνητα, οχήματα πυροσβεστικής, περιπολικά αστυνομίας, κ.λ.π.). Το πρόβλημα είναι να διαλέξουμε τιμές για τα y_k, $k = 1,\ldots,N$, έτσι ώστε να βελτιστοποιήσουμε κάποιο μέτρο απόδοσης. Σε μια πρώτη μελέτη που έγινε το 1965 για την πυροσβεστική υπηρεσία της Νέας Υόρκης, χρησιμοποιήθηκαν τα εξής στοιχεία.*

A_i *= μέγεθος της περιοχής i, $i = 1,\ldots,N$*

λ_i *= ο μέσος όρος των συναγερμών ανά μονάδα χρόνου στην περιοχή i*

s_i *= ο μέσος όρος του χρόνου που μια μονάδα απασχολείται για ένα συναγερμό στην περιοχή i.*

Σαν μέτρο απόδοσης χρησιμοποιήθηκε ο μέσος χρόνος ανταπόκρισης σ ένα συναγερμό για όλη την πόλη, όπου στατιστικά στοιχεία έδειξαν ότι μια καλή προσέγγιση για τον μέσο χρόνο ανταπόκρισης στην περιοχή i δίνεται απο μια συνάρτηση της μορφής

$$r_i(y_i) = \alpha_i + b_i \frac{A_i}{y_i - \lambda_i s_i}\,,$$

όπου α_i, b_i είναι παράμετροι που βρίσκονται με στατιστικές μεθόδους.

Ετσι το πρόβλημα για τον σχεδιασμό της πυροσβεστικής υπηρεσίας για την πόλη είναι

$$(min) \quad \sum_{i=1}^{N} \frac{\lambda_i}{\sum_{j=1}^{N}\lambda_j} \; r_i(y_i)$$

περιορισμός

$$\sum_{1}^{N} y_i \;\leq\; B\,,$$

$$y_i \;\geq\; 0\,.$$

6.3. Το Πρόβλημα του Κοψίματος Προιόντων. *(Cutting Stock Problem). Ενα εργοστάσιο παράγει ράβδους σιδήρου σταθερού μήκους* l. *Πελάτες απαιτούν κομάτια που να έχουν μήκος* l_1 *και* l_2 *σε ποσότητες* m_1 *και* m_2 *αντίστοιχα. Το εργοστάσιο έχει στη διάθεσή του ένα συγκεκριμένο αριθμό διαφορετικών τρόπων για να κόβει τις ράβδους (cutting patterns).*

Εστω P_j *ο* j *τρόπος κοψίματος ο οποίος ορίζεται ως εξής.* $P_j = (p_{1j}\,, p_{2j})$, *όπου* p_{kj} = *αριθμός κομματιών μήκους* l_k *που κόβονται απο την ράβδο μεγέθους* l, $k = 1,2$, *π.χ.* $P_1 = (10,0)$, $P_2 = (8,4)$, $P_3 = (0,20)$. *Οταν κόβονται οι αρχικές ράβδοι με τους διαφορετικούς τρόπους, συχνά απομένουν κομμάτια, τα οποία περισσεύουν και θεωρούμε ότι χάνονται. Το πρόβλημα λοιπόν είναι να περιορίσουμε κατά το δυνατόν το σύνολο των περισσευμάτων, και να ικανοποιήσουμε την ζήτηση. Ισοδύναμα, αρκεί να ελαχιστοποιήσουμε τον συνολικό αριθμό των ράβδων που θα κόψει το εργοστάσιο για να ικανοποιήσει την ζήτηση.*

Εστω λοιπόν y_i *ο αριθμός των ράβδων που θα κοπούν με τον τρόπο* P_j, $i = 1,\ldots,N$, *όπου* N *είναι ο συνολικός αριθμός διαφορετικών τρόπων κοψίματος που χρησιμοποιεί το εργοστάσιο. Το πρόβλημα λοιπόν είναι*

$$(min) \sum_{i=1}^{N} y_i\,,$$

περιορισμοί

$$P_1 y_1 + \cdots + P_N y_N \;\geq\; \begin{pmatrix} n_1 \\ n_2 \end{pmatrix},$$

$$y_i \;\geq\; 0\,.$$

Η γενίκευση του προβλήματος για την περίπτωση που οι απαιτήσεις των πελατών είναι της μορφής $l_1, l_2, \ldots, l_n$, *σε ποσότητες* $m_1, m_2, \ldots, m_n$ *δεν είναι δύσκολο να γίνει.*

7. Πιο Πολύπλοκα Προβλήματα.

7.1. Πρόβλημα Απόσταξης.

B

Δ_2 $\qquad$ Δ_3 $\qquad$ *Εικόνα 9*

Δ_1

Οι βασικές εργασίες σέ ένα εργοστάσιο απόσταξης είναι οι εξής :

1) Σε ένα δοχείο Δ_1 *τοποθετείται μια ποσότητα Β υλικού (π.χ. νερό και σταφύλια, νερό και τεύτλα, κ.λ.π.).*

2) Το δοχείο Δ_1 *θερμαίνεται, μέχρι που η πίεση μέσα σ' αυτό να φτάσει κάποιο γνωστό επίπεδο. Τότε η ένταση της φωτιάς ελέγχεται, έτσι ώστε η πίεση στο δοχείο να διατηρηθεί σταθερή.*

3) Το υλικό στο δοχείο Δ_1 *βράζει, αρχίζει η ζύμωση με την βοήθεια καταλυτών, και στη μορφή ατμών αρχίζει να ανεβαίνει τον πύργο απόσταξης* $\Delta 2$ *όπου αρχίζει και ψύχεται. Ο ατμός περνά απο την βαλβίδα (Β) όπου χωρίζεται σε δυο ροές. Ετσι, ένα μέρος γυρίζει πίσω στο δοχείο* Δ_1 *ενώ ένα άλλο μέρος καταλήγει αφού υγροποιηθεί στο δοχείο* Δ_3 *. Η βαλβίδα μας παρέχει την δυνατότητα να ασκήσουμε έλεγχο πάνω στη διαδικασία της παραγωγής, όπως θα δούμε παρακάτω.*

Η νέα έννοια που εισάγουμε με αυτό το πρόβλημα είναι εκείνη της ***εποχής αλλαγής*** *(transition epoch). Συγκεκριμένα, είναι φανερό ότι η παραγωγική διαδικασία που εξετάζουμε είναι μια διαδικασία που αναπτύσεται σε συνεχή χρόνο, δηλ. στο διάστημα* $(0,\infty)$ *.*

$t_0 = 0 \qquad t_1 \qquad t_2 \qquad t_3 \quad \cdots\cdots$

$\tau_0 = 0 \qquad \tau_1 \qquad \tau_2 \qquad \tau_3 \qquad \tau_4 \quad \cdots\cdots$

Για να χρησιμοποιήσουμε την μεθοδολογία του δυναμικού προγραμματισμού είναι απαραίτητο να χωρίσουμε το διάστημα $(0,\infty)$ *σε διακριτά μέρη (discretize τον χρόνο). Η πιο προφανής μέθοδος είναι να χρησιμοποιήσουμε διαστήματα ίσου (και μικρού) μήκους π.χ. στην παραπάνω εικόνα το μήκος του k διαστήματος* $t_k - t_{k-1}$ *είναι δυνατόν να θεωρηθεί ίσο με 1 sec για κάθε* $k = 1,2$. *Το ερώτημα βέβαια για το ποιό χρονικό διάστημα είναι μικρό έτσι που η προσέγγιση να είναι καλή, εξαρτάται απο την όλη διαδικασία.*

Μια δεύτερη μέθοδος για να χωρίσουμε το διάστημα $(0,\infty)$*, σε διακριτά μέρη, βρίσκεται όταν αφήνουμε την ίδια την διαδικασία που μελετάμε να ορίσει τα διαστήματα* (τ_0,τ_1) , (τ_1,τ_2) ,... *τα οποία σ' αυτή την περίπτωση δεν έχουν αναγκαστικά το ίδιο μήκος. Οι χρονικές στιγμές* $\tau_0,\tau_1,\tau_2,\ldots$ *ονομάζονται* ***εποχές αλλαγής*** *(transition epochs), και ανάλογα με το πρόβλημα που μελετάμε, υπάρχουν πολλοί τρόποι για να τις ορίσουμε.*

Για το συγκεκριμένο πρόβλημα, δεχόμαστε ότι είναι επιτρεπτό να αλλάξουμε την τιμή της απόφασης (control) μόνο κάθε φορά που μια μονάδα τελικού προιόντος μαζεύεται στο δοχείο Δ_3 *, ενώ κατά την διάρκεια των χρονικών περιόδων* (τ_{k-1},τ_k) *, διατηρούμε σταθερή την τιμή της απόφασης (άνοιγμα της βαλβίδας) .*

Ετσι λοιπόν , τ_1 *είναι ο χρόνος που απαιτείται μέχρι που η πρώτη μονάδα (π.χ. ένα kgr) προιόντος μαζευτεί στο δοχείο*Δ_3*, κ.λ.π. Παρατηρούμε ότι οι εποχές αλλαγής καθορίζονται απο την αλληλοεπίδραση της παραγωγικής διαδικασίας* ***και της πολιτικής*** *που ακολουθείται.*

Για να περιγράψουμε το πρόβλημα, ας αρχίσουμε με τον ορισμό μερικών αναγκαίων μεταβλητών. Εστω

x_t := *ποσότητα αλκοόλ, σε kg , που υπάρχει στα δοχεία* Δ_1,Δ_2 *, και στις εσωτερικές σωληνώσεις, στην αρχή της περιόδου* (τ_t,τ_{t+1}) .

y_t := ποσότητα,σε kg , μη ζυμωμένου υλικού που υπάρχει στο Δ_1 στην αρχή της περιόδου (τ_t,τ_{t+1}) .

z_t := ποσότητα, σε kg , αλκοόλ που υπάρχει στο Δ_3 και στις εξωτερικές σωληνώσεις, στην αρχή της περιόδου (τ_t,τ_{t+1}) .

B := αρχική ποσότητα υλικού στο δοχείο.

α_t := το ποσοστό του ατμού που η βαλβίδα (B) επιτρέπει να προχωρήσει προς το δοχείο Δ_3 , κατά την περίοδο t , (δηλ. το διάστημα (τ_t,τ_{t+1})) . Οι τιμές που το α_t μπορεί να πάρει, είναι στο διάστημα $[0,1]$.

Οι χημικοί μηχανικοί μας δίδουν της εξής σχέσεις–συναρτήσεις, που περιγράφουν την όλη διαδικασία .

i) $T(x_t,y_t,\alpha_t)$:= χρόνος που απαιτείται για να μαζευτεί μια επιπλέον μονάδα προιόντος (π.χ. μονάδα = kg) στο Δ_3 όταν $(x_t,\ y_t)$ είναι η κατάσταση στο δοχείο, και α_t είναι το άνοιγμα της βαλβίδας κατά τη χρονική περίοδο (τ_t,τ_{t+1}). Παρατηρούμε εδώ ότι

$$\tau_{t+1} = \tau_t + T(x_t,y_t,\alpha_t) .$$

ii) $h(x_t,y_t)$:= Η ποσότητα αλκοόλ που περιέχεται στην επόμενη μονάδα προιόντος στο Δ_3 , δεδομένων των x_t,y_t .

iii) $g(x_t,y_t,\alpha_t)$:= ποσότητα υλικού (σε kg) στο Δ_3 , που θα ζυμωθεί σε αλκοόλ κατά την περίοδο (τ_t,τ_{t+1}) , σαν συνάρτηση των (x_t,y_t,α_t) .

Η παραγωγική διαδικασία επιβάλλει τις εξής σχέσεις στις παραπάνω ποσότητες.

1) $x_t + y_t + z_t = B , \quad t = 0,1,2,\dots ,$

2) $z_{t+1} = z_t + h(x_t,y_t) , \quad t = 0,1,\dots ,$

3) $x_{t+1} = x_t + g(x_t,y_t,\alpha_t) - h(x_t,y_t) ,$

4) $y_{t+1} = y_t - g(x_t,y_t,\alpha_t) - (1 - h(x_t,\ y_t)\ ) .$

Η εξίσωση (1) εκφράζει την διατήρηση της μάζας και το γεγονός ότι σε κάθε χρονική περίοδο η ποσότητα προιόντος στο Δ_3 αυξάνει κατά μία μονάδα, η εξίσωση (2) εκφράζει το γεγονός ότι η ποσότητα του αλκοολ στο Δ_3 αυξάνει σε ποσότητες μεγέθους $h(x_t,y_t)$, η εξίσωση (3) εκφράζει το γεγονός ότι η ποσότητα αλκοόλ στα δοχεία Δ_1 , Δ_2 αυξάνει λόγω ζύμωσης και ελαττούται λόγω εξόδου στο δοχείο Δ_3 . Τέλος, η εξίσωση (4) εκφράζει το γεγονός ότι το υλικό στο δοχείο Δ_1 ελατώνεται λόγω ζύμωσης και της εξόδου προιόντος στο δοχείο Δ_3, όπου σε κάθε μονάδα προιόντος που πηγαίνει στο Δ_3, $h(x_t,y_t)$ είναι η ποσότητα του αλκοόλ και $(1 - h(x_t,y_t))$ είναι η ποσότητα των άλλων ουσιών.

Οι εξισώσεις (1) – (4) καθορίζουν την δυναμική του συστήματος, όπου η κατάσταση του συστήματος ορίζεται απο την τριάδα (x_t,y_t,z_t) και η απόφαση $\alpha_t \in D(x_t,y_t,z_t) = [0,1]$.

Σχηματικά έχουμε

$$\boxed{x_t,y_t,z_t} \ - - - \overset{\alpha_t}{-} - - \longrightarrow \ \boxed{x_{t+1},y_{t+1},z_{t+1}}$$

όπου τα x_{t+1} , y_{t+1} , z_{t+1} ορίζονται απο τις σχέσεις (1) μέχρι (4) .

Ας υποθέσουμε ότι ο στόχος μας είναι να παράγουμε N μονάδες προιόντος που να περιέχουν Z μονάδες (kg) αλκοόλ (έξοδος στο Δ_3) ασο το δυνατόν πιο γρήγορα.

Εχουμε ήδη ορίσει τις καταστάσεις - αποφάσεις και την δυναμική του συστήματος. Αρα, αρκεί να ορίσουμε τις συναρτήσεις τρέχοντος και τερματικού κόστους. Δεν είναι δύσκολο να δούμε ότι η συνάρτηση τρέχοντος κόστους ορίζεται απο την συνάρτηση $T(x_t,y_t,\alpha_t)$ *ενώ η συνάρτηση τερματικού κόστους ορίζεται ως εξής*

$$\hat{c}(N,x_N,y_N,z_N) \;=\; \begin{cases} -M & \alpha\nu\;\; z_N \neq z \\ 0 & \alpha\nu\;\; z_N = z \end{cases} \qquad \forall\, x_N\,,\;\; y_N\,,$$

όπου M *είναι ένας μεγάλος αριθμός.*

Η αρχική κατάσταση είναι η $(0,x_0,y_0,z_0) \;=\; (0,0,B,0)$.

Παρατήρηση. *Λόγω της (1) η γνώση της* y_t *και της χρονικής περιόδου* t *αρκούν για να ορίσουμε την μεταβλητή* $x_t := B - y_t - t$. *Αρα γνωρίζουμε επαρκώς την κατάσταση του συστήματος όταν γνωρίζουμε τις τιμές της τριάδας* (t,y_t,z_t) .

Η εξισώσεις βελτιστοποίησης για το πρόβλημα είναι οι εξής

$$v(N,y,z) \;=\; \hat{c}(N,y,z)\,,$$

$$v(t,y,z) \;=\; \min_{\alpha \in [0,1]} \{T(B-y_t-t,y_t,\alpha) \;+\; v(t+1,y',z')\}\,,$$

όπου

$$y' \;=\; y \;-\; g(B-y-t,y,\alpha) \;-\; (1-h(B-y-t,y))\,,$$

$$z' \;=\; z + h(B-y-t,y)\,.$$

Η λύση του προβλήματος δίδεται απο την τιμή $v(0,B,0)$.

7.2. Πρόβλημα Σταθεροποίησης. *Συχνά σε προβλήματα μηχανικών εμφανίζονται καταστάσεις όπου θέλουμε να διατηρήσουμε μια ποσότητα σταθερή, μεταβάλλοντας μια άλλη ποσότητα η οποία την επηρεάζει. Για παράδειγμα, αναφέρουμε το πρόβλημα της σταθεροποίησης της πίεσης στο παραγωγικό σύστημα της προηγούμενης παραγράφου που θέλουμε να διατηρήσουμε σταθερή μεταβάλλοντας την ένταση της φωτιάς κάτω απο το δοχείο* Δ_1 .

Η γενική μορφή του προβλήματος έχει ως εξ#ής. Να βρεθεί μια πεπερασμένη ακολουθία $\{x_1,x_2,\ldots,x_N\}$ *που ελαχιστοποιεί την συνάρτηση*

$$F(x_1,\ldots,x_N) \;=\; \sum_{k=1}^{N} f_k(x_k \;-\; r_k) + \sum_{k=1}^{N} h_k(x_k \;-\; x_{k-1})\,,$$

όπου $\{r_k\,,\; k=1,\ldots,N\}$ *είναι μια ακολουθία γνωστών αριθμών, οι συναρτήσεις* $f_k(x)$, $h_k(x)$ *είναι συνεχείς για πεπερασμένα* x , *και* $f_k(x) \to \infty$, $h_k(x) \to \infty$, *όταν* $x \to \infty$. *Επίσης,* $x_0 = c$ *είναι ένας γνωστός αριθμός. Το όνομα του προβλήματος οφείλεται στο γεγονός ότι το να ελαχιστοποιή-σουμε την* F *σημαίνει ότι θέλουμε να ακολουθήσουμε την ακολουθία* $\{r_k\}$ *(μικρή τιμή στο άθροισμα των* $f_k(x_k - r_k)$*) με μικρές διακυμάνσεις της ακολουθίας* $\{x_k\}$, *(μικρή τιμή στο άθροισμα των* $h_k(x_k - x_{k-1})$*). Το πρόβλημα γίνεται πιο ρεαλιστικό αν υποθέσουμε ότι οι αριθμοί* x_k *πρέπει να πληρούν τους παρακάτω περιορισμούς*

$$x_{k+1} \;=\; g(x_k,\alpha_k)\;,\;\; \alpha_k \in D(x_k)\;,\;\; k = 0,\ldots, N-1\,,$$

και ότι οι μεταβλητές οι οποίες είναι κάτω απο τον έλεγχο μας, είναι οι α_k . *Επίσης, δίδεται η αρχική κατάσταση* x_0 .

Διατύπωση του Προβλήματος.

1) Καταστάσεις - Αποφάσεις. *Η μεταβλητή που περιγράφει την κατάσταση του συστήματος είναι:* η $\overrightarrow{x_t} = (x_t, x_{t-1})$, για $t \geq 1$ και $\overrightarrow{x_0} = (0,c)$.
Παρατηρούμε ότι ο δεύτερος όρος που εμφανίζεται στην αντικειμενική συνάρτηση F, *μας αναγκάζει να περιλάβουμε στην περιγραφή της κατάστασης* $\overrightarrow{x_t}$, *μέρος της ιστορίας του συστήματος (το* x_{t-1}*). Η μέθοδος της περίληψης μέρους της ιστορίας στη περιγραφή της κατάστασης του συστήματος είναι χρήσιμη σε προβλήματα που η αντικειμενική συνάρτηση ή/και οι περιορισμοί μας επιβάλλουν να παίρνουμε υπ όψη μέρος της ιστορίας του συστήματος στο καθορισμό βέλτιστων πολιτικών. Ετσι, οι βέλτιστες πολιτικές είναι Μαρκοβιανές ως προς την δυναμική του συστήματος που περιγράφεται απο τις μεταβλητές* $\overrightarrow{x_t}$, *ενώ δεν είναι Μαρκοβιανές ως προς την δυναμική που περιγράφεται απο τις μεταβλητές* x_t.

Για τις αποφάσεις, δεχόμαστε ότι περιορίζονται απο σύνολα της μορφής $D(x_t)$. *Στην περίπτωση που δεν υπάρχει ο περιορισμός* $g(x_t,\alpha_t) = x_{t+1}$ *τότε οι αποφάσεις είναι ταυτόσημες με τις μεταβλητές* x_t *δηλ. έχουμε* $x_{t+1} = \alpha_t$, $t = 0, 1, \ldots, N-1$.

2. Δυναμική του Συστήματος. *Η δυναμική περιγράφεται σχηματικά ως εξής*

$$\boxed{x_0 = c} \xrightarrow[\alpha_0]{} \boxed{x_1,x_0} \xrightarrow[\alpha_1]{} \boxed{x_2,x_1} \xrightarrow[\alpha_2]{} \cdots \longrightarrow \boxed{x_{N-1},x_{N-2}} \xrightarrow[\alpha_{N-1}]{} \boxed{x_N,x_{N-1}}$$

$$x_0 \qquad \overrightarrow{x}_1 \qquad \overrightarrow{x}_2 \qquad \qquad \overrightarrow{x}_{N-1} \qquad \overrightarrow{x}_N$$

3. Συναρτήσεις Κόστους.

$$c_t(\overrightarrow{x_t},\alpha) = f_{t+1}(x_{t+1} - r_{t+1}) + h_{t+1}(x_{t+1} - x_t) ,$$

όπου $x_{t+1} = g(x_t,\alpha_t)$,

$$\hat{c}(\overrightarrow{x}_N) = 0 .$$

4. Εξίσωση Βελτιστοποίησης.

$$v(N,\overrightarrow{x}_N) = 0 ,$$

$$v(t,\overrightarrow{x}_t) = \min_{\alpha \in D(x_t)} \{f_{t+1}(x_{t+1} - r_{t+1}) + h_{t+1}(x_{t+1} - x_t) + v(t+1,\overrightarrow{x}_{t+1})\} ,$$

όπου $x_{t+1} = g(x_t,\alpha_t)$, $t = 0, 1, 1, \ldots, N - 1$.

Η λύση δίδεται απο την: $v(0,x_0)$.

7.3 Πρόβλημα Ελέγχου Ταχύτητας. *Μία απλή μορφή του προβλήματος είναι η εξής. Η ταχύτητα ενός πυραύλου που ανεβαίνει κατακόρυφα ελέγχεται μεταβάλοντας τον ρυθμό καυσίμων στις μηχανές που παράγουν την δύναμη ώθησης. Ο στόχος μας είναι να προσδιορίσουμε τη πολιτική ελέγχου του ρυθμού παροχής καυσίμων έτσι που ο πύραυλος να βρίσκεται στο μέγιστο δυνατό ύψος σε κάποια δεδομένη χρονική στιγμή* T_0 . *Ας ορίσουμε τις μεταβλητές*

$x_1(t)$ = *ύψος στο οποίο βρίσκεται ο πύραυλος την χρονική στιγμή* t ,
$x_2(t)$ = *η (κάθετος) ταχύτητα την χρονική στιγμή* t ,
$x_3(t)$ = *το βάρος του πυραύλου (* $\cong$ *βάρος απομένοντος καυσίμου) την χρονική στιγμή* t .

Οι δυνάμεις που επιδρούν στο πύραυλο είναι οι εξής.

$F_1(t) = x_3(t)\,\ddot{x}_1(t) = x_3(t)\,\dot{x}_2(t) =$ αδράνεια
$F_2(t) = c_D\rho\,(x_1(t)(x_2(t))^2 =$ αντίσταση αέρα
$F_3(t) = g\,x_3(t) =$ έλξης γής
$F_4(t) = c_T\,x_3(t) =$ δύναμη ώθησης

Εστω $\alpha(t)$ = ρυθμός παροχής καυσίμου στις μηχανές την χρονική στιγμή t. Οι εξισώσεις κίνησης του πυραύλου είναι οι εξής.

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_3(t) = -\,\alpha(t) \tag{1}$$

$$\dot{x}_2(t) = -\,g - \frac{c_D}{x_3(t)}\,\rho(x_1(t)\,)\,(x_2(t))^2 + \frac{c_T}{x_3(t)}\,\alpha_t$$

Οι σταθερές g, c_D, c_T είναι αντίστοιχα η επιτάχυνση της βαρύτητας, ο συντελεστής τριβής με τον αέρα, ο συντελεστής απόδοσης των μηχανών. Τέλος, $\rho(x_1(t))$ είναι η πυκνότητα του αέρα σε ύψος $x_1(t)$.

Διατύπωση του Προβλήματος.

1. Καταστάσεις - Αποφάσεις. *Η κατάσταση του συστήματος περιγράφεται απο το διάνυσμα $\vec{x}\,(t) = (x_1(t), x_2(t), x_3(t)$. Οι αποφάσεις συμβολίζονται με $\alpha(t)$ και υποθέτουμε ότι $\alpha(t) \in [\,0,M\,]$.*

2. Δυναμική του Συστήματος. *Το σύστημα $\{\vec{x}\,(t)\,,\, t \ge 0\}$ μεταβάλλεται σε συνεχή χρόνο σύμφωνα με το σύστημα των διαφορικών εξισώσεων (1) (οι οποίες εκφράζουν την εξίσωση $F = m\gamma$ της Φυσικής).*

3. Συναρτήσεις Κόστους. *Η συνάρτηση τρέχοντος κόστους είναι παντού ίση με το μηδέν, ενώ η συνάρτηση τερματικού κόστους είναι: $\hat{c}\,(x_1(T_0)) = -\,x_1(T_0)$*
Συμβολικά το πρόβλημα μπορεί να γραφεί ως εξής

$$(max) \qquad x_1(T_0)$$

περιορισμοί

$$\vec{x}\,(t) = f(\,\vec{x}\,(t),\alpha(t)\,)\,, \quad 0 \le t \le T_0\,,$$

$$\alpha(t) \in [\,0,M\,]\,,\ x_3(t) \ge 0\,,$$

$$\vec{x}\,(0) = (\,0,0,B\,)\,.$$

όπου B συμβολίζει το αρχικό βάρος του πυραύλου, και η f είναι μια συνάρτηση απο το $\mathbb{R}^4$ που ορίζεται από το δεξιό μέρος του συστήματος (1).

Μια πρακτική μέθοδος λύσης του προβλήματος (2) βρίσκεται αν δεχθούμε ότι η συνάρτηση $\alpha(t)$ παραμένει σταθερή εντός διαστημάτων της μορφής: $(t_0,t_1]$, $(t_1,t_2],\ldots,(t_{N-1},t_N]$ (όπου $t_0=0$, $T_0 = t_N$) και είναι δυνατόν να μεταβάλλεται μόνο κατά της χρονικές στιγμές $t_0, t_1, \ldots, t_{N-1}$. Εστω λοιπόν $\alpha(k) = \alpha(t)$, $t \in (t_k,t_{k+1})$, $k = 0,1,\ldots,N-1$. Τώρα το πρόβλημα (3) είναι ισοδύναμο με το παρακάτω πρόβλημα.

$$(\,max\,) \quad x_1(t_N)$$

$$\vec{x}\,(t_{k+1}) = g\,(k,\vec{x}\,(t_k),\alpha(k))$$

$$\alpha(k) \in [0,M]\,, \quad x_3(t_k) \ge 0\,, \quad k = 0,1,\ldots,N\,.$$

$$\vec{x}\,(0) = (0,0,B)\,.$$

όπου η $g(k,\vec{x}\,(t_k),\alpha(k))$ *μας δίδει την κατάσταση του συστήματος κατα την χρονική περίοδο* t_{k+1} . *Η συνάρτηση g μπορεί να υπολογιστεί εύκολα απο το σύστημα (1) . Το πρόβλημα (3) είναι ένα σύνηθες πρόβλημα δυναμικού προγραμματισμού.*

7.4. Πρόβλημα λειτουργίας Σταθμών Παραγωγής Ενέργειας. *Το πρόβλημα τούτο, είναι γνωστό στην αγγλική βιβλιογραφία σαν* **economic dispatch problem.** *Μια απλή μορφή του προβλήματος είναι η εξής : (Εικ.10)*

$\boxed{M_1}$

$\boxed{M_2}$ $\quad > \quad d(t)$ *Εικόνα 9*

$\boxed{M_3}$

Ενας σταθμός παραγωγής ηλεκτρικής ενέργειας, αποτελείται απο 3 μονάδες παραγωγής M_i , $i = 1, 2, 3.$

Η παραγωγή της μονάδας i είναι y_i *σε M.W.* , $i = 1, 2, 3$. *Το κόστος λειτουργίας της μονάδας i για μια χρονική περίοδο είναι* $c_i(y_i) = b_i\, y_i^2$ *όπου* b_i *είναι γνωστές σταθερές,* $i = 1,2,3.$

Η μονάδα 1 δουλεύει συνεχώς, ενώ οι μονάδες 2 και 3 είτε δουλεύουν είτε τις κρατάμε σε αναμονή (σβηστές) ανάλογα με την ζήτηση σε ενέργεια $d(t)$, $t = 0,1,2, \ldots. T$.

Το πρόβλημα είναι να βρεθεί η πολιτική λειτουργίας των μονάδων έτσι ώστε το ολικό κόστος λειτουργίας να ελαχιστοποιείται.

Διατύπωση του Προβλήματος.

1. Καταστάσεις–Αποφάσεις. *Δεχόμαστε ότι οι δυνατότητες παραγωγής των μονάδων περιγράφονται απο περιορισμούς της μορφής*

$$m_1 \le y_1 \le M_1,$$

και,

$$0 \le y_i \le M_i \quad i = 2, 3$$

όπου m_i , M_i , $i = 1, 2, 3$ *είναι γνωστοί αριθμοί. Η κατάσταση του συστήματος περιγράφεται απο το διάνυσμα* $(t,y_1(t),y_2(t),y_3(t))$ *και η απόφαση απο το διάνυσμα* $(\alpha_1(t),\alpha_2(t),\alpha_3(t))$ *οπου*

$$\alpha_1(t) \in [\,m_1,M_1] \,,\ \alpha_i(t) \in [0,M_i] \,,\ i = 2, 3$$

2. Η δυναμική του συστήματος *σχηματικά περιγράφεται ως εξής*

$$\boxed{t,\, y_1(t),\, y_2(t),\, y_3(t)} \longrightarrow \boxed{t+1,\, y_1(t+1),\, y_2(t+1),\, y_3(t+1)}$$

όπου,

$$y_i(t+1) = \alpha_i(t) \,,$$

$$\sum_1^3 y_i(t) \ge d(t) \,.$$

Συναρτήσεις κόστους.

$$c\,(t, y_1(t), y_2(t), y_3(t)) \;=\; \sum_{1}^{3} b_i\,(y_i(t))^2\,.$$

$$\hat{c}\,(T, y_1(T), y_2(T), y_3(T)) \;=\; 0\,.$$

Το πρόβλημα τώρα μπορεί να λυθεί με τις μεθόδους του Δυναμικού Προγραμματισμού.

8. Χρήση Βέλτιστων Πολιτικών στην Πράξη.

Οπως έχουμε δεί οι βέλτιστες πολιτικές είναι συναρτήσεις του χρόνου και των καταστάσεων $\alpha = \alpha(t,x_t)$. *Στην πράξη έχουμε δυο διαφορετικούς τρόπους για να εφαρμόσουμε μια βέλτιστη πολιτική.*

Ο πρώτος τρόπος συνίσταται στο να λύσουμε το πρόβλημα που μας ενδιαφέρει και να κρατήσουμε στην μνήμη ενός υπολογιστή όλες τις τιμές της πολιτικής $\alpha(t,x_t)$. *Κατόπιν, στην πράξη καθώς παρατηρούμε τις καταστάσεις του συστήματος στις διάφορες χρονικές περιόδους ανακαλούμε απο την μνήμη τις ανάλογες αποφάσεις. Το πλεονέκτημα αύτης της μεθόδου είναι ότι όλοι οι υπολογισμοί γίνονται εκ των προτέρων (off-line) και κατά την πρακτική εφαρμογή απλώς ανακαλούμε τις αποφάσεις που μας χρειάζονται. Σχηματικά, η μέθοδος περιγράφεται ως εξής*

ΜΝΗΜΗ Η/Υ

(t,x) $\alpha(t,x)$

$\alpha(t,x)$

ΕΛΕΓΚΤΗΣ   ΣΥΣΤΗΜΑ

(t,x)

Ο δεύτερος τρόπος συνίσταται στο να εκτελούμε τους απαραίτητους υπολογισμούς για τον καθορισμό των αποφάσεων $\alpha(t,x_t)$ *ταυτόχρονα (on line) με την λειτουργία του συστήματος. Η μέθοδος αυτή δεν παρέχει σοβαρά πλεονεκτήματα σε σχέση με την προηγούμενη όταν η δυναμική του συστήματος είναι προσδιοριστική (deterministic). Παρέχει όμως σοβαρά πλεονεκτήματα στην περίπτωση που δυναμική κίνηση του συστήματος είναι πιθανοθεωρητική (stochastic).*

Στην περίπτωση της προσδιοριστικής δυναμικής η μέθοδος τούτη πλεονεκτεί της πρώτης όταν λόγω της πολυπλοκότητας του προβλήματος χρησιμοποιούμε διάφορους κανόντες προσέγγισης των βέλτιστων αποφάσεων $\hat{\alpha}(t,x_t)$.

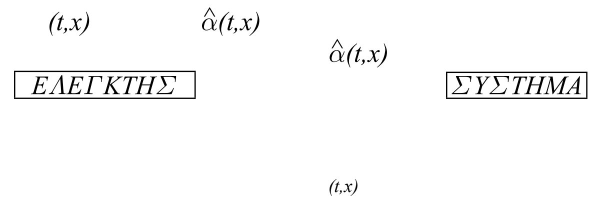
Επιπλέον, πολλές φορές στην πράξη όταν οι μεταβολές της κατάστασης x_t *είναι μικρές δεν ξαναυπολογίζουμε νέες τιμές για τις αποφάσεις.*

Σχηματικά, η μέθοδος περιγράφεται με την παρακάτω εικόνα, όπου υποθέτουμε ότι ο κόμβος $\boxed{K}$ *συμβολίζει ένα κανόνα ο οποίος ορίζει πότε η μεταβολή της κατάστασης είναι μεγάλη αρκετά έτσι ώστε να απαιτούνται νέοι υπολογισμοί.*

$\hat{\alpha}(t,x)\ ,\ t=1,\ldots,N-1$

Η/Υ   Κ

(t,x)



9. *Γενικά Αποτελέσματα για Πεπερασμένους Χώρους .*

Σε τούτο το τμήμα εξετάζουμε το γενικό πρόβλημα του πεπερασμένου ορίζοντα. Θα αποδείξουμε, για πεπερασμένους χώρους καταστάσεων και αποφάσεων, ότι η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο τιμών συγκλίνει σε πεπερασμένο αριθμό επαναλήψεων, καθώς και το γεγονός ότι το πρόβλημα είναι δυνατό να λυθεί με την μέθοδο του γραμμικού προγραμματισμού. Η σύγκλιση της μεθόδου των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών σε πεπερασμένο αριθμό βημάτων είναι αποτέλεσμα της σχέσης της με τη μέθοδο του γραμμικού προγραμματισμού. Οι μέθοδοι των διαδοχικών προσεγγίσεων συγκλίνουν κάτω από ορισμένες συνθήκες ομαλότητας (όχι αναγκαστικά σε πεπερασμένο αριθμό βημάτων), όταν οι χώροι καταστάσεων και αποφάσεων δεν ειναι πεπερασμένοι. Οι αποδείξεις ειναι ανάλογες με εκείνες του επόμενου κεφαλαίου και θά τις δούμε εκεί.

Το πρόβλημα του πεπερασμένου ορίζοντα με πεπερασμένους χώρους καταστάσεων και αποφάσεων ειναι ισοδύναμο με εκείνο του προσδιορισμού της διαδρομής ελάχιστου μήκους σε ένα δίκτυο που δεν περιέχει αρνητικούς κύκλους. Εστω λοιπόν ένα δίκτυο με N *κόμβους, όπου ονομάζουμε* 1 *τον αρχικό και* N *τον τελικό κόμβο. Για κάθε κόμβο* x *μας δίδονται τα σύνολα των βελών* $D(x)$ *καθώς και τα αντίστοιχα μήκη* $c(x,\overline{\alpha})$ *(ή* $c(\alpha)$ *),* $\alpha \in D(x)$ *) . Η εξίσωση βελτιστοποίησης είναι η εξής.*

$$v(x) = \min_{\alpha \in D(x)} \{c(x,\overline{\alpha}) + v(\overline{\alpha})\}, \quad x = 1, \dots, N-1, \tag{1}$$

$$v(N) = 0$$

ή ισοδύναμα

$$v = Tv \tag{2}$$

Ας θεωρήσουμε επίσης το σύστημα των ανισοτήτων

$$\xi \le T\xi \tag{3}$$

όπου ορίζουμε $\xi_1 \ge \xi_2$ *όταν* $\xi_1(x) \ge \xi_2(x)$, $x = 1,\dots, N$ *και* $\xi_1 > \xi_2$ *όταν* $\xi_1 \ge \xi_2$ *και* $\xi_1 \ne \xi_2$. Παρακάτω *θα μας χρειαστεί ο εξής συμβολισμός. Για δύο διανύσματα* ξ_1 *και* ξ_2 *ορίζουμε το διάνυσμα* $\xi_1 \vee \xi_2$ *(αντίστοιχα το* $\xi_1 \wedge \xi_2$ *) έτσι ώστε:* $(\xi_1 \vee \xi_2)(x) = \max\{\xi_1(x), \xi_2(x)\}$ *(αντίστοιχα ,* $(\xi_1 \wedge \xi_2)(x) = \min\{\xi_1(x), \xi_2(x)\}$*) για* $x = 1,\dots, N$.

Εστω S *το σύνολο των λύσεων της (3) , δηλαδή* $S = \{\xi \in R^n : \xi \le T\xi\}$ *. Θα αποδείξουμε οτι ισχύουν τα εξής αποτελέσματα:*

1. *Αν,* $\xi_1 \le \xi_2$ *τότε* $T\xi_1 \le T\xi_2$ *, δηλαδή ο τελεστής* T *ειναι* ***μονότονος*** *.*
2. *Αν,* ξ_1 *,* $\xi_2 \in S$ *τότε* $\xi_1 \vee \xi_2 \in S$ *.*
3. *Εστω* ξ^* *το μεγαλύτερο στοιχείο του* S *, συγκεκριμένα*

$$\xi^*(x) = \max\{\xi(x),\ \xi \le T\xi\}\ x = 1,\dots, N, \tag{4}$$

τότε το ξ^* *είναι η μεγαλύτερη λύση της (2) . Επίσης, αν* $\widetilde{\xi} \le \xi^*$ *τότε*

$$T\widetilde{\xi} \le \xi^*.$$

4. *Εστω* $\widetilde{\xi} \le \xi^*$, $E = \{x : \widetilde{\xi}(x) = \xi^*(x)\}$ *τότε* $\exists\, x_0 \in \{1,\dots,N\} - E :$ $(T\widetilde{\xi})(x_0) = \xi^*(x_0)$.

5. Το σημείο ξ^* είναι η μοναδική λύση του προβλήματος γραμμικού προγραμματισμού

$$(max) \quad \sum_{1}^{N} \lambda(x)\,\xi(x)$$

περιορισμοί

$$\xi(x) - \xi(\overline{\alpha}) \le c(x,\overline{\alpha}),\ \forall \alpha \in D(x),\ x = 1,\dots,N-1,$$
$$\xi(N) = 0.$$

6. Εαν το δίκτυο δεν περιέχει αρνητικούς κύκλους, τότε η εξίσωση (2) έχει μοναδική λύση και άρα: $\xi^* = v$

Τα αποτελέσματα 4, 6 συνεπάγονται ότι η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο τιμών συγκλίνει σε πεπερασμένο (το πολύ N) αριθμό βημάτων όταν η αρχική προσέγγιση ξ_0 είναι: $\xi_0 < \xi^*$.

Λήμμα 1. *i) Ο τελεστής T είναι μονότονος.*
ii) Αν $\xi_1,\ \xi_2 \in S$ τότε $\xi_1 \vee \xi_2 \in S$.

Απόδειξη. *i) Εστω $\xi_1 \le \xi_2$. Τότε*

$$(T\xi_1)(x) = \min_{\alpha \in D(x)} \left\{ c(x,\overline{\alpha}) + \xi_1(\overline{\alpha}) \right\} \le \min_{\alpha \in D(x)} \left\{ c(x,\overline{\alpha}) + \xi_2(\overline{\alpha}) \right\} =$$

$$= (T\xi_2)(x),\ \forall x = 1,\dots,N-1.$$

ii) Εστω $\xi_3 = \xi_1 \vee \xi_2$ όπου,

$$\xi_i(x) \le (T\xi_i)(x) = \min_{\alpha \in D(x)} \{c(x,\overline{\alpha}) + \xi_1(\overline{\alpha})\},\ i = 1, 2, x = 1,\dots,N-1.$$

Ας υποθέσουμε ότι $\xi_3(x) = \xi_1(x)$ τότε

$$\xi_3(x) = \xi_2(x) \le c(x,\overline{\alpha}) + \xi_2(\overline{\alpha}) \le c(x,\overline{\alpha}) + \xi_3(\overline{\alpha})\ \forall x, \alpha$$

δηλαδή,

$$\xi_3(x) \le (T\xi_3)(x) \tag{5}$$

Ανάλογα παίρνουμε ότι η (5) ισχύει όταν $\xi_3(x) = \xi_1(x)$ και η απόδειξη είναι πλήρης.

Λήμμα 2. *Αν $(x,x) \notin D(x)\ \forall x \in \{1,\dots,N-1\}$, τότε:*
i) Τό ξ^* είναι η μεγαλύτερη λύση της εξίσωσης $T\xi = \xi$.
ii) Αν $\widetilde{\xi} \le \xi^*$ τότε $T\widetilde{\xi} \le \xi^*$.

Απόδειξη. *i) Τό ξ^* είναι εξ ορισμού η μεγαλύτερη λύση των ανισοτήτων $\xi \le T\xi$, άρα αρκεί να αποδείξουμε ότι $\xi^* = T\xi^*$. Ας υποθέσουμε το αντίθετο, δηλαδή ας υποθέσουμε ότι*

$$\exists\, x_0 : \quad \xi^*(x_0) < (T\xi^*)(x_0) \tag{6}$$

Τότε μπορούμε να ορίσουμε το σημείο $\widetilde{\xi}$ έτσι ώστε

$$\widetilde{\xi}(x) = \begin{cases} \xi^*(x), & x \ne x_0 \\ (T\xi^*)(x_0), & x = x_0 \end{cases} \tag{7}$$

Εξ ορισμού έχουμε ότι

$$\xi^* < \widetilde{\xi} \le T\xi^* \tag{8}$$

Παρατηρούμε όμως, ότι

$$(T\xi^*)(x_0) = \min_{\alpha \in D(x_0)} \{c(x,\overline{\alpha}) + \xi^*(\overline{\alpha})\} \overset{(7)}{=} (\overline{\alpha}\, x_0,\ (x_0, x_0) \notin D(x_0))$$

$$= \min_{\alpha \in D(x_0)} \{c(x,\overline{\alpha}) + \widetilde{\xi}(\overline{\alpha})\} = (T\widetilde{\xi})(x_0) . \tag{9}$$

Επίσης,

$$(T\xi^*)(x) = \min_{\alpha \in D(x_0)} \{c(x,\overline{\alpha}) + \xi^*(\overline{\alpha})\} \leq (\widetilde{\xi} > \xi^*)$$

$$= \min_{\alpha \in D(x_0)} \{c(x,\overline{\alpha}) + \xi^*(\overline{\alpha})\}$$

$$= (T\widetilde{\xi})(x) . \tag{10}$$

Από τις (9), (10) παίρνουμε ότι

$$T\xi^* \leq T\widetilde{\xi} \tag{11}$$

και λόγω της (8) έχουμε ότι

$$\widetilde{\xi} \leq T\xi^* \leq T\widetilde{\xi}$$

δηλαδή $\widetilde{\xi} \in S$ *και* $\widetilde{\xi} > \xi^*$, *πράγμα που αποτελεί αντίφαση λόγω του ορισμού του* ξ^*.

ii) Πραγματικά, όταν το ξ^* *υπάρχει τότε* $\xi^* = T\xi^*$, *αλλά από το Λήμμα 1, (i)* $\widetilde{\xi} \leq \xi^* \Rightarrow T\widetilde{\xi} \leq T\xi^* = \xi^*$.

Πρόταση 1. *Εστω ότι υπάρχει το* ξ^* *και έστω ότι* $\widetilde{\xi} \leq \xi^*$. *Ορίζουμε το σύνολο :*

$$E = \{x : \widetilde{\xi}(x) = \xi^*(x)\} .$$

Τότε $\exists x_0 \notin E$, *έτσι ώστε* $(T\widetilde{\xi})(x_0) = \xi^*(x_0)$.

Απόδειξη. *Από το λήμμα 2 έχουμε ότι* $T\widetilde{\xi} \leq \xi^*$. *Εστω λοιπόν,* $\delta = \xi^* - T\widetilde{\xi} \leq (0,\ldots,0)$. *Το ζητούμενο ισχύει εάν* $\exists x_0 \notin E$, *έτσι ώστε* $\delta(x_0) = 0$.
Ας υποθέσουμε λοιπόν το αντίθετο, δηλαδή : $\forall x \notin E$ *έχουμε* $\delta(x) < 0$.
Εστω $r = max\{\delta(x) / \forall x : \delta(x) < 0\} < 0$, *και ας ορίσουμε το διάνυσμα* $\hat{\xi}$ *ως εξής :*

$$\hat{\xi}(x) = \begin{cases} \xi^*(x), & x \in E \\ \xi^*(x) - r, & x \notin E \end{cases} \tag{12}$$

Θα αποδείξουμε τώρα ότι : $\hat{\xi} \in S$, $(\delta\eta\lambda.\ \hat{\xi} \leq T\hat{\xi})$, *οπότε, δεδομένου ότι από την κατασκευή του* $\hat{\xi}$ *ισχύει* $\hat{\xi} > \xi^*$, *θα έχουμε καταλήξει σε άτοπο.*

Πραγματικά, $\hat{\xi} \in S$ *είναι ισοδύναμο με:*

$$\xi^*(x) \leq \min_{\alpha \in D(x)} \{c(x,\overline{\alpha}) + \hat{\xi}(\overline{\alpha})\} \qquad x = 1,\ldots,N-1 \tag{13}$$

Εξετάζουμε τις δυνατές περιπτώσεις για το x

1. $x \in E$. *Τότε:*

$$\hat{\xi}(x) = \xi^*(x) = T\xi^*(x) \leq T\hat{\xi}(x) \tag{14}$$

όπου η τελευταία ανισότητα οφείλεται στο ότι ο T είναι μονότονος.

2. $x \notin E$. *Εξετάζουμε τις υποπεριπτώσεις :*

i) $T\hat{\xi}(x) = c(x,\overline{\alpha}_0) + \hat{\xi}(\overline{\alpha}_0)$, *όπου* $\overline{\alpha}_0 \notin E$

τότε

$$T\hat{\xi}(x) = c(x,\overline{\alpha}_0) + \xi^*(\overline{\alpha}_0) - r$$

$$\geq \min_{\alpha \in D(x)} \{c(x,\overline{\alpha}) + \xi^*(\overline{\alpha})\} - r = \xi^*(x) - r = \hat{\xi}(x) ,$$

δηλαδή

$$T\hat{\xi}(x) \geq \hat{\xi}(x)\,. \tag{15}$$

(ii) Εχουμε ότι

$$T\hat{\xi}(x) = c(x,\overline{\alpha}_0) + \hat{\xi}(\overline{\alpha}_0)\,, \quad \text{όπου } \overline{\alpha}_0 \in E \tag{16}$$

τότε

$$T\hat{\xi}(x) = c(x,\overline{\alpha}_0) + \xi^*(\overline{\alpha}_0) \tag{17}$$

αλλά,

$$\begin{aligned} r \geq \delta(x) &= \xi^*(x) - T\hat{\xi}(x) \\ &= \xi^*(x) - min\, c(x,\overline{\alpha}) + \tilde{\xi}(\overline{\alpha})\, \} \\ &\geq \xi^*(x) - (c(x,\overline{\alpha}_0) + \tilde{\xi}(\overline{\alpha}_0)) \\ &= \xi^*(x) - (c(x,\overline{\alpha}_0) + \xi^*(\overline{\alpha}_0)) \end{aligned}$$

$$\Rightarrow \quad c(x,\overline{\alpha}_0) + \xi^*(\overline{\alpha}_0) \geq \xi^*(x) - r = \hat{\xi}(x) \tag{18}$$

Οι (17) και (18) συνεπάγονται ότι :

$$T\hat{\xi}(x) \geq \hat{\xi}(x) \tag{19}$$

Αρα, από τις (14) , (15) και (19) παίρνουμε

$$\hat{\xi}(x) \leq T\hat{\xi}(x) \quad \forall x = 1,2,\ldots,N-1\,,$$

δηλαδή $\hat{\xi} \in S$.

Πριν αποδείξουμε το αποτέλεσμα 5 θα εισάγουμε τον εξής ορισμό :

Εστω $S \subset \mathbb{R}^n$, *και έστω* $f = (f_1,\ldots,f_m) : S \to \mathbb{R}^m$.

Ενα σημείο $x \in S$ *λέγεται* ***αποδοτικό*** *(efficient) ως προς την* f *αν και μόνο αν :*
$\nexists\ y \in S : f(y) > f(x)$.

Η έννοια των αποδοτικών σημείων χρησιμοποιείται συχνά στην πολυκριτήρια ανάλυση. Διαισθητικά, αν υποθέσουμε ότι $f_1,f_2,\ldots,f_m$ *είναι διαφορετικά μέτρα απόδοσης, τότε ένα σημείο* x *είναι αποδοτικό αν δεν υπάρχει κάποιο άλλο σημείο* y, *το οποίο πετυχαίνει καλύτερες τιμές για* <u>*όλα*</u> *τα μέτρα απόδοσης* $f_1,\ldots,f_m$.

Ο παρακάτω χαρακτηρισμός των αποδοτικών σημείων χρησιμοποιείται για τον υπολογισμό τους.

<u>***Πρόταση 2.***</u> *Εστω* $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}^m$, $f(x) \leq M$, *και ας υποθέσουμε ότι το σύνολο* $f(X) - \mathbb{R}^m_+$ *είναι κυρτό. Τότε το σημείο* x_0 *είναι αποδοτικό αν και μόνο αν* $\exists\, \lambda = \lambda(x_0) \in \mathbb{R}^m$, $\lambda > \vec{0}$: $\lambda^T f(x) \leq \lambda^T f(x_0)$ $x \in S$.

<u>***Απόδειξη.***</u> *(* $\rightarrow$ *) Το* x_0 *είναι αποδοτικό* $\Leftrightarrow$ $f(x_0) \in b(f(X) - \mathbb{R}^m_+)$, *όπου* $b(E)$ *είναι το σύνορο του συνόλου* E.

$$f(X) - \mathbb{R}^m_+ = \{f(x) - (z_1,\ldots,z_m)\,,\ x \in S\,,\ z_i \in \mathbb{R}'_+\}\,.$$

Εφόσον το σύνολο $f(X) - \mathbb{R}^m_+$ *είναι κυρτό* $\Rightarrow$ $\exists$ *ένα στηρίζον υπερεπίπεδο στο σημείο* $f(x_0)$, *δηλαδή*

$$\exists\, \lambda = \lambda(f(x_0)) \in \mathbb{R}^m\,,\ \lambda \neq \vec{0} \ : \ \lambda^T w \leq \lambda^T f(x_0) \ \ \forall\, w \in f(X) - \mathbb{R}^m_+\,,$$

αλλά $w \in f(X) - \mathbb{R}^m_+ \Leftrightarrow w = f(x) - t$, *για κάποιο* $t \in \mathbb{R}^m_+$, *άρα*

$$\lambda^T(f(x) - t) \leq \lambda^T f(x_0)\,,\ \ \forall\, x \in S,\ \forall\, t \in \mathbb{R}^m_+\,. \tag{20}$$

Τώρα, αν πάρουμε $t = 0 \in \mathbb{R}^m_+$ *στην (20) , έχουμε ότι :*

$$\lambda^T f(x) \leq \lambda^T f(x_0) \quad \forall\, x \in S\,,$$

ενώ αν πάρουμε $x = x_0$ *στην (20) , έχουμε ότι :*

$$-\lambda^T t \leq 0 \ \ \forall\, t \in \mathbb{R}^m_+, \text{ και επειδή } \lambda \neq \vec{0}\,, \text{ έπεται ότι } \lambda > \vec{0}\,.$$

($\leftarrow$) Εστω ότι $\exists\ \overrightarrow{\lambda} > 0 : \lambda f(x) \leq \lambda f(x_0)\ \forall x \in S$, *και ας υποθέσουμε ότι το* x_0 *δεν είναι αποδοτικό σημείο, δηλ.* $\exists\, \hat{x} : f(\hat{x}) > f(x_0)$. *Τότε όμως :*

$$\lambda^T f(\hat{x}) > \lambda^T f(x_0) \quad \forall \lambda \in \mathbb{R}^m_+,\ \lambda_i > 0 \quad (\forall i),$$

πράγμα άτοπο, γιατί έχουμε δεχθεί ότι το x_0 *είναι αποδοτικό.*

Παρατήρηση. *Εχουμε δει ότι το διάνυσμα* ξ^*, *όταν υπάρχει, είναι αποδοτικό στο σύνολο* $S := \{\xi : \xi < T\xi\ ,\ \xi(N) = 0\}$. *Αμεση συνέπεια, λοιπόν, της πρότασης 2 (με* $X = S$, $f(\xi) = \xi$) *είναι η παρακάτω πρόταση.*

Πρόταση 2. *Το σημείο* ξ^* *είναι λύση του παρακάτω προβλήματος γραμμικού προγραμματισμού:*

$$(\max)\ z = \sum_{x=1}^{N-1} \lambda(x)\xi(x)$$

$$\xi(x) - \xi(\overline{\alpha}) \leq c(x,\overline{\alpha}),\quad x = 1,\ldots,N-1,\ \alpha \in D(x)$$
$$\xi(N) = 0.$$

Μένει να αποδείξουμε ότι όταν το δίκτυο δεν έχει κύκλους, τότε η εξίσωση βελτιστοποίησης $v = Tv$, $v(N) = 0$, *έχει μια μόνο λύση. Τότε, σύμφωνα με τα παραπάνω, θα ισχύει* $v = \xi^*$.

Πρόταση 4. *Εαν για κάθε κόμβο* $x = 1,\ldots, N-1$, *υπάρχει μια διαδρομή πεπερασμένου μήκου απο τον κόμβο* x *στον κόμβο* N, *τότε η εξίσωση βελτιστοποίησης* $v = Tv$, $v(N) = 0$, *έχει μια μόνο λύση.*

Απόδειξη. *Ας υποθέσουμε το αντίθετο, δηλαδή* $\exists\ v_1, v_2$, $v_1 \neq v_2$, *λύσεις της* $v = Tv$, $v(N) = 0$, *και έστω* π_1, π_2 *οι αντίστοιχες πολιτικές. Τότε:*

$$v_1(x) = c(x,\overline{\pi}_1(x)) + v_1(\overline{\pi}_1(x)) \quad \forall x \neq N,$$

και

$$v_2(x) \geq c(x,\overline{\pi}_1(x)) + v_2(\overline{\pi}_1(x)), \quad x \neq N.$$

Από τις παραπάνω σχέσεις παίρνουμε ότι

$$v_1(x) - v_2(x) \leq v_1(\overline{\pi}_1(x)) - v_2(\overline{\pi}_1(x)). \tag{21}$$

Επειδή το δίκτυο δεν περιέχει κύκλους, η πολιτική π_1 *οδηγεί στον τερματικό κόμβο* N *σε ένα πεπερασμένο αριθμό βημάτων. Ετσι, επαναλαμβάνοντας τον συλλογισμό που οδήγησε στην σχέση (21), παίρνουμε ότι*

$$v_1(x) - v_2(x) \leq 0,\quad x \neq N. \tag{22}$$

Ανάλογα, χρησιμοποιώντας την πολιτική π_2 *παίρνουμε ότι*

$$v_1(x) - v_2(x) \geq 0,\quad \forall x \neq N. \tag{23}$$

Ετσι, οι σχέσεις (22), (23) οδηγούν στο συμπέρασμα

$$v_1(x) = v_2(x),\quad \forall x \neq N,$$

και η απόδειξη είναι πλήρης.

Παρατηρήσεις. *1. Αν μελετήσουμε προσεκτικά την παραπάνω απόδειξη, βλέπουμε ότι η υπόθεση που απαιτείται για να ισχύει το αποτέλεσμα είναι ότι οι βέλτιστες πολιτικές δεν περιέχουν κύκλους, αλλά οδηγούν στον τερματικό κόμβο* N *ορίζοντας διαδρομές πεπερασμένου μήκους.*

2. Η ύπαρξη μοναδικής λύσης για την εξίσωση βελτιστοποίησης καθώς και η ύπαρξη βέλτιστων πολιτικών είναι άμεσες για το πρόβλημα με πεπερασμένους χώρους αποφάσεων και καταστάσεων. Ανάλογα αποτελέσματα ισχύουν, άμεσα, για το πρόβλημα με συμπαγείς χώρους αποφάσεων και καταστάσεων όταν οι συναρτήσεις κόστους ειναι συνεχείς. Σε πιο γενικά προβλήματα απαιτούνται επιπλέον υποθέσεις, βλπ. Dynkin and Yuskevich (1979).

10 . Ασκήσεις.

Ασκηση 1 . ***(α)*** *Να λυθεί με τις μεθόδους των διαδοχικών προσεγγίσεων,*

(i) στο χώρο τιμών

(ii) στο χώρο πολιτικών

το πρόβλημα διαδρομής ελάχιστου μήκους του σχήματος 1.

(β) *Να δοθεί η διατύπωση του προβλήματος μέσω γραμμικού προγραμματισμού.*

Σχήμα 1

Ασκηση 2 . *Εστω ένα πεπερασμένο δίκτυο, και έστω* $\delta_1 = (x_o, \ldots, y, \ldots z)$ *μια βέλτιστη διαδρομή απο το* x_o *στο* z *(τερματικό).*
Εστω $\delta_2 = (x_o, \ldots, y)$ *και* $\delta_3 = (y, \ldots, z)$ *δυο κομμάτια της διαδρομής.*

α) *Είναι η* δ_3 *βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το* y *και τελικό το* z *; (δώστε απόδειξη).*

β) *Είναι η* δ_2 *βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το* x_0 *και τελικότο* y *; δώστε απόδειξη).*

γ) *Αν η απάντηση στο (β) είναι καταφατική, υπάρχει εξίσωση βελτιστοποίησης που βασίζεται στην ιδιότητα (β) ; (αν ναι, να βρεθεί).*

Ασκηση 3 . *Ας θεωρήσουμε το γενικό πρόβλημα της διαδρομής ελάχιστου μήκους σε ένα δίκτυο. Να βρεθεί μια ικανή συνθήκη κάτω απο την οποία η βέλτιστη πολιτική είναι η μυωπική.*

Ασκηση 4 . *Ενας φοιτητής έχει αγοράσει ένα "όχι και πολύ καινούργιο" αυτοκίνητο, και σκοπεύει να κάνει τις διακοπές του πηγαίνοντας από την πόλη 1 στην πόλη 8 της εικόνας 1 του εισαγωγικού προβλήματος 1 . Σκέφτεται όμως ότι αν ακολουθήσει μια διαδρομή ανάμεσα σε δύο πόλεις όπου η θερμοκρασία είναι πολύ υψηλή, η μηχανή του αυτοκινήτου έχει κίνδυνο να χαλάσει. Υπολογίζει επίσης ότι* c_{ij} *είναι η μέγιστη θερμοκρασία που θα συναντήσει κατά τη διαδρομή από την πόλη* i *στην πόλη* j *, όπου* c_{ij} *είναι οι αριθμοί που στο παράδειγμα της εικόνας 1 χρησιμοποιήθηκαν σαν κόστη των αντίστοιχων δρόμων. Θέλει επομένως να βρει εκείνη τη διαδρομή που ελαχιστοποιεί τη μεγαλύτερη θερμοκρασία που θα συναντήσει.*

α) *Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού , και να εφαρμοσθεί για τα δεδομένα της εικόνας 1 .*

β) *Να διατυπωθεί η αντίστοιχη εξίσωση βελτιστοποίησης στη γενική περίπτωση δικτύων που δεν περιέχουν κύκλους .*

Ασκηση 5 . *Θεωρούμε το πρόβλημα εύρεσης της διαδρομής ελάχιστου μήκους από ένα αρχικό κόμβο (πηγή) σ' ένα τελικό (προορισμό), σ' ένα δίκτυο που δεν περιέχει κύκλους. Χρησιμοποιήστε τη μέθοδο του Δυναμικού Προγραμματισμού για να βρείτε τη δεύτερη–καλύτερη (αμέσως μεγαλύτερη από την ελάχιστη) διαδρομή από την πηγή στον προορισμό. (Υπόδειξη : 1. Η δεύτερη–καλύτερη διαδρομή πρέπει να διαφέρει από τη βέλτιστη τουλάχιστον σε ένα βέλος. 2. Εναλλακτικά, μπορείτε να αποδείξετε ότι η δεύτερη–καλύτερη διαδρομή από τον κόμβο* j *στον προορισμό επιτυγχάνεται αν πάμε από το* j *σε κάποιο* k *μέσω του βέλους* (j,k) *, και μετά στον προορισμό ακολουθώντας είτε την ελάχιστη, είτε τη δεύτερη–καλύτερη διαδρομή από τον* k *στον προορισμό).*

Ασκηση 6. *Θεωρούμε ένα* $m \times n$ *πίνακα* $A = \{a_{ij}\}$. *Θέλουμε να ξεκινήσουμε από τη θέση (1,1) και να φτάσουμε στη (m,n) προχωρώντας κάθε φορά ένα βήμα είτε προς τα δεξιά είτε προς τα κάτω, αλλά κατά τέτοιο τρόπο ώστε το άθροισμα των* a_{ij} *των στοιχείων που θα διασχίσουμε να είναι όσο το δυνατό μικρότερο. Δείξτε πώς μπορούμε να υπολογίσουμε τις βέλτιστες διαδρομές.*

Ασκηση 7. *Εστω το πρόβλημα*

$$(\min) \quad w = \sum_{t=0}^{3} [2x^2(t) + \alpha^2(t)]$$

$$\begin{aligned} \pi. \quad & x(t+1) = x(t) - \alpha(t) \\ & x(t) \in X = [0,1] \quad \forall\, t, \\ & \alpha(t) \in A(x(t)) = [0,1], \\ & x(0) = 1. \end{aligned}$$

Να βρεθεί η βέλτιστη πολιτική $\pi = (\alpha_o, \alpha_1, \alpha_2)$.

Ασκηση 8. *Εστω το πρόβλημα*

$$(\max) \quad \sum_{j=1}^{n} c_j x_j$$

$$\begin{aligned} \pi. \quad & \alpha_1 \le x_1 \le b_1 \\ & \alpha_j x_j \le x_{j+1} \le b_j x_j, \qquad j = 1,2,\ldots,n-1, \end{aligned}$$

όπου c_j, α_j, b_j *είναι γνωστοί αριθμοί.*
Να δοθεί η διατύπωση του προβλήματος μέσω Δυναμικού Προγραμματισμού.

Ασκηση 9. *Εστω το πρόβλημα*

$$\begin{aligned} \max \quad & 3x + 5y \\ \pi. \quad & 2x + y \le 10 \\ & x + 2y \le 6 \\ & x \in \{0,1,4,6\} \\ & y \in [0,\infty]. \end{aligned}$$

Να δοθεί η διατύπωση του προβλήματος με τη μέθοδο του Δυναμικού Προγραμματισμού.

Ασκηση 10. *Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης*

$$\min \quad \sum_{j=1}^{N} \frac{x_j^2}{c_j},$$

$$\pi. \quad \sum_{j=1}^{N} x_j = A,$$

$$x_j \ge 0, \quad j = 1,\ldots,N,$$

όπου $c_j > 0$. *Χρησιμοποιώντας μεθόδους Δ.Π., δείξτε ότι αν ξέρουμε τις τιμές* $x_1, x_2, \ldots, x_{t-1}$, *η βέλτιστη τιμή για το* x_t *είναι :*

$$x_t = c_t \cdot \frac{A - x_1 - \cdots - x_{t-1}}{c_t + \cdots + c_N}.$$

Ασκηση 11. *Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης*

$$\max \quad x_1^{c_1} \cdot x_2^{c_2} \cdots x_N^{c_N} \;,$$

$$\pi. \quad \sum_{j=1}^{N} x_j \;=\; A \;,$$

$$x_j \;\geq\; 0\;, \quad j = 1,\ldots,N\;,$$

όπου $c_j \;>\; 0$ *. Χρησιμοποιώντας μεθόδους Δ.Π., βρείτε τη μορφή της βέλτιστης λύσης* $x_1^*, x_2^*, \ldots, x_N^*$ *. Δείξτε πώς εξαρτάται η βέλτιστη τιμή της αντικειμενικής συνάρτησης από το* A *.*

Ασκηση 12. *Θεωρούμε τους αριθμούς* $\alpha_1,\alpha_2,\ldots,\alpha_n$, *όπου* $\alpha_\kappa \geq 0 \;\; k = 1,2,\ldots,n$, *και* $\sum_{\kappa=1}^{n} \alpha_\kappa = A$. *Να αποδείξετε με επιχειρήματα Δυναμικού Προγραμματισμού ότι*

$$[(\tfrac{1}{n}\sum_{\kappa=1}^{n} \alpha_\kappa\}]^n \geq \prod_{\kappa=1}^{n} \alpha_\kappa \;.$$

Ασκηση 13. *Εστω το πρόβλημα*

$$\max \quad \sum_{t=1}^{n} (x_t)^2$$

$$\pi. \quad \sum_{t=1}^{n} (x_t) = 1$$

$$x_t \geq 0$$

Να λυθεί με τη μέθοδο του Δυναμικού Προγραμματισμού.

Ασκηση 14. *Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες* $E_1,E_2,\ldots E_n$. *Ο χρόνος εκτέλεσης της* E_i *είναι γνωστή σταθερά* T_i, $i = 1,2,\ldots,n$. *Για την εκτέλεση των εργασιών υπάρχει ένας σταθμός εξυπηρέτησης. Υποθέτουμε ότι η εκτέλεση των εργασιών αρχίζει τη χρονική στιγμή* $t = 0$ *. Κριτήριο βελτιστοποίησης είναι ο συνολικός χρόνος ροής (flow–time), που ορίζεται σαν το άθροισμα* $\sum_{i=1}^{n} t_i$ *, όπου* t_i *είναι η χρονική στιγμή που τελείωσε η εκτέλεση της εργασίας* E_i *.*

Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που ελαχιστοποιεί το συνολικό χρόνο ροής.

Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

Παράδειγμα. *Αν* $n = 3$, $T_1 = 2$, $T_2 = 4$, $T_3 = 5$, *και η εξυπηρέτηση γίνει με τη σειρά* 2,3,1 , *ο χρόνος ροής θα είναι* $4 + (4 + 5) + (4 + 5 + 2) = 4 + 9 + 11 = 24$.

Ασκηση 15. *Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες* $E_1,E_2,\ldots E_n$. *Ο χρόνος εκτέλεσης της* E_i *είναι γνωστή σταθερά* T_i, $i = 1,2,\ldots,n$. *Μπορούμε να εκτελούμε το πολύ μια εργασία (υπάρχει ένας μόνο σταθμός εξυπηρέτησης) και αν η εργασία* E_i *τελειώσει τη χρονική στιγμή t έχουμε ένα κέρδος* $R_i(t)$. *Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που μεγιστοποιεί το συνολικό κέρδος.*
Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

Ασκηση 16. *Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες* $E_1,E_2,\ldots E_n$. *Ο χρόνος εκτέλεσης της* E_i *είναι γνωστή σταθερά* T_i, $i = 1,2,\ldots,n$. *Για την εκτέλεση των εργασιών υπάρχουν k όμοιοι σταθμοί εξυπηρέτησης. Υποθέτουμε ότι η εκτέλεση των εργασιών αρχίζει τη χρονική στιγμή* $t = 0$ *. Κριτήριο βελτιστοποίησης είναι ο συνολικός χρόνος που απαιτείται μέχρις ότου να τελειώσουν όλες οι εργασίες (makespan).*

Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που ελαχιστοποιεί τον παραπάνω συνολικό χρόνο .

Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

Ασκηση 17. *Ενα αεροπλάνο πρέπει να μεταφέρει υλικά σε μια επιστημονική μονάδα στο Νότιο Πόλο. Υπάρχουν τρεις τύποι υλικών, που για ευκολία τους ονομάζουμε 1 , 2 και 3 . Επειδή κάθε είδος υλικού έχει διαφορετική αξία ως προς το βαθμό αναγκαιότητάς του, μοντελοποιούμε τις διαφορές αυτές υποθέτοντας ότι βάρος* α *από το υλικό τύπου* i *έχει αξία ίση με* $R_i(\alpha)$ *, όπου :*

$$R_1(\alpha) = \alpha^2 , \quad R_2(\alpha) = \alpha^2 + 2\alpha , \quad R_3(\alpha) = 2\alpha .$$

Το μεγαλύτερο βάρος που μπορεί να μεταφέρει το αεροπλάνο είναι 10 τόνοι.

Να βρεθούν οι ποσότητες από κάθε τύπο υλικού που πρέπει να φορτωθούν στο αεροπλάνο, έτσι ώστε να μεγιστοποιηθεί η συνολική αξία των μεταφερόμενων στη βάση υλικών, χρησιμοποιώντας τη μέθοδο του Δυναμικού Προγραμματισμού.

Ασκηση 18 . *Θεωρούμε ένα δίκτυο επικοινωνίας* N *κόμβων, όπου οι κόμβοι συμβολίζουν πόλεις και τα βέλη τηλεφωνικές γραμμές μεταξύ των πόλεων. Κατά τα γνωστά συμβολίζουμε με* $D(x)$ *το σύνολο των βελών που έχουν σαν αρχικό κόμβο τον* x *,* $x = 1,2,\ldots,N-1$ *. Υποθέτουμε ότι θέλουμε να στείλουμε ένα μήνυμα από τον κόμβο 1 στον κόμβο* N*. Για κάθε βέλος* $(x,\alpha) \in D(x)$ *, η πιθανότητα ότι το μήνυμα, περνώντας από τον* x *, θα φτάσει με επιτυχία στον* $\overline{\alpha}(x)$ *, είναι γνωστή και ίση με* $p(x,\alpha)$ *.*

Θέλουμε να βρούμε τη διαδρομή από τον 1 στον N *, που έχει τη μεγαλύτερη πιθανότητα επιτυχούς μετάδοσης του μηνύματος.*

Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

Ασκηση 19 . *Σε μια επιτροπή συμμετέχουν* R *αντιπρόσωποι. Ολη η χώρα διαιρείται σε* s *περιοχές, με* $s < R$ *. Κάθε περιοχή* j *έχει πληθυσμό* p_j *. Για να είναι αυστηρά αναλογική η αντιπροσώπευση κάθε περιοχής στην επιτροπή, θα έπρεπε η περιοχή* j *να πάρει* $r_j = R \cdot p_j / \sum_{i=1}^{s} p_i$ *αντιπροσώπους. Αυτή όμως η κατανομή δεν είναι δυνατή, γιατί γενικά τα* r_j *δεν έχουν ακέραιες τιμές. Το πρόβλημα επομένως είναι να βρεθεί μια κατανομή αντιπροσώπων* y_j *,* $j = 1,\ldots,s$ *στις περιοχές, κατά τρόπο ώστε να ελαχιστοποιηθεί η μέγιστη απόλυτη διαφορά μεταξύ του* y_j *και της "ιδανικής" τιμής* r_j *, που παρατηρείται στις περιοχές, δηλαδή να ελαχιστοποιηθεί η παράσταση* $\max (|y_1 - r_1| , |y_2 - r_2| ,\ldots, |y_s - r_s|)$ *.*

α) *Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.*

β) *Να γίνει αριθμητική εφαρμογή όταν* $R = 4$ *,* $s = 3$ *,* $r_1 = 0.6$ *,* $r_2 = 2.2$ *,* $r_3 = 1.2$*. Να συζητηθεί ποιοτικά η λύση που προκύπτει.*

Ασκηση 20 . *Θεωρούμε την εξής επέκταση για το γνωστό πρόβλημα μεταφοράς του γραμμικού προγραμματισμού. Υπάρχουν* m *σταθμοί αποστολής, καθένας από τους οποίους έχει διαθέσιμη ποσότητα* S_i *, και* n *σταθμοί ζήτησης, καθένας από τους οποίους ζητά ποσότητα* D_j *. Ισχύει η γνωστή παραδοχή ότι* $\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} D_j$ *, όπως επίσης ότι τα* S_i *,* D_j *, και οι ποσότητες* x_{ij} *που αποστέλλονται από το σταθμό αποστολής* i *στο σταθμό ζήτησης* j *, πρέπει να είναι ακέραιοι αριθμοί. Το κόστος μεταφοράς* x_{ij} *μονάδων είναι γενικά μη γραμμική συνάρτηση* $c_{ij}(x_{ij})$ *,* $i = 1,\ldots,m$ *,* $j = 1,\ldots,n$ *.*

α) *Μοντελοποιήστε το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού, για* $m = 2$ *. (<u>Υπόδειξη :</u> Μπορείτε να χρησιμοποιήσετε την ισότητα συνολικής αποστολής και συνολικής ζήτησης, για να πάρετε μονοδιάστατη μεταβλητή κατάστασης).*

β) *Δείξτε πώς μεταβάλλεται η μοντελοποίηση στο α) όταν* $m = 3$ *.*

Ασκηση 21 . *Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης .*

$$\min \quad z = \sum_{i=1}^{N} x_i \ ,$$

$$\pi. \quad x_i + x_{i+1} \ \geq \ d_i \ , \ i = 1,2,\dots,N-1 \ , \quad x_i \ \geq \ 0 \ .$$

Δείξτε ότι η ελάχιστη τιμή της f *είναι ίση με* $f^* = \max_{i=1,\dots,N} d_i$, *αρκεί ένα από τα* d_i *να είναι θετικό.*

Ασκηση 22. *Επαναλάβετε την άσκηση 21 όταν οι πρώτοι* $N-1$ *περιορισμοί αντικατασταθούν από τους* $x_i + x_{i+1} + x_{i+2} \ \geq \ d_i \ , \ i = 1,2\dots,N-2$.

Ασκηση 23. *Χρησιμοποιώντας τη μέθοδο του Δυναμικού Προγραμματισμού και με την εισαγωγή κατάλληλων βοηθητικών μεταβλητών, μοντελοποιήστε το παρακάτω πρόβλημα ελαχιστοποίησης*

$$\min_{x_i \geq 0} \left\{ \frac{x_1}{x_2+x_3} + \frac{x_2}{x_3+x_4} + \cdots + \frac{x_{n-1}}{x_n+x_1} + \frac{x_n}{x_1+x_2} \right\} .$$

Ασκηση 24. *Θεωρούμε το πρόβλημα ελαχιστοποίησης*

$$\min \quad \sum_{k=1}^{N} g_k(x_k,x_{k+1}) + \sum_{k=1}^{N} h_k(x_k) \ ,$$

$$\pi. \quad x_{n+1} \ = \ x_1 \ , \quad 0 \leq x_k \leq b_k \ , \ k = 1,\dots,N \ , \quad \sum_{k=1}^{N} \phi_k(x_k) \ \geq \ c \ ,$$

όπου κάθε $\phi_k(x)$ *είναι μια γνωστή γνήσια αύξουσα συνάρτηση του* x , *με* $\phi_k(0) = 0$.
Εισάγουμε το παρακάτω βοηθητικό πρόβλημα

$$\min \quad g(y,x_2) \ + \ g(x_2,x_3) \ + \cdots + \ g(x_{N-1},x_N) \ + \ g(x_N,z) \ + \ \sum_{k=2}^{N} h_k(x_k) \ ,$$

$$\pi. \quad \sum_{k=2}^{N} \phi_k(x_k) \ \geq \ c \ , \quad 0 \leq x_k \leq b_k \ , \ k = 2,\dots,N \ .$$

Δείξτε ότι αν το ελάχιστο του παραπάνω προβλήματος ως προς $x_2, \dots , x_N$ *είναι ίσο με* $v(y,z,c)$, *τότε το ελάχιστο του αρχικού προβλήματος είναι ίσο με*

$$\min_{0 \leq x_1 \leq b_1} v(x_1,x_1,c-\phi_1(x_1)) \ .$$

Ασκηση 25. *Για την εύρεση του ελάχιστου στην άσκηση 24 , θεωρείστε την παρακάτω ακολουθία συναρτήσεων για* $t = 2,3,\dots,N-1$

$$v_t(y,z,c) \ = \min_{x_k,\, k=t,\dots N} \left\{ g(y,x_t) + g(x_t,x_{t+1}) + \cdots + g(x_{N-1},x_N) + g(x_N,z) + \sum_{k=t}^{N} h_k(x_k) \right\}$$

$$v_N(y,z,c) \ = \min_{x_N} \left\{ g(y,x_N) + g(x_N,z) + h_N(x_N) \right\} ,$$

όπου τα $x_2,\dots,x_N$ *υπόκεινται στους περιορισμούς που αναφέρθηκαν στην άσκηση 24 , ενώ για τις επιτρεπτές τιμές του* c *ισχύει η υπόθεση ότι* $\sum_{k=t}^{N} \phi_k(b_k) \ \geq \ c$, *όπου* b_k *είναι γνωστοί θετικοί αριθμοί.*

Δείξτε ότι η ισοδύναμη με το παραπάνω πρόβλημα εξίσωση βελτιστοποίησης είναι

$$v_t(y,z,c) \;=\; \min_{x_t \in D_t} \left[g(y,x_t) + h_t(x_t) + v_{t+1}(x_t,z,c - \phi_t(x_t)) \right] \;,$$

όπου ο χώρος αποφάσεων D_t *για το* x_t *καθορίζεται από τις παρακάτω σχέσεις :*

$$0 \le x_t \le b_t$$
$$\sum_{k=t+1}^{N} \phi_k(b_k) \;\ge\; c - \phi_t(x_t) \,.$$

Ασκηση 26. *Μοντελοποιήστε τα παρακάτω προβλήματα με τη μέθοδο του Δυναμικού Προγραμματισμού, όταν το κριτήριο βελτιστοποίησης είναι i. ελαχιστοποίηση και ii. μεγιστοποίηση. Με* z *συμβολίζουμε την αντικειμενική συνάρτηση.*

α) $z \;=\; cx_1^2 \;+\; (x_1 + cx_2)^2 \;+\; (x_1 + x_2 + cx_3)^2 \;+ \cdots +\; (x_1 + x_2 + \cdots + x_{N-1} + cx_N\,)^2 \,,$
υ.π. $x_1^2 \;+\; x_2^2 \;+ \cdots +\; x_N^2 \;=\; 1 \,.$

β) $z \;=\; cx_1^2 \;+\; (x_1 + cx_2)^2 \;+\; (x_1 + cx_2 + c^2x_3)^2 \;+ \cdots +\; (x_1 + cx_2 + \cdots + c^{N-2}x_{N-1} + c^{N-1}x_N\,)^2,$
υ.π. $x_1^2 \;+\; x_2^2 \;+ \cdots +\; x_N^2 \;=\; 1 \,.$

γ) $z \;=\; cx_1^2 \;+\; (x_1 + cx_2)^2 \;+\; [x_1 + cx_2 + (c + d)x_3]^2 \;+ \cdots$
$\quad +\; [x_1 + cx_2 + (c + d)x_3 + (c + 2d)x_4 + \cdots + (c + (N - 2)d)x_N\,]^2 \,,$
υ.π. $x_1^2 \;+\; x_2^2 \;+ \cdots +\; x_N^2 \;=\; 1 \,.$

Ασκηση 27. *Να λυθεί με τη μέθοδο του Δυναμικού Προγραμματισμού το πρόβλημα*

$$\max \qquad \sum_{i=1}^{N} c_i x_i \,,$$

$$\pi. \qquad \sum_{i=1}^{N} x_i^2 \;=\; 1 \,,$$
$$0 \le x_1 \le x_2 \le \ldots \le x_N \,,$$

όπου $c_i\,,\; i = 1,\ldots,N$ *είναι γνωστοί μη αρνητικοί συντελεστές.*

Ασκηση 28. *Σε προβλήματα εύρεσης ελάχιστων και μέγιστων ιδιοτιμών πινάκων (οπως π.χ. στον πίνακα Jacobi) προκύπτει συχνά η παρακάτω τετραγωνική μορφή, που πρέπει να ελαχιστοποιηθεί ή να μεγιστοποιηθεί ως προς* $x_i\,,\; i = 1,\ldots,N$*.*

$$f_N(\underline{x}) \;=\; \sum_{i=1}^{N} c_i x_i^2 \;+\; 2\sum_{i=1}^{N-1} b_i x_i x_{i+1} \,,$$

όπου τα x_i *είναι περιορισμένα στη N–διάστατη μοναδιαία σφαίρα* S*, δηλαδή* $\sum_{i=1}^{N} x_i^2 \;=\; 1$*.*

Θεωρούμε τις παρακάτω ακολουθίες

$$v_N(c) \;=\; \max_{\underline{x} \in S} \left\{ f_N(\underline{x}) \;+\; 2cx_N \right\} ,$$

$$w_N(c) \;=\; \min_{\underline{x} \in S} \left\{ f_N(\underline{x}) \;+\; 2cx_N \right\} .$$

Χρησιμοποιώντας μεθόδους του Δυναμικού Προγραμματισμού βρείτε αναδρομικές σχέσεις που συνδέουν τα $v_N(c)\,,\; w_N(c)$ *με τα* $v_{N-1}(c)\,,\; w_{N-1}(c)$ *αντίστοιχα.*

Ασκηση 29. *Λύστε την άκηση 28 όταν*

$$f_N(\underline{x}) \;=\; \sum_{i=1}^{N} c_i x_i^2 \;+\; 2\sum_{i=1}^{N-1} b_i x_i x_{i+1} \;+\; 2\sum_{i=1}^{N-2} d_i x_i x_{i+2}\,.$$

Ασκηση 30. *Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης*

$$\begin{array}{ll} \min & F_N(x_1,x_2,\ldots,x_N) \;=\; \phi_1(x_1) + \phi_2(x_2) + \cdots + \phi_N(x_N) \\ \pi. & x_i \ge 0 \quad i = 1,\ldots,N \;, \\ & x_1 \ge r_1 \;, \\ & x_1 + x_2 \ge r_2 \;, \\ & \vdots \\ & x_1 + x_2 + \cdots x_N \ge r_N \;. \end{array}$$

Ορίζουμε την ακολουθία

$$f_k(z) \;=\; \min_{\underline{x}} \sum_{i=k}^{N} \phi_i(x_i)\;,\ \text{με νέους περιορισμούς :}$$

$$\begin{array}{l} x_i \ge 0 \quad i = 1,\ldots,N \;, \\ x_k \ge r_k - z \;, \\ x_k + x_{k+1} \ge r_{k+1} - z \;, \\ \vdots \\ x_k + \cdots + x_N \ge r_N - z \;, \end{array}$$

όπου $z \ge 0$, $k = 1,2,\ldots,N$. *Δείξτε ότι :*

$$f_k(z) \;=\; \min_{x_k \ge 0,\; x_k \ge r_k - z} \left[\phi_k(x_k) \;+\; f_{k+1}(z + r_k)\right],\ k = 1,2,\ldots,N\,,$$

και επομένως ότι $\min_{\underline{x}} F_N(x_1,\ldots,x_N) \;=\; f_1(0)$.

Ασκηση 31. *Χρησιμοποιώντας επιχειρήματα Δυναμικού Προγραμματισμού βρείτε μια μέθοδο για να προσεγγίσετε τη συνάρτηση* $f(x)$ *στο διάστημα* $[a,b]$ *μέσω της γραμμικής συνάρτησης* $rx + s$, *σύμφωνα με τα εξής μέτρα απόκλισης :*

α) $\int_a^b (f(x) - rx - s)^2\, dx$

β) $\max_{a \le x \le b} |f(x) - rx - s|$.

<u>Υπόδειξη</u>. *Βρείτε μια αναδρομική διαδικασία για να προσδιορίσετε τα* r,s *που ελαχιστοποιούν τα παραπάνω μέτρα απόκλισης.*

Ασκηση 32. *Γενικεύστε το αποτέλεσμα της προηγούμενης άσκησης, για συνάρτηση προσέγγισης όχι γραμμική, αλλά πολυωνυμική βαθμού* n .

ΚΕΦΑΛΑΙΟ 2
ΠΡΟΒΛΗΜΑΤΑ ΑΠΕΙΡΟΥ ΟΡΙΖΟΝΤΑ ΜΕ ΠΡΟΣΔΙΟΡΙΣΤΙΚΗ ΔΥΝΑΜΙΚΗ

1. Εισαγωγή.

Το γενικό πρόβλημα άπειρου ορίζοντα έχει ως εξής

$$\begin{array}{ll} (\min\ \acute{\eta}\ \max) & w = w(l_\infty) \\ \acute{o}\pi o \upsilon & x_{t+1} = \overline{\alpha}_t\,,\quad t = 0,1,2,\ldots, \\ & x_t \in X\,,\quad \alpha_t \in D(x_t) \qquad t = 0,1,\ldots. \end{array}$$

Αρχικό σημείο : x_0.

Το $l_\infty = (x_0,\alpha_0,x_1,\alpha_1,\ldots)$ *είναι μια διαδρομή απείρου (αριθμήσιμου) μήκους και* $w = w(l_\infty)$ *είναι ένα μέτρο απόδοσης του συστήματος π.χ.* $w(l_\infty) = \sum_0^\infty c_t(x_t,\alpha_t)$, *το οποίο αντιστοιχεί στο μέτρο απόδοσης* $\sum_0^{N-1} c_t(x_t,\alpha_t) + \hat{c}_N(x_N)$ *του προβλήματος με πεπερασμένο ορίζοντα.*

Το πρόβλημα του μη πεπερασμένου ορίζοντα αποτελεί μία προσέγγιση για ένα πρόβλημα πεπερασμένου ορίζοντα, όταν το μήκος του ορίζοντα N είναι μεγάλο. Σαν μαθηματική διαδικασία είναι καλά ορισμένο και αποτελεί ένα ισχυρό εργαλείο για την μελέτη διαδικασιών που εκτυλίσσονται πάνω σε ένα ορίζοντα (χρονικό διάστημα) μεγάλου μήκους. Εξετάζουμε τα πλεονεκτήματα / μειονεκτήματα της προσέγγισης προβλημάτων με την μέθοδο του μη πεπερασμένου ορίζοντα στο επόμενο συγκεκριμένο πρόβλημα.

1.1 Πρόβλημα Κατανομής Πόρων. *Ας θεωρήσουμε τη παρακάτω γενίκευση του γνωστού προβλήματος πεπερασμένου ορίζοντα.*

$$(\max) \qquad w = \sum_{k=0}^{N-1} R_k(x_k,y_k)$$

$$\pi. \qquad x_{k+1} = g_k(x_k,y_k) = c\,(x_k - y_k) + d\,y_k\,,$$

$$y_k \in [0,x_k]\,,\ k = 0,1,\ldots,N-1\,,$$

Αρχικό σημείο: $x_0 = B$.

Οπου c, d *είναι γνωστές σταθερές. Στό κεφάλαιο 1 εξετάσαμε την ειδική περίπτωση πού* $c = 1$ *και* $d = 0$. *Ας ονομάσουμε* $P(B,N)$ *το παραπάνω πρόβλημα. Οταν το* N *είναι μεγάλος αριθμός, τότε το πρόβλημα* $P(B,N)$ *μπορεί να προσεγγιστεί από το* $P(B)$ $(:= P(B,\infty))$ *παρακάτω :*

$$(\max) \qquad w = \sum_{k=0}^{\infty} R_k(x_k,y_k)$$

$$\pi. \qquad x_{k+1} = g(x_k,y_k) = c\,(x_k - y_k) + d\,y_k\,,$$

$$y_k \in [0,x_k]\,,\ k = 0,1,\ldots\,,$$

Αρχικό σημείο: $x_0 = B$.

Ας ορίσουμε τις παρακάτω συναρτήσεις τιμών

$$v(t,N,x) = \sup_{y_t,\ldots y_{N-1}} \left\{ \sum_{k=t}^{N-1} R_k(x_k,y_k) \,/\, x_{k+1} = g_k(x_k,y_k)\,,\ y_k \in [0,x_k]\,,\ k = t,\ldots,N\text{-}2\,,\ x_k = x \right\} \qquad (1)$$

$$v(x) \;=\; \sup_{y_0,y_1,\ldots} \left\{ \sum_{k=0}^{\infty} R_k(x_k,y_k) \,/\, x_{k+1} \;=\; g_k(x_k,y_k)\,,\;\; y_k \in [0, x_k]\,,\;\; k=0,1,\ldots\;, x_0 = x \right\} \tag{2}$$

Αν κάνουμε την επιπλέον υπόθεση ότι οι συναρτήσεις $R_k(\cdot\,,\cdot)$ είναι συνεχείs τότε το sup στην (1) μπορεί να αντικατασταθεί με max . Η συνάρτηση $v(t,N,x)$ και η βέλτιστη πολιτική υπολογίζονται μέσω των παρακάτω εξισώσεων βελτιστοποίησης

$$v(t,N,x) \;=\; \max_{\alpha \in [0,x]} \{R_t(x,\alpha) \;+\; v(g_k(x,\alpha)\} \quad , \qquad t \;=\; 0,\ldots,N-1 \tag{3}$$

$$v(N,N,x) \;=\; 0\,. \tag{4}$$

Η λύση του προβήματοs δίδεται από τη τιμή $v(0,N,B)$.

Αs εξετάσουμε την πρώτη “δυσκολία" που υπάρχει στα προβλήματα μη–πεπερασμένου ορίζοντα. Στο πρόβλημα $P(B,N)$ ζητάμε το μέγιστοποιόν “σημείο" : $\{y_1,\ldots,y_N\}$, ενώ στο πρόβλημα $P(B)$ το μέγιστοποιόν “σημείο" (όταν υπάρχει) είναι μία μη πεπερασμένη ακολουθεία: $\{y_0,y_1,\ldots\}$. Ετσι, το μέγιστο τηs παράστασηs στο δεξιό μέροs τηs (2) δεν είμαστε βέβαιοι εκ των πρωτέρων ότι υπάρχει κάτω απο τιs υποθέσειs που έχουμε κάνει. Αντίθετα, στην (1) η ύπαρξη του μέγιστου είναι σίγουρη επειδή η συνάρτηση $\sum_{k=0}^{N-1} R_k(x_k,y_k)$ είναι συνεχήs πάνω στο συμπαγέs σύνολο $[0,B]^N$, και αυτό λόγω τηs υπόθεσηs ότι οι $R_k(\cdot\,,\cdot)$ είναι συνεχείs. Στην (2) όμωs, η υπόθεση τηs συνέχειαs τηs R δεν φτάνει για να είμαστε βέβαιοι ότι η συνάρτηση $\sum_{k=0}^{\infty} R_k(x_k,y_k)$ έχει μέγιστο. Το νόημα τηs μη ύπαρξηs του μεγίστου στην (2) είναι ότι δεν υπάρχει μια πολιτική κατανομήs του πόρου η οποία πετυχαίνει την τιμή $v(x)$.

Ειναι εύκολο να δούμε ότι

$$\begin{aligned} v(x) \;&=\; \sup_{y_0} \; \sup_{y_0,y_1,\ldots} \left\{ R_0(x_0,y_0) + \sum_{k=1}^{\infty} R_k(x_k,y_k) \,/\, x_{k+1} = g_k(x_k,y_k)\,,\; y_k \in [0, x_k]\,,\, k \,\geq\, 0\,,\, x_0{=}x \right\} \\ &=\; \sup_{y_0} \left\{ R_0(x_0,y_0) + \sup_{y_1,y_2,\ldots} \{\sum_{k=1}^{\infty} R_k(x_k,y_k) \,/\, x_{k+1} = g_k(x_k,y_k)\,,\; y_k \in [0, x_k]\,,\, k \,\geq\, 1\,\},\, x_0{=}x \right\} \\ &=\; \sup_{y_0} \left\{ R(x_0,y_0) + v(g(x_0,y_0)) \right\} \end{aligned} \tag{5}$$

Οπου στην τελευταία ισότητα έχουμε υποθέσει ότι

$$R_k(\cdot\,,\cdot) \;=\; R(\cdot\,,\cdot)\,,\;\; g_k(\cdot\,,\cdot) \;=\; g(\cdot\,,\cdot) \;\; \forall\, k \,\geq\, 0\,. \tag{6}$$

Παραδοχέs τηs παραπάνω μορφήs ονομάζονται υποθέσειs ***στασιμότηταs*** για τη δομή κέρδουs (ή κόστουs) και τη δυναμική του συστήματοs. Χρήσιμα αποτελέσματα για προβλήματα άπειρου ορίζοντα υπάρχουν μόνο κάτω από τέτοιεs παραδοχέs. Η εξίσωση (5) ισχύει για οποιαδήποτε αρχική κατάσταση $x_0{=}x$ και είναι μία εξίσωση τηs μορφήs

$$v(x) \;=\; \sup_{\alpha} \left\{ R(x_0,\alpha) + v(\overline{\alpha}) \right\}, \tag{7}$$

όπου

$$\overline{\alpha} \;=\; g(x_0,y_0)\,. \tag{8}$$

Η ενδιαφέρουσα τιμή είναι βέβαια η $v(B)$.

Παρατηρήσεις. **1.** *Αν κάνουμε επιπλέον υποθέσεις, τότε μπορούμε να είμαστε σίγουροι ότι η τιμή v(x) είναι τουλάχιστον πεπερασμένη. Για παράδειγμα, ας υποθέσουμε ότι* $\exists\, M \in \mathbb{R}^+ : R(x,y) = R(y) \leq My \quad \forall y$ *και* $g(x,y) = x - y$. *Τότε*

$$v(x) \leq MB < \infty . \tag{9}$$

Πραγματικά : $\forall y_k \geq 0 , \quad \sum_{k=0}^{\infty} y_k \leq B \quad \sum_{k=0}^{\infty} R(y_k) \leq \sum_{k=0}^{\infty} My_k \leq M\sum_{k=0}^{\infty} y_k \leq MB .$

2. Μπορέσαμε να γράψουμε την εξίσωση βελτιστοποίησης (5) ακολουθώντας μια μέθοδο ανάλογη με εκείνη που οδήγησε στην (3) στην περίπτωση του πεπερασμένου ορίζοντα, όπου η κατανομή στη δραστηριότητα 0, y_0, *είναι η απόφαση στο αρχικό στάδιο. Συμβολικά έχουμε αντιστοιχίσει στο πρόβλημα P(B) τα υποπροβλήματα P(x) παρακ`ατω*

$$(max) \qquad w = \sum_{k=1}^{\infty} R(x_k,y_k)$$

$$\pi. \qquad x_{k+1} = g(x_k,y_k) = c\,(x_k - y_k) + d\,y_k ,$$

$$y_k \in [0, x_k] , \; k = 1, 2, \ldots ,$$

Αρχικό σημείο: $x_1 = x$.

3. Η ποσότητες που μας ενδιαφέρουν πρωταρχικά είναι οι βέλτιστες πολιτικές και όχι οι τιμές v(x) . Στο δεδομένο πρόβλημα μία βέλτιστη πολιτική είναι μία ακολουθία $\{y_0,y_1,\ldots\}$ *που να έχει τιμή v(x) . Η ύπαρξη μίας βέλτιστης πολιτικής είναι ισοδύναμη με το γεγονός ότι μπορούμε να αντικαταστέισουμε το sup με max στις σχέσεις (2) και (7) παραπάνω.*

4. Ας υποθέσουμε ότι γνωρίζουμε τη λύση v(x) της (7) $\forall x$ *και μία βέλτιστη πολιτική* $\{y_0,y_1,\ldots\}$. *Τότε η λύση στο αρχικό πρόβλημα δίδεται από την ακολουθία:* $y_0 = y_0(B)$, $y_1 = y_1(g(x_0,y_0))$, $y_2 = y_2(g(x_1,y_1))$, *κλπ.*

Επεκτείνοντας την ορολογία που αναπτύξαμε στο προηγούμενο κεφάλαιο, μία πολιτική π *ορίζει μια μοναδική απόφαση, για κάθε χρονική στιγμή t, σύμφωνα με ένα κανόνα :* $\alpha_t = \pi(h_t)$, *όπου* h_t *είναι η ιστορία του συστήματος μέχρι την χρονική στιγμή t* ($h_t = (x_0,\alpha_0, x_1,\ldots,\alpha_{t-1}, x_t)$). *Η απόδοση της* π , *όταν* $x_0 = x$, *ορίζεται ως εξής*

$$w(\pi, x) = \sum_{t=0}^{\infty} R(x_t,\pi(h_t)) . \tag{10}$$

όπου $x_{t+1} = g(x_t,\pi(h_t))$, $x_0 = x$.

Το πρώτο αποτέλεσμα που θα αποδείξουμε είναι το εξής.

Πρόταση 1. *Εστω* $R(\cdot,\cdot) \geq 0$. *Ε`αν ισχύει η παρακάτω ανισότητα*

$$w(\pi, x) \geq \sup_{\alpha \in [0,x]} \Big\{ R(x,\alpha) + w(\pi,g(x,\alpha)) \Big\} . \tag{11}$$

για κάθε x στο διάστημα [0, B] , όπου π *είναι μια απλή πολιτική, τότε η* π *είναι βέλτιστη. Δηλαδή υπάρχει η τιμή v(x)* $\forall x$ *καί* $w(\pi, x) = v(x)$ *(ιδιαίτερα* $w(\pi, B) = v(B)$*) .*

Απόδειξη. *Εστω* $\pi' = \{\alpha'_t , \; t = 0,1,\ldots\}$ *μια άλλη πολιτική (δηλαδή μια εφικτή κατανομή του πόρου στις δραστηριότητες) και έστω* $\{x'_t , \; t = 0,1,\ldots\}$ *η αντίστοιχη διαδρομή,* $x'_0 = x$, $x'_1 = g(x,\alpha'_0)$, $x'_2 = g(x'_1,\alpha'_1)$, *κλπ. Τότε*

$$w(\pi, x) \geq R(x,\alpha'_0) + w(\pi,x'_1) \qquad (\text{λόγω της } (11))$$

$$\geq \; R(x,\alpha_0') + R(x_1',\alpha_1') + w(\pi,x_2') \qquad\qquad (\pi\acute{\alpha}\lambda\iota \;\lambda\acute{o}\gamma\omega \;\tau\eta s \;\; (11)\,)$$

$$\geq \ldots \geq R(x,\alpha_0') + R(x_1',\alpha_1') + \cdots + R(x_t',\alpha_t') + w(\pi, x_{t+1}')\,. \tag{12}$$

Τώρα η υπόθεση $R(\cdot\,,\cdot) \;\geq\; 0$ *συνεπάγεται ότι*

$$w(\pi, x_{t+1}') \;\geq\; 0\,. \tag{13}$$

Ετσι, οι (12) και (13) συνεπάγονται την παρακάτω ανισότητα

$$w(\pi, x) \;\geq\; \sum_{k=0}^{t} R(x_t',\alpha_t') \qquad \forall\, t \;\geq\; 0\,. \tag{14}$$

Απο την (14) παίρνουμε αμέσως την εξής ανισότητα :

$$w(\pi,\; x) \;\geq\; \sum_{k=0}^{\infty} R(x_t',\alpha_t') \;=\; w(\pi\,', x)\,, \tag{15}$$

άρα η π *είναι βέλτιστη πολιτική.*

1.2 Μελέτη Ειδικών Περιπτώσεων. *Ας υποθέσουμε ότι* $R(x,y) = R(y)$, $R(0) = 0$, $R \geq 0$.

Παράδειγμα 1. *Ας υποθέσουμε επιπλέον ότι* $\exists\, \frac{\partial^2 R}{\partial y^2} \;<\; 0$.

Ας θεωρήσουμε την πολιτική $\pi^0 : \pi^0(h_t) \;=\; x_t$, *δηλαδή η* π^0 *κατανέμει όλη την υπάρχουσα ποσότητα του πόρου αμέσως. Εχουμε λοιπόν,*

$$w(\pi^0, x) \;=\; R(x) \;+\; R(0) \;+\; R(0) \;+ \cdots \;=\; R(x)\,.$$

Θέλουμε να δούμε αν η ανισότητα (8) της πρότασης 1 ισχύει. Εχουμε,

$$w(\pi, x) \;\geq\; R(\alpha) + w(\pi, x-\alpha) \qquad \forall\; \alpha \;\in\; [0, x] \tag{1}$$

Αρκεί, λοιπόν να δείξουμε ότι

$$R(x) \;\geq\; R(\alpha) + R(x-\alpha) \qquad\qquad \forall\; \alpha \;\in\; [0, x]\,. \tag{2}$$

Είναι τώρα εύκολο να δούμε ότι η (13) ισχύει. Πράγματι, έστω

$$h(\alpha) \;=\; R(\alpha) + R(x-\alpha)\,. \tag{3}$$

Εχουμε

$$\frac{\partial^2 h(\alpha)}{\partial\alpha^2} \;=\; \frac{\partial^2 R(\alpha)}{\partial\alpha^2} + \frac{\partial^2 R(x\text{-}\alpha)}{\partial\alpha^2} \;<\; 0\,. \tag{4}$$

Ετσι, η h είναι αυστηρά κυρτή και το μέγιστο σημείο της πάνω στο διάστημα $[0, x]$ *επιτυγχάνεται σε ένα απ`ο τα άκρα του διαστήματος* $\alpha \;=\; 0$ *ή* $\alpha \;=\; x$.
Τώρα $h(0) \;=\; h(x) \;=\; R(x)$, *και έτσι η (13) ισχύει.*

Η βέλτιστη λύση του προβλήματος είναι η εξής.

$$\alpha_t^0 \;=\; \begin{cases} B\,, & \gamma\iota\alpha \quad t = to \\ 0\,, & \gamma\iota\alpha \quad t \neq to \end{cases}$$

όπου to είναι οποιαδήποτε σταθερά τιμή του t, $t_0 \;\geq\; 0$.

Παρατηρούμε, λοιπόν, ότι η βέλτιστη λύση είναι η ίδια με εκείνη που είχαμε βρει για την περίπτωση του πεπερασμένου ορίζοντα όπου

$$\alpha_t^0(N) = \begin{cases} B, & \text{για } t = to \\ 0, & \text{για } t \neq to \end{cases} \qquad to \in \{0,\dots,N-1\}.$$

Παράδειγμα 2 : *Ας υποθέσουμε επιπλέον ότι* $\exists \ \frac{\partial^2 R}{\partial y^2} > 0$.

Ας θυμηθούμε κατ' αρχή ότι για το πρόβλημα του πεπερασμένου ορίζοντα είχαμε βρεί ότι η βέλτιστη πολιτική είχε την μορφή

$$(\alpha_t^0 = \alpha_t^0(N) = \tfrac{B}{N}, \qquad t = 0,1,\dots,N-1. \tag{5}$$

Παρατηρούμε τώρα ότι

$$\alpha_t^0(N) \rightarrow 0 \quad (N \rightarrow \infty). \tag{6}$$

Η (6) υπανίσσεται ότι η βέλτιστη λύση για το πρόβλημα του άπειρου ορίζοντα θα είναι $\alpha_t^* = 0 \ \forall t \geq 1$. *Αλλά τότε :* $w(\pi^*,B) = \sum_{k=1}^{\infty} R_k(0) = 0$, *και είναι εύκολο να δούμε ότι υπάρχουν πολιτικές π ' που έχουν μεγαλύτερη απόδοση.*

Τα παραπάνω οδηγούν στο συμπέρασμα ότι δεν πέπει να υπάρχει λύση για το πρόβλημα του άπειρου ορίζοντα.

Απόδειξη. *Ας υποθέσουμε ότι υπάρχει μια βέλτιστη πολιτική* $\pi^0 = \{\alpha_t^0, t \geq 0\}$. *Τότε πρέπει να υπάρχει τουλάχιστον ένα ζεύγος* $\alpha_k^0, \alpha_{k+1}^0$ *αποφάσεων έτσι ώστε :* $\alpha_k^0 \neq \alpha_{k+1}^0$. *Πραγματικά, η απαίτηση* $\alpha_t^0 = \alpha \ \forall t$, *οδηγεί σε μη εφικτή πολιτική αν* $\alpha > 0$ $(\sum_{k=0}^{\infty} \alpha = \infty \neq B)$, *και όχι βέλτιστη πολιτική αν* $\alpha = 0$.

Τώρα

$$\begin{aligned} w(\pi^0,B) &= \sum_{0}^{k-1} R(\alpha_t^0) + R(\alpha_k) + R(\alpha_{k+1}) + \sum_{k+2}^{\infty} R(\alpha_t) \\ &< \sum_{0}^{k-1} R(\alpha_t^0) + 2R(\tfrac{\alpha_k+\alpha_{k+1}}{2}) + \sum_{k+2}^{\infty} R(\alpha_t), \end{aligned} \tag{7}$$

όπου η τελευταία ανισότητα ισχύει λόγω του γεγονότος ότι η R είναι κοίλη $(\partial^2 R/\partial\alpha^2 > 0)$. *Ετσι, αν ορίσουμε την πολιτική* $\pi' = \{\alpha_t', \ t \geq 0\}$, # *όπου :*

$$\alpha'_t = \begin{cases} \alpha_t, & t \neq k,k+1 \\ \frac{\alpha_k+\alpha_{k+1}}{2}, & t = k,k+1 \end{cases},$$

τότε η (7) είναι ισοδύναμη με :

$$w(\pi^0,B) < w(\pi',B). \tag{8}$$

Δηλαδή η π' *είναι καλύτερη απο την* π^0, *πράγμα που είναι άτοπο δεδομένου ότι δεχθήκαμε ότι η* π^0 *είναι βέλτιστη.*

1.3 Μελέτη της Εξίσωσης Βελτιστοποίησης για το Πρόβλημα Κατανομής Πόρων.

Ας θεωρήσουμε το πρόβλημα της §1.1

$$(max) \qquad w = \sum_{k=0}^{\infty} R(x_k, y_k)$$

$$\pi. \qquad x_{k+1} = g(x_k, y_k) = c\,(x_k - y_k) + d\,y_k\,,$$

$$y_k \in [0, x_k]\,,\ k = 0, 1, \ldots\,,$$

Αρχικό σημείο: $x_0 = B$.

Οι βασικές ερωτήσεις σε σχέση με την εξίσωση βελτιστοποίησης (1.7) είναι ανάλογες με εκείνες που εθέσαμε για το πρόβλημα του πεπερασμένου ορίζοντα. Υπάρχει μοναδική λύση της (1.1.7); Μπορούμε να αντικαταστήσουμε το sup με max; Υπάρχουν μέθοδοι υπολογισμού της λύσης; Παρακάτω δίνουμε απαντήσεις στα ερωτήματα αυτά για το συγκεκριμένο πρόβλημα. Η μεθοδολογία που θα αναπτύξουμε είναι χρήσιμη για το γενικό πρόβλημα.

Θα χρησιμοποίσουμε τον εξής συμβολισμό. Εστω

$$\mathfrak{F}(\mathbb{R}) = \left\{ u \,/\, u\colon \mathbb{R} \to \mathbb{R} : \ \|u\| := \sup_x \{\, |\, u(x)\, |\, \} < \infty \right\} \tag{1}$$

και οι τελεστές

$$T_\alpha : \qquad \mathfrak{F}(\mathbb{R}) \to \mathfrak{F}(\mathbb{R}) \ : \ (T_\alpha u)(x) := R(x,\alpha) + u(\overline{\alpha}) \tag{2}$$

$$T : \qquad \mathfrak{F}(\mathbb{R}) \to \mathfrak{F}(\mathbb{R}) \ : \ (Tu)(x) := \sup_{\alpha \in [0,x]} \{R(x,\alpha) + u(\overline{\alpha})\} = \sup_{\alpha \in [0,x]} \{(T_\alpha u)(x)\} \tag{3}$$

όπου

$$\overline{\alpha} = x_1 = g(x_0, \alpha)\,. \tag{4}$$

Με τη βοήθεια των παραπάνω τελεστών οι σχέσεις (1.11) και (1.7) μπορούν αντίστοιχα να γραφούν ως εξής.

$$w(\pi,x) \geq (T_\alpha w(\pi,\cdot))(x) \ \ \forall\, x \tag{5}$$

$$v(x) = (Tv)(x) = \sup_{\alpha \in [0,x]} \{(T_\alpha v)(x)\} \tag{6}$$

Το κύριο αποτέλεσμα είναι το εξής.

Πρόταση 1. *Ας κάνουμε τις παρακάτω υποθέσεις.*

1. $R(x,y)$ *είναι αύξουσα συνάρτηση ως πρός* x *και* y *και* $R(0,0)=0$

2. *Εστω* $b = max\{c,d\}$, *τότε* $b \in [0,1)$.

3. *Εστω* $m(x) = \max_{y \in [0,x]} \{\, |\, R(x,y)\, |\, \}$ *τότε* $\sum_{t=0}^{\infty} m(b^t x) < \infty$.

Τότε:

1. *Υπάρχει μία μοναδική λύση της εξίσωσης βελτιστοποίησης (6).*

2. *Η λύση της (6) είναι συνεχής συν`αρτηση και* $v(0) = 0$.

3. *Η λύση της (6) είναι η συνάρτηση τιμών του προβλήματος* $P(x)$.

4. *Υπάρχουν βέλτιστες πολιτικές.*

Απόδειξη. *Ας ορίσουμε την εξής ακολουθία των διαδοχικών προσεγγίσεων*

$$\phi^{(0)}(x) = 0 \tag{7}$$

$$\phi^{(n+1)}(x) = (T\phi^{(n)})(x) = \sup_{\alpha \in [0,x]} \{(T_\alpha \phi^{(n)})(x)\}, \; n = 0,1,\ldots . \tag{8}$$

Θα αποδείξουμε ότι οι παραπάνω υποθέσεις συνεπάγωνται ότι οι συναρτήσεις $\phi^{(n)}(\cdot)$ είναι συνεχείς και ότι συγκλίνουν ομοιόμορφα σε μία συνάρτηση $\phi(\cdot)$. Σαν αποτέλεσμα της ομοιόμορφης σύγκλισης έχουμε ότι η συνάρτηση $\phi(x)$ είναι συνεχής. Η συνέχεια τώρα της $\phi(\cdot)$ εύκολα συνεπάγεται τη συνέχεια της $(T_\alpha\phi)(x)$, ως πρός α, γεγονός που αρκεί γιά να είναι δύνατη η αντικατάσταση του sup με max στην (6).

Η απόδειξη της συνέχειας των $\phi^{(n)}(\cdot)$ είναι άμεση συνέπεια των (7), (8) και της υπόθεσης 1. Για να αποδείξουμε την ομοιόμορφη σύγκλιση της ακολουθίας $\phi^{(n)}(\cdot)$ ας ορίσουμε τη συνάρτηση

$$u_n(x) = \max_{\alpha \in [0,x]} \{ \mid \phi^{(n)}(\alpha) - \phi^{(n+1)}(\alpha) \mid \} . \tag{9}$$

Θά αποδείξουμε στο τέλος τον εξής

Ισχυρισμός 1.

$$\sum_{n=0}^{\infty} u^{(n)}(x) \leq \sum_{n=0}^{\infty} m(b^n x) < \infty . \tag{10}$$

Αμεσο αποτέλεσμα της (10) είναι η ομοιόμορφη σύγκλιση της $\sum_{n=0}^{\infty} u^{(n)}(x)$ σε κάθε πεπερασμένο διάστημα. Από γνωστό θεώρημα της Ανάλυσης (βλπ. Rudin (1976) σελ. 147) έχουμε ότι η ακολουθία $\phi^{(n)}(\cdot)$ συγκλίνει ομοιόμορφα σε κάθε πεπερασμένο διάστημα, γεγονός πού με τη σειρά του συνεπάγεται την ύπαρξη και τη συνέχεια της $\phi(\cdot)$ (βλπ. Rudin (1976) σελ. 149).

Επιπλέον είναι εύκολο να αποδείξουμε επαγωγικά (μέσω των (7) καί (8)) ότι $\phi^{(n)}(0) = 0$ πράγμα που συνεπάγεται ότι $\phi(0) = 0$.

Για να αποδείξουμε ότι η (6) έχει μία μόνο λύση ας υποθέσουμε κατ' αρχή το αντίθετο. Εστω λοιπόν, $\phi_1(\cdot)$, $\phi_2(\cdot)$ δύο λύσεις της (6) για τις οποίες, σύμφωνα με τα παραπάνω, δεχόμαστε ότι

$$\phi_1(0) = \phi_2(0) = 0 \tag{11}$$

Ας ορίσουμε τώρα τη συνάρτηση

$$u(x) = \max_{\alpha \in [0,x]} \{ \mid \phi_1(\alpha) - \phi_2(\alpha) \mid \} . \tag{12}$$

Θά αποδείξουμε στο τέλος τον εξής

Ισχυρισμός 2.

$$u(x) \leq u(bx) \leq \cdots \leq u(b^n x), \; n \geq 1 . \tag{13}$$

Τώρα οι (11) και (12) συνεπάγονται ότι $u(0) = 0$ και έτσι, η υπόθεση ότι $b \in [0,1)$ και η (13) συνεπάγονται ότι $u(x) \leq 0 \; \forall x$. Η τελευταία μαζί με τον ορισμό της $u(\cdot)$ συνεπάγονται ότι $u(x) = 0 \; \forall x$ και συνεπώς $\phi_1(x) = \phi_2(x) \; \forall x$.

Για να αποδέιξουμε ότι η μοναδική λύση της (6) είναι η συνάρτηση τιμών όπως ορίστικε με τη σχέση (1.1.2) φτάνει να παρατηρήσουμε ότι από τον ορισμό της $\phi^{(n)}(x)$ έχουμε ότι

$$\phi^{(n)}(x) = v(0,n-1,x) . \tag{14}$$

Τέλος, η ύπαρξη βέλτιστης (ων) πολιτικής(ων) είναι άμεσο αποτέλεσμα της συνέχειας της $(T_\alpha\phi)(x)$, ως πρός α , όπου α ανήκει στο συμπαγές σύνολοο $[0,x]$.

Γιά να είναι πλήρης η απόδειξη πρέπει να αποδείξουμε τους ισχυρισμούς 1. και 2. .

Απόδειξη του Ισχυρισμού 1. Είναι εύκολο να δείξουμε επαγωγικά ότι οι $\phi^{(n)}(x)$ είναι συνεχείς πράγμα που συνεπάγεται (λόγω τηςσυνέχειας της $(T_\alpha\phi^{(n)})(x)$) ότι υπάρχουν αποφάσεις $\alpha_n^0 = \alpha_n^0(x)$ έτσι πού

$$\phi^{(n+1)}(x) = (T\phi^{(n)})(x) = (T_{\alpha_n^0(x)}\phi^{(n)})(x) . \tag{15}$$

Εχουμε λοιπόν

$$\phi^{(n+1)}(x) = (T_{\alpha_n^0(x)}\phi^{(n)})(x) \qquad \geq (T_{\alpha_{n+1}^0(x)}\phi^{(n)})(x) , \tag{16}$$

$$\phi^{(n+2)}(x) = (T_{\alpha_{n+1}^0(x)}\phi^{(n+1)})(x) \qquad \geq (T_{\alpha_n^0(x)}\phi^{(n+1)})(x) . \tag{17}$$

από τις (16) και (17) παίρνουμε

$$\begin{aligned} |\phi^{(n+1)}(x) - \phi^{(n+2)}(x)| &\leq \max\Big\{ |(T_{\alpha_{n+1}^0(x)}\phi^{(n)})(x) - (T_{\alpha_{n+1}^0(x)}\phi^{(n+1)})(x)| , \\ &\qquad\qquad |(T_{\alpha_n^0(x)}\phi^{(n)})(x) - (T_{\alpha_n^0(x)}\phi^{(n+1)})(x)| \Big\} \\ &\leq \max_{\alpha\in[0,x]}\Big\{ |(T_\alpha\phi^{(n)})(x) - (T_\alpha\phi^{(n+1)})(x)| \Big\} \\ &\leq \max_{\alpha\in[0,x]}\Big\{ |\phi^{(n)}(g(x,\alpha)) - \phi^{(n+1)}(g(x,\alpha))| \Big\} \\ &\leq \max_{z\in[0,bx]}\Big\{ |\phi^{(n)}(z) - \phi^{(n+1)}(z)| \Big\} \\ &= u_n(bx) . \end{aligned} \tag{18}$$

Οπου η προτελευταία ανισώτητα παραπάνω είναι αποτέλεσμα της υπόθεσης

$$z = g(x,\alpha) \leq bx \ \forall\, a \in [0,x] , \ \ \forall\, x . \tag{19}$$

Η (18) καί η παρατήρηση ότι οι συναρτήσεις $u_n(x)$ είναι αύξουσες συνεπάγονται ότι

$$u_{n+1}(x) \leq u_n(bx) . \tag{20}$$

Για τη $u_0(x)$ έχουμε

$$u_0(x) \leq \max_{\alpha\in[0,x]}\Big\{ |\phi^{(0)}(\alpha) - \phi^{(1)}(\alpha)| \Big\} = \max_{\alpha\in[0,x]}\Big\{ |R(x,\alpha)| \Big\} = m(x) \tag{21}$$

Το αποτέλεσμα τώρα είναι άμεση συνέπεια των (20) και (21) .

Απόδειξη του Ισχυρισμού 2. Εχουμε δείξει ότι οι λύσεις $\phi_i(\cdot)$ είναι συνεχείς, συνεπώς υπάρχουν βέλτιστες αποφάσεις $\alpha_i = \alpha_i(x)$. Οπως και παραπάνω έχουμε τις σχέσεις

$$\phi_1(x) \;=\; (T\phi_1)(x) \;=\; (T_{\alpha_1}\phi_1)(x) \;\geq\; (T_{\alpha_2}\phi_1)(x)\,, \tag{22}$$

$$\phi_2(x) \;=\; (T\phi_2)(x) \;=\; (T_{\alpha_2}\phi_2)(x) \;\geq\; (T_{\alpha_1}\phi_2)(x)\,. \tag{23}$$

Από τις(22) και (23) ανάλογα με τα παραπάνω παίρνουμε

$$u(x) \;=\; \max_{\alpha \in [0,x]} \{\, |\, \phi_1(\alpha) \,-\, \phi_2(\alpha)\, |\, \} \;\leq\; u(bx)\,. \tag{24}$$

Επαναλαμβάνοντας την (24) παίρνουμε τη σχέση

$$u(x) \;\leq\; u(b^n x) \;\; \forall\, n \;\geq\;\; \forall\, x\,. \tag{25}$$

Από τον ορισμό τη $u(\cdot)$*, την (25) και την υπόθεση ότι* $b \in [0,1)$ *παίρνουμε ότι*

$$0 \;\leq\; u(x) \;\leq\; u(0) \quad \forall\, x\,. \tag{26}$$

Και η απόδειξη είναι πλήρης αν παρατηρήσουμε ότι $u(0) \;=\; 0$*, δεδομένου ότι γνωρίζουμε ότι οι* $\phi_i(x)$ *είναι συνεχείς και* $\phi_i(0) \;=\; 0$*.*

2. Λύση του Γενικού Προβλήματος.

Ας θεωρήσουμε το πρόβλημα της §1 με μέτρο απόδοσης της μορφής (1), όπου $b \in [0,1]$ *είναι μια γνωστή σταθερά,*

$$w(l_\infty) \;=\; \sum_{t=0}^{\infty} b^t c(x_t, \alpha_t)\,, \tag{1}$$

και ας θεωρήσουμε το σύνολο των υποπροβλημάτων $P(x)$*,* $x \in S$*,*

(min) $\qquad w(l_\infty)$

περιορισμοί : $\qquad x_{t+1} \;=\; \alpha_t\,,\;\; t \;\geq\; 0$

$\qquad x_t \;\in\; S\,,\;\; \alpha_t \;\in\; D(x_t)$

$\qquad x_0 \;=\; x\,.$

Θα ορίσουμε τις εξής υποθέσεις.

Υπόθεση 1. *Η δυναμική του συστήματος και η δομή κόστους είναι* ***στάσιμες*** *(μη μεταβαλόμενες στο χρόνο), δηλαδή* $c_t(x_t,\alpha_t) = c(x_t,\alpha_t)$ *και εάν* $x_t \;=\; x_s$*,* $\alpha_t \;=\; \alpha_s$ *τότε* $D(x_t) \;=\; D(x_s)$ *και* $\overline{\alpha}_t \;=\; \overline{\alpha}_s$*.*

Υπόθεση 2. *Τα προβλήματα* $P(x)$ *έχουν τιμή, δηλαδή*

$$\exists\; v(x) \;=\; \inf\{w(l_\infty) \;\; /\; x_{t+1} \;=\; \overline{\alpha}_t\,,\; t \;\geq\; 0\,,\; x_t \;\in\; S\,,\; \alpha_t \;\in\; D(x_t)\,,\; x_0 \;=\; x\} \tag{2}$$

Υπόθεση 3. *Ισχύει η υπόθεση 2 με minimum στη θέση του infimum.*

Το πρώτο αποτέλεσμα που αποδεικνύουμε είναι ότι η συνάρτηση $v(\cdot)$ *ικανοποιεί, και μπορεί να υπολογιστεί από μια εξίσωση του τύπου* $v = Tv$*.*

Πρόταση 1. *1. Κάτω από τις υποθέσεις 1 και 2 η συνάρτηση* $v(\cdot)$ *είναι λύση της παρακάτω εξίσωσης*

$$v(x) \;=\; \inf_{\alpha \in D(x)} \{c(x,\alpha) + bv(\overline{\alpha})\} \;, \quad \forall\, \alpha \;\in\; D(x)\,, \;\; \forall\, x \;\in\; S \tag{3}$$

2. Κάτω από τις υποθέσεις 1 και 3 η συνάρτηση $v(\cdot)$ είναι η μοναδική λύση της παραπάνω εξίσωσης με minimum στη θέση του infimum.

Απόδειξη. Η απόδειξη είναι ανάλογη των επιχειρημάτων που εχρησιμοποιήσαμε στην §1.2. για πληρότητα επαναλαμβάνουμε τα κύρια βήματα. Υπάρχουν λοιπόν οι παρακάτω δύο τρόποι γιά να αποδείξουμε το αποτέλεσμα.

Μέθοδος 1 . Από τον ορισμό της $v(\cdot)$ έχουμε

$$\begin{aligned}
v(x) \;&=\; \inf_{\alpha_0,\alpha_1,\ldots} \left\{ \sum_0^\infty b^t c(x_t,\alpha_t) \;\; / \;\; x_t \;\in\; S\,, \;\; x_{t+1} = \overline{\alpha}_t\,, \;\; \alpha_t \;\in\; D(x_t)\,, \;\; x_0 \;=\; x \right\} \\
&=\; \inf_{\alpha_0} \;\; \inf_{\alpha_1,\alpha_2,\ldots} \left\{ c(x,\alpha_0) + \sum_1^\infty b^t c(x_t,\alpha_t) \;\; / \;\; x_t \in S\,, \; x_{t+1} = \overline{\alpha}_t\,, \; \alpha_t \in D(x_t),\; x_0 = x \right\} \\
&=\; \inf_{\alpha_0 \in D(x)} \Big\{ c(x,\alpha_0) + b \inf_{\alpha_1,\alpha_2,\ldots} \{ c(x,\alpha_1) + \sum_1^\infty b^t c(x_{t+1},\alpha_{t+1}) \;\; / \;\; x_t \;\in\; S\,, \; x_{t+1} = \overline{\alpha}_t\,, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \alpha_t \in D(x_t)\,, \;\; x_0 \;=\; x \,\} \Big\} \\
&=\; \inf_{\alpha_0 \in D(x)} \{ c(x,\alpha_0) + b\, v(\overline{\alpha}_0) \}\,.
\end{aligned}$$

Μέθοδος 2 . Εστω $v(t,x,N)$ η συνάρτηση τιμών του προβλήματος πεπερασμένου ορίζοντα (όπου το μήκος του ορίζοντα είναι N) , δηλαδή

$$v(t,x,N) \;=\; \inf_{\alpha_t,\ldots,\alpha_N} \{ \sum_{k=t}^{N} b^k c(x_k,\alpha_k) \} \tag{4}$$

όπου $x_t \;=\; x$.

Τότε η $v(t,x,N)$ ικανοποιεί την εξίσωση

$$v(t,x,N) \;=\; \inf_{\alpha} \{ c(x,\alpha) + v(t+1,\overline{\alpha},N)\,. \tag{5}$$

Ας υποθέσουμε ότι υπάρχει το όριο : $\lim_{N\to\infty} v(t,x,N)$ και έστω

$$u(t,x) \;=\; \lim_{N\to\infty} v(t,x,N) \tag{6}$$

Από την (19) όμως έχουμε ότι

$$u(t,x) \;=\; \inf_{k \geq t} \left\{ \sum_{k=t}^{\infty} b^k c(x_k,\alpha_k) \right\} \qquad (x_t \;=\; x) \tag{7}$$

και

$$\begin{aligned}
u(t+1,x) \;\;&=\; \inf_{k \geq t+1} \left\{ \sum_{t+1}^{\infty} b^k c(x_k,\alpha_k) \right\} \qquad (x_{t+1} \;=\; x) \\
&=\; bu(t,x)\;\;. \qquad\qquad\qquad\qquad\qquad\qquad\qquad (8)
\end{aligned}$$

Απ`ο την (8) βρίσκουμε ότι

$$u(t,x) \;=\; b^t u(0,x) \;=\; b^t u(x) \tag{9}$$

όπου ορίσαμε την συνάρτηση $u(x) \;=\; u(0,x)$.

Αν τώρα, πάρουμε τα όρια και των δ`υο μελών της (9) , βρίσκουμε ότι

$$u(t,x) \;=\; \inf_{\alpha}\{c(x,\alpha) + u(t+1,\overline{\alpha})\} \,. \tag{10}$$

Τώρα οι (10) , (8) συνεπάγονται (παίρνοντας $\alpha = 0$) ότι

$$u(x) = \inf_{\alpha \in D(x)} \{c(x,\alpha) + bu(\overline{\alpha})\} \,. \tag{11}$$

Η απόδειξη είναι πλήρης όταν παρατηρήσουμε ότι λόγω της (6) ισχύει

$$u(x) \;=\; v(x) \;\; \forall\, x \,\in\, S \,. \tag{12}$$

Θα δούμε παρακάτω συνθήκες που επιτρέπουν την αντικατάσταση του inf (ή sup) με min (ή max) στις εξισώσεις. Οπως και στο πρόβλημα του πεπερασμένου ορίζοντα, στο πρόβλημα του άπειρου ορίζοντα η αντικατάσταση του sup με max , στην βασική εξίσωση, εξασφαλίζει την ύπαρξη βέλτιστης πολιτικής, συγκεκριμένα αποδεικνύουμε την εξής.

Πρόταση 2. Ας υποθέσουμε ότι ισχύουν οι υποθέσεις 1 και 3. Εστω π^0 η απλή πολιτική για την οποία ισχύει η παρακάτω εξίσωση.

$$v(x) \;=\; c(x,\, \pi^0(x)) + bv(\overline{\pi}^0(x)) \,,\;\; \forall\, x \,\in\, S. \tag{13}$$

Τότε η π^0 είναι βέλτιστη.

Απόδειξη. Εχουμε δεχτεί ότι

$$v(x) := \;\; \inf\left\{\sum_{0}^{\infty} c(x_t,\alpha_t)b^t\right\} \;=\; \min\left\{\sum_{0}^{\infty} c(x_t,\alpha_t)b^t\right\} \tag{14}$$

κα`ι

$$v(x) := \;\; \min\left\{c(x,\alpha) + bv(\overline{\alpha})\right\} \;=\; c(x,\, \pi^0(x)) + bv(\overline{\pi}^0(x))\right\} \tag{15}$$

$$v(x) \;>\; c(x,\alpha) + bv(\overline{\alpha}) \,,\;\; \forall\, \alpha \,\in\, D(x) \,,\;\; \alpha \,\neq\, \overline{\pi}^0(x) \,. \tag{16}$$

Ας υποθέσουμε ότι π είναι μια αυθαίρετη πολιτική και ότι $\pi \neq \pi^0$, δηλαδή $\exists\, x_0 \in S$: $\pi(x_0) \,\neq\, \pi^0(x_0)$. Τότε, λόγω της (32) έχουμε ότι :

$$v(x_0) \;>\; c(x_0,\pi(x_0)) + v(\overline{\pi}(x_0)) \,. \tag{17}$$

Απο τον ορισμό της $v(\cdot)$ έχουμε ότι :

$$w(\pi,x) \;=\; \sum_{t=0}^{\infty} c(x,\pi(x))b^t \;\leq\; v(x) \,. \tag{18}$$

Τώρα, οι (17) και (18) οδηγούν στην

$$w(\pi,x_0) \;<\; v(x_0) \tag{19}$$

και η (19) σημαίνει ότι η π δεν είναι βέλτιστη για το σημείο x_0 , δηλαδή έχουμε ότι

$$w(\pi,x) \;\leq\; w(\pi^0,x) \;=\; v(x) \;\; \forall x \in S \,,\;\; \text{και}\;\; w(\pi,x^0) \;<\; w(\pi^0,x^0) \,, \tag{20}$$

άρα η π δεν είναι βέλτιστη.

Παρατήρηση. Οι σχέσεις (20) εκφράζουν το γεγονός ότι όταν υπάρχει μια βέλτιστη πολιτική π^0 για το πρόβλημα του απείρου ορίζοντα, τότε η π^0 είναι **ομοιόμορφα** βέλτιστη, δηλαδή βέλτιστη για κάθε αρχική κατάσταση $x_0 = x$.

Η παρακάτω πρόταση, ανάλογη της πρότασης 1 της §1.1 είναι χρήσιμη στό να αποδείξουμε ότι μια πολιτική π^* για την οποία υποψιαζόμαστε ότι είναι βέλτιστη είναι πράγματι βέλτιστη.

Πρόταση 3. Εστω ότι ισχύουν οι ανισότητες

$$w(\pi^*, x) \geq c(x, \alpha) + bw(\pi^*, \overline{\alpha}), \quad \forall \alpha \in D(x), \ \forall x \in S \tag{21}$$

και επιπλέον,

$$\lim_{t \to \infty} b^t v(\overline{\pi}_t(x)) = 0, \ \forall x \in S, \ \forall \pi, \tag{22}$$

όπου $\pi_1(x) = \pi(x)$, $\pi_{t+1}(x) = \pi(\overline{\pi}_t(x))$. Τότε η πολιτική π^* είναι βέλτιστη.

Απόδειξη. Αρκεί να αποδείξουμε ότι

$$w(\pi^*, x) \geq w(\pi, x) \quad \forall x, \ \pi \neq \pi^*, \tag{23}$$

όπου π είναι μια αυθαίρετη πολιτική, $\pi = (\alpha_0, \alpha_1, \alpha_2, \ldots)$. Η (23) οδηγεί στην

$$\begin{aligned} w(\pi^*, x) &\geq c(x, \alpha_0) + bw(\pi^*, \overline{\alpha}_0) \\ &\geq c(x, \alpha_0) + bc((\overline{\alpha}_0, \alpha_1) + bw(\pi^0, \overline{\alpha}_1)) \\ &\geq \sum_{t=0}^{\infty} c(x, \pi_t(x)) b^t + \lim_{t \to \infty} b^t w(\pi^0, \overline{\pi}_t(x)) \\ &= \sum_{t=0}^{\infty} c(x, \pi_t(x)) = w(\pi, x). \end{aligned}$$

Τελειώνουμε την παράγραφο με την εξής γενίκευση των αποτελεσμάτων των προτάσεων 1,2 και 3.

Θεώρημα 1. (1) Ας δεχθούμε ότι ισχύουν οι η υποθέσεις 1 και 2. Ας δεχθούμε επίσης ότι υπάρχει μία συνάρτηση $u(\cdot)$ για την οποία ισχύει η εξής σχέση

$$\lim_{t \to \infty} b^t u(\overline{\pi}_t(x)) = 0, \ \forall x \in S, \ \forall \pi, \tag{24}$$

όπου $\pi_1(x) = \pi(x)$, $\pi_{t+1}(x) = \pi(\overline{\pi}_t(x))$ και επιπλέον, η $u(\cdot)$ είναι λύση της παρακάτω εξίσωσης

$$u(x) = \inf_{\alpha \in D(x)} \{c(x,\alpha) + bu(\overline{\alpha})\}, \quad \forall x \in S, \tag{25}$$

τότε

$$u(x) \leq v(x), \quad \forall x \in S. \tag{26}$$

(2) Ας δεχθούμε ότι $b < 1$. Εάν για την απλή πολιτική π^0 ισχύουν οι ανισότητες

$$u(x) \geq c(x, \pi^0(x)) + bu(\overline{\pi}^0(x)) - \epsilon\,, \ \forall x \in S. \tag{27}$$

για κάποια σταθερά $\epsilon > 0$, τότε

$$u(x) \geq w(\pi^0, x) - \tfrac{\epsilon}{1-\beta}\,, \ \forall x \in S \tag{28}$$

οπότε η πολιτική π^0 είναι $\frac{\epsilon}{1-\beta}$ βέλτιστη (δηλαδή $w(x,\pi^0) \in (v(x), v(x)+\frac{\epsilon}{1-\beta})$).

(3) Εαν οι (25) παραπάνω ισχύουν με min στην θέση των inf και εάν στην κατάσταση x τα ελάχιστα επιτυγχάνεται για την απ`οφαση $\alpha(x) = \pi^0(x)$, τότε η πολιτική π^0 είναι βέλτιστη.

Απόδειξη. Εστω π μια αυθαίρετη πολιτική, και έστω $x_0 = x, \alpha_0\, x_1\, \alpha_1\, x_2\, \alpha_2 \ldots$ η τροχιά του συστήματος που αντιστοιχεί στην π , όπου $x_{t+1} = \overline{\alpha}_t = \pi(h_t)$.
Εχουμε λοιπόν

$$w(x,\pi) = \sum_{t=0}^{\infty} b^t c(x_t, \alpha_t)\,. \tag{29}$$

Αρκεί να αποδείξουμε ότι

$$u(x) \leq w(x,\pi^0)\,, \tag{30}$$

διότι τότε, δεδομένου ότι η π^0 είναι αυθαίρετη, η (30) συνεπάγεται την (26) .
Εχουμε λοιπόν

$$\begin{aligned} u(x) &\overset{(25)}{\geq} c(x, \alpha_0) + bu(\overline{\alpha}_0) \\ &\overset{(25)}{\geq} c(x, \alpha_0) + b\Big(c(\overline{\alpha}_0, a_1) + bu(\overline{\alpha}_1)\Big) \\ &\overset{(25)}{\geq} \cdots \geq \sum_{t=0}^{\infty} b^t c(x_t, \alpha_t) + \lim_{t\to\infty} b^t u(\overline{\pi}_t(x)) = w(x,\pi^0)\,. \end{aligned}$$

Για το 2. έχουμε

$$\begin{aligned} u(x) &\overset{(27)}{\geq} c(x, \alpha_0) + bu(\overline{\alpha}_0) - \epsilon \\ &\overset{(27)}{\geq} c(x, \alpha_0) + b\Big(c(\overline{\alpha}_0, a_1) + bu(\overline{\alpha}_1) - \epsilon\Big) - \epsilon \\ &\overset{(27)}{\geq} \cdots \geq \sum_{t=0}^{\infty} b^t c(x_t, \alpha_t) + \lim_{t\to\infty} b^t u(\overline{\pi}_t(x)) - \sum_{t=0}^{\infty} b^t \epsilon \\ &= w(x,\pi^0) - \tfrac{\epsilon}{1-\beta}\,. \end{aligned}$$

Η απόδειξη του 3. είναι ανάλογη.

3. Κύρια Προβλήματα.

Σε εφαρμογές είναι εύκολο να δούμε εάν ισχύει η υπόθεση 1 . Είναι επιθυμητό να έχουμε απλούς ελέγχους για τις υποθέσεις 2 ή 3 . Επιπλέον όπως φαίνεται από τα αποτελέσματα της προηγούμενης παραγράφου υπάρχουν απλές βέλτιστες πολιτικές όταν το infimum (αντίστοιχα, το supremum για προβλήματα μεγιστοποίησης) μπορεί να αντικατασταθεί με minimum (αντίστοιχα, maximum) στις εξισώσεις βελτιστοποίησης. Τυπικές υποθέσεις κάτω από τις οποίες είναι εύκολο να ελέγξουμε την ισχή της υπόθεσης 3 είναι οι παρακάτω δύο :

1. οι χώροι καταστάσεων καί αποφάσεων είναι πεπερασμένα σύνολα,

2. οι χώροι καταστάσεων καί αποφάσεων είναι συμπαγή σύνολα και οι συναρτήσεις κοστους είναι συνεχείς.

Οι παραπάνω υποθέσεις δεν είναι ικανές για την ύπαρξη βέλτιστων λύσεων. Παρακάτω παραθέτουμε σύνολα επιπλέον συνθηκών οι οποίες σε συνδιασμό με μία από τις παραπάνω υποθέσεις κάνουν βέβαιη τη ύπαρξη βέλτιστων λύσεων, βλπ. Ross(1982), Dynkin and Yushkevich(1979) γιά πιό εκτεταμένη παρουσίαση. Τονίζουμε ότι οι παρακάτω συνθήκες δεν είναι οι πι`ο γενικές σε κάθε περίπτωση. Τα σύνολα αυτών των συνθηκών έχουν μελετηθεί στη βιβλιογραφία και έχουν επικρατήσει οι παρακάτω ονομασίες .

3.1 Το Πρόβλημα του Εκπτωτικού Κόστους (Discounted Cost Problem).

(max ή min) $w = \sum_{t=0}^{\infty} b^t c(x_t, a_t)$ κάτω από τις παραδοχές

1. $|c(x,\alpha)| \le C \ \forall x \in S, \ \forall \alpha \in D(x)$,

2. $b < 1$.

3.2. Το Πρόβλημα τής Πρώτης Μετάβασης (First Passage Problem).

(max ή min) $w = \sum_{t=0}^{\infty} b^t c(x_t, a_t)$ κάτω από τις παραδοχές

1. $\exists x_0, \ \alpha_0 \in D(x_0) : c(x_0, \alpha_0) = 0$,

2. $\forall x \in S \ \exists$ μια διαδρομή πεπερασμένου κόστους (μήκους) που οδηγεί στο σημείο x_0. Δηλαδή $\exists \{\alpha_t, \ t = 0,1,\ldots, N(x)\}$, έτσι ώστε $\alpha_0 \in D(x)$, $\alpha_t \in D(\overline{\alpha}_{t-1})$, $\overline{\alpha}_{N(x)} = x_0$ και $\sum_{t=0}^{N_\pi(x)} b^t c(x_t, \alpha_t)$ πεπερασμένο, όπου $N_\pi(x)$ είναι ο χρόνος της πρώτης μετάβασης στο x_0 κάτω από την πολιτική π .

Παρατηρούμε ότι η συνθήκη 2., παραπάνω, περιέχει ένα περιορισμό πάνω την δυναμική κίνηση του συστήματος. Το πρόβλημα είναι καλά ορίσμένο και στή γενικότερη περίπτωση όπου η εξέλιξη του συστήματος σταματά όταν το σύστημα "φτάσει" σε οποιαδήποτε κατάσταση ενός υποσυνόλου του χώρου των καταστάσεων.

3.3. Το Πρ`οβλημα τής Βέλτιστης Διακοπής (Optimal Stopping Problem).

(min) $w = \sum_{t=0}^{\infty} b^t c(x_t, a_t)$ κάτω από τις παραδοχές

1. $D(x) = \{\alpha_0, \alpha_1\} \ \forall x$. Η απόφαση α_0 έχει την ερμηνεία ότι διακόπτουμε τη διαδικασία που περιγράφεται από το σύστημα ενώ η απόφαση $\alpha_1 = 0$, έχει την ερμηνεία ότι επιτρέπουμε στη διαδικασία να συνεχίσει την εξέλιξή της γιά μία ακόμη χρονική περίοδο.

2. $c(x,\alpha_0) = r(x) \ge 0$, $c(x,\alpha_1) = c(x) \le 0$, όπου

$r(x) \le R < \infty$ καί , $c(x) \ge C > -\infty$.

3.4. Η Περίπτωση του Θετικού Προγραμματισμού (Positive Programming).

$(min)\ \ w\ =\ \sum_{t=0}^{\infty} c(x_t,a_t)$ κάτω από τή παραδοχή

$c(x,\alpha) \geq 0\ \ \forall\, x\ \ \forall\, \alpha \in D(x)$.

Η περίπτωση της ελαχιστοποίησης της απόδοσης

1. $c(x,\alpha) \geq\ 0\,, \forall\, x\ \in\ S\,,\ \forall\, \alpha\ \in\ D(x)$.

3.5. Η Περίπτωση του Αρνητικού Προγραμματισμού (Negative Programming).

$(max)\ \ w\ =\ \sum_{t=0}^{\infty} c(x_t,a_t)$ κάτω από τή παραδοχή

$c(x,\alpha) \leq 0\ \ \forall\, x\ \ \forall\, \alpha \in D(x)$.

3.6. Η Περίπτωση του Κριτηρίου του Μέσου Κόστους. Ενα κριτήριο που συχνά εμφανίζεται σε εφαρμογές είναι εκείνο του μέσου κόστους, το οποίο δεδομένης μιάς αρχικής κατάστασης x_0 και μιας πολιτικής π , ορίζεται ως εξής

$$\phi(x,\pi) = \lim_{n\to\infty} w_n(x,\pi)/n\ ,$$

όπου

$$w_n(x,\pi)\ =\ \sum_{t=0}^{n-1} c(x_t,\alpha_t)$$

Παρακάτω στη §6 θα εξετάσουμε με λεπτομέρεια ερωτήσεις που έχουν σχέση με τον προσδιορισμό βέλτιστων πολιτικών ως προς αυτό το κριτήριο.

4. Μέθοδοι Λύσης της τής Εξίσωσης Βελτιστοποίησης.

Σε τούτη τη παράγραφο δεχόμαστε ότι ισχύουν οι υποθέσεις 1 και 3 της §2. Οι μέθοδοι πού παρουσιάζουμε εδώ έχουν σχέση με τα προβλήματα των §3.1 - §3.5 όπου το μέτρο απόδοσης (κριτήριο) είναι της μορφής$(min)\ \ w\ =\ \sum_{t=0}^{\infty} b^t c(x_t,a_t)\ ,\ \ b \leq 1$.
Ας θεωρήσουμε την εξίσωση βελτιστοποίησης

$$v(x)\ \ =\ (Tv)(x) \tag{1}$$

$$=\min_{\alpha\in D(x)}\Big\{(T_\alpha v)(x)\Big\}\ , \tag{2}$$

όπου ανάλογα με τις σχέσις (2), (3) της §1.3 ορίζουμε τους τελεστές T_α , T ως εξής

$$T_\alpha :\ \ \mathfrak{F}(\mathbb{R})\ \to\ \mathfrak{F}(\mathbb{R})\ :\ (T_\alpha u)(x) := c(x,\alpha) + bu(\overline{\alpha})\ , \tag{3}$$

$$T :\ \ \mathfrak{F}(\mathbb{R})\ \to\ \mathfrak{F}(\mathbb{R})\ :\ (Tu)(x) := \min_{\alpha\in D(x)}\{c(x,\alpha) + bu(\overline{\alpha})\} = \sup_{\alpha\in D(x)}\{(T_\alpha u)(x)\}\ , \tag{4}$$

4.1. Μέθοδος των Διαδοχικών Προσεγγίσεων στο Χώρο Τιμών. Η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο τιμών αποτελε`ιται από τα εξής βήματα

i) Εστω u^0 μ`ια αρχική προσέγγιση για την v .
ii) Υπολογίζουμε την ακολουθία των συναρτήσεων $u^{(n)}$ και αποφάσεων $\alpha^{(n)}$ ως εξής

$$u^{(n)}(x) = Tu^{(n-1)}(x) = \min_{\alpha \in D(x)} \{c(x,\alpha) + bu^{(n-1)}(\overline{\alpha})\}, \qquad (1)$$

$\alpha^{(n)} = \alpha^{(n)}(x)$ *είναι η απόφαση για την οποία επιτυγχάνεται το minimum στην (3). Συμβολικά γράφουμε*

$$\alpha^{(n)}(x) = \arg\left\{\min_{\alpha \in D(x)} \{c(x,\alpha) + bu^{(n-1)}(\overline{\alpha})\}\right\}, \qquad (2)$$

Τότε, κάτω απο κατάλληλες συνθήκες, βλπ. Θεώρημα 1 παρακάτω, έχουμε ότι

$$u^{(n)}(x) \to v(x), \quad \forall x \quad (n \to \infty) \qquad (3)$$

$$\alpha^{(n)}(x) \to \pi^*(x), \ \forall x \quad (n \to \infty) \qquad (4)$$

όπου π^* *είναι μία βέλτιστη πολιτική.*

Θεώρημα 1 (Picard Fixed Point Theorem). *Εστω* $(\mathfrak{F},\rho)$ *ένας πλήρης μετρικός χώρος καί* $T : \mathfrak{F} \to \mathfrak{F}$ *ένας τελεστής που έχει τη παρακάτω ιδιότητα:*

$$\rho(Tu_1,Tu_2) \le \rho(u_1,u_2) \ \forall\, u_1, u_2 \in \mathfrak{F}. \qquad (5)$$

Τότε

(1) *η εξίσωση*

$$u = Tu \qquad (6)$$

έχει μια μόνο λύση u^*, *καί*

(2) *η ακολουθία* $\{u_n\}$ *που ορίζεται ως εξής* $u_n = Tu_{n-1} = T^n u_0$ *συκλίνει στη μοναδική λύση* u^*, $\forall\, u_0 \in \mathfrak{F}$.

Γιά εκτενή συζήτηση των εφαρμογών του παραπάνω θεωρήματος γιά αποδείξεις ισχυρισμών της μορφής (3), (4) βλπ. Dernado (1967).

4.2. Μέθοδος των Διαδοχικών Προσεγγίσεων στο Χώρο των Πολιτικών. *Ας θεωρήσουμε μία απλή πολιτική* π^0, *καί την απόδοση αυτής* $w(\pi^0,x)$, *όπου* x *συμβολίζει την αρχική κατάσταση της διαδικασίας. Η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο των πολιτικ`ων συνίσταται στη δημιουργία μίας ακολουθίας πολιτικών* π^n, *που συκλίνει στη βέλτιστη πολιτική, ακολουθώντας τα εξής βήματα*

(1) Προσδιορισμός τιμών (value determination): *υπολογισμός των αποδ`οσεων* $w(\pi^n,x)\ \forall x \in S$, *σαν τη λύση του εξής συστήματος εξισώσεων*

$$w(\pi^n,x) = c(x,\pi^n(x)) + bw(\pi^n,\overline{\pi}^n(x)), \ \forall\, x \in S, \qquad (1)$$

(2) Προσδιορισμός καλύτερης πολιτικής (policy improvement): *ορίζουμε τη πολιτική* π^{n+1}, *ως εξής*

$$\pi^{n+1}(x) = \arg\left\{\min_{\alpha \in D(x)} \{c(x,\alpha) + bw(\pi^0,\overline{\pi}^n(x))\right\}, \ \forall\, x \in S, \qquad (2)$$

Συγκεκριμένα έχουμε το εξής αποτέλεσμα

Πρόταση 1. *Η ακολουθία των πολιτικών που ορίζεται από τις (1), (2) παραπάνω έχει τις εξής ιδιότητες:*

(1) $$w(\pi^{n+1},x) \le w(\pi^n,x)\,,\ \forall x \in S\,, \tag{3}$$

(2) *αν*

$$w(\pi^{n+1},x) = w(\pi^n,x)\,,\ \forall x \in S\,, \tag{4}$$

τότε $w(\pi^{n+1},x) = w(\pi^n,x) = v(x),\ \forall x \in S$, *δηλαδή οι πολιτικές* π^{n+1}, π^n *είναι βέλτιστες.*

Απόδειξη. *(1): Αρκεί να αποδείξουμε το αποτέλεσμα γιά* $n = 0$. *Από τις (1), (2) έχουμε*

$$\begin{aligned} c(x,\pi^1(x)) + bw(\pi^0,\overline{\pi}^1(x)) &= \min_{\alpha \in D(x)} \{c(x,\alpha) + bw(\pi^0,\overline{\pi}^0(x))\} \\ &\le w(\pi^0,\overline{\pi}^0(x))\,,\ \forall x \in S\,, \end{aligned} \tag{5}$$

Επαναλαμβάνοντας την (5) παίρνουμε

$$\begin{aligned} &c(x,\pi^1(x)) + bc(\overline{\pi}^1(x),\pi^1(\overline{\pi}^1(x))) + b^2 w(\pi^0,\overline{\pi}^1(\overline{\pi}^1(x))) \\ &\le c(x,\pi^1(x)) + bw(\pi^0,\overline{\pi}^1(x)) \le w(\pi^0,\overline{\pi}^0(x))\,,\ \forall x \in S\,, \end{aligned} \tag{6}$$

και επαγωγικά παίρνουμε τη σχέση

$$\sum_{t=0}^{k} b^t c(x_t,\pi_t^1(x_t)) + b^{k+1} w(\pi^0,\overline{\pi}_k^1(x_k)) \le w(\pi^0,x)\,,\ \forall x \in S\,, \tag{7}$$

όπου έχουμε ορίσει $\pi_t^1(x_t) = \pi^1(\pi_{t-1}^1(x_{t-1}))$. *Από την (7) παίρνουμε αμέσως το αποτέλεσμα (1) όταν*

$$b^{k+1} w(\pi^0,\overline{\pi}_k^1(x_k)) \to 0\,,\ (k \to \infty)\,,\ \forall x \in S\,, \tag{8}$$

δεδομένου ότι

$$\sum_{t=0}^{k} b^t c(x_t,\pi_t^1(x_t)) \to w(\pi^1,x)\,,\ (k \to \infty)\,,\ \forall x \in S\,, \tag{9}$$

(2): Αρκεί να παρατηρήσουμε ότι όταν $w(\pi^1,x) = w(\pi^0,x)$, *τότε οι συναρτήσεις απόδοσης ικανοποιούν την εξίσωση βελτιστοποίησης, που υποθέτουμε ότι έχει μοναδική λύση.*

4.3. Μέθοδος του Γραμμικού Προγραμματισμού. *Ας υποθέσουμε ότι οι χώροι καταστάσεων και αποφάσεων είναι πεπερασμένοι και ας θεωρήσουμε το εξής γραμμικό πρόγραμμα.*

(max) $$\sum_{x \in S} \lambda(x) u(x)$$

π. $$c(x,\alpha) + u(\overline{\alpha}) \ge u(x)\,,\ \forall x \in S\,, \forall \alpha \in D(x)\,, \tag{1}$$

$u(x)$ *μη περιορισμένα στο πρόσημο.*

Εστω $u^*(x)$, $x \in S$ *η βέλτιστη λύση του παραπάνω προγράμματος. Τότε*

$$u^*(x) = v(x)\,,\ x \in S\,. \tag{2}$$

Η απόδειξη του παραπάνω αποτελέσματος βασίζεται στην εξής πρόταση.

Πρόταση 1. Η συνάρτηση τιμών είναι η μεγαλύτερη λύση των περιορισμών (1).

Απόδειξη. Εστω $u(x)$, $x \in S$ μία αυθαίρετη λύση των περιορισμών. Πρέπει να δείξουμε ότι

$$u(x) \;\leq\; v(x)\,,\;\; \forall\, x \in S\,. \tag{3}$$

Πραγματικά η υπόθεση ότι η $u(x)$, $x \in S$ είναι λύση των περιορισμών (2) είναι ισοδύναμη με τη παρακάτω σχέση

$$u(x) \;\leq\; (Tu)(x)\,,\;\; \forall\, x \in S\,. \tag{4}$$

όπου T είναι ο γνωστός τελεστής. Αρκεί τώρα να παρατηρήσουμε ότι η συνάρτηση τιμών $v(x)$ είναι λύση των περιορισμών (2) καί

$$(T^n u)(x) \;\rightarrow\; v(x)\,,\;\; \forall\, x \in S\,. \tag{5}$$

5. Εφαρμογές..

5.1. Ενα Πρόβλημα Πρώτης Μετάβασης. Ας θεωρήσουμε το εξής πρόβλημα

$$(min) \qquad w \;=\; \sum_{1}^{\infty} (x^2(t) + \alpha^2(t)) \tag{1}$$

$$\pi. \qquad x(t+1) \;=\; x(t) + \alpha(t) \tag{2}$$

$$S \;=\; [0,5]\,,\;\; D(x) \;=\; [-5,0]\,, \tag{3}$$
$$x_0 = \;5 \;\in\; S\,. \tag{4}$$

Παρατηρούμε ότι ισχύουν οι αναγκαίες συνθήκες του προβλήματος πρώτης μετάβασης, εάν πάρουμε σαν το ζεύγος (x_0,α_0) του ορισμού το σημείο $(0,0)$.

Ας πάρουμε σαν αρχική προσέγγιση την συνάρτηση $u^0(x) \;=\; 0$. Τότε

$$u^{(1)}(x) \;=\; \min_{\alpha \in D(x)} \{x^2 + \alpha^2 + u^{(o)}(x+\alpha)\} \;=\; \min_{\alpha \in D(x)} \{x^2 + \alpha^2\} \;=\; x^2 \tag{5}$$

και

$$\alpha^{(1)} \;=\; \alpha^{(1)}(x) \;=\; 0\,,\;\; \forall x\,. \tag{6}$$

$$u^{(2)}(x) \;=\; \min_{\alpha \in D(x)} \{x^2 + \alpha^2 + u^{(1)}(x+\alpha)\} \;=\; \min_{\alpha \in [-5,0]} \{x^2 + \alpha^2 + (x+\alpha)^2\}\,. \tag{7}$$

Εστω $\quad h(\alpha) \;=\; x^2 + \alpha^2 + (x+\alpha)^2 \qquad (8)$

$$h'(\alpha) \;=\; 2\alpha + 2(x+\alpha) \tag{9}$$

$$h''(\alpha) \;=\; 4 \;>\; 0 \tag{10}$$

(9), (10) $\Rightarrow$ $\alpha^0 = \; -x/2 \;\in\; D(x)$ ελαχιστοποιεί την h, άρα

$$u^{(2)}(x) \;=\; x^2 + \left(\tfrac{x}{2}\right)^2 + \left(\tfrac{x}{2}\right)^2 + \tfrac{3}{2}x^2\,, \tag{11}$$

και

$$\alpha^{(2)} \;=\; \alpha^{(2)}(x) \;=\; -x/2\,. \tag{12}$$

$$u^{(3)}(x) \ = \min_{\alpha \in D(x)} \{x^2 + \alpha^2 + \tfrac{3}{2}\,(x+\alpha)^2\}\ . \tag{13}$$

$$h(\alpha) \ = \ x^2 + \alpha^2 + \tfrac{3}{2}\,(x+\alpha)^2 \tag{14}$$

$$h'(\alpha) \ = \ 2\alpha + 3(x+\alpha) \tag{15}$$

$$h''(\alpha) \ = \ 2+3 \ = \ 5 \ > \ 0 \tag{16}$$

(15), (16) $\Rightarrow$ $\alpha^0 = -3x/5$ ελαχιστοποιεί την h. Αρα

$$u^3(x) \ = \ x^2 + (\tfrac{3}{5}x)^2 + \tfrac{3}{2}(x - \tfrac{3}{5}x)^2 \ = \ \tfrac{8}{5}x^2 \ = \ 1.6x^2\ , \tag{17}$$

και

$$\alpha^3(x) \ = \ -\tfrac{3}{5}x\ . \tag{18}$$

Ας υποθέσουμε ότι

$$u^{(n)}(x) \ = \ (1+c_n)x^2 \tag{19}$$

και

$$\alpha^{(n)}(x) \ = \ -c_n x^2 \tag{20}$$

Τότε

$$u^{(n+1)}(x) \ = \min_{\alpha \in [-5,0]} \{x^2 + \alpha^2 + c_n(x+\alpha)^2\} \ = \ (1 + \tfrac{c_n}{1+c_n})x^2 \tag{21}$$

και

$$\alpha^{(n+1)}(x) \ = \ -\ \tfrac{c_n}{1+c_n}x\ . \tag{22}$$

Εχουμε λοιπόν αποδείξει, επαγωγικά την (45) και επιπλέον παρατηρούμε ότι η ακολουθία των $\{c_n\}$ ορίζεται επαναληπτικά ως εξής

$$c_1 \ = \ 0\ ,\ c_2 \ = \ 1\ ,\ c_{n+1} \ = \ 1 + \tfrac{c_n}{1+c_n}\ . \tag{23}$$

Παρατηρούμε λοιπόν ότι

i) $\{c_n\} \uparrow_{(n)}$

ii) $1 \ \leq \ c_n \ \leq \ z$

Αρα $\exists \ c \ = \lim_{n \to \infty} c_n$.

Τώρα απο την (47) παίρνουμε ότι :

$$c \ = \ 1 + \tfrac{c}{1+c} \ \Leftrightarrow \ c^2 - c - 1 \ = \ 0 \ \Leftrightarrow \ c \ = \ \tfrac{\alpha 1 \pm \sqrt{1+4}}{2} \ \Leftrightarrow$$

τα πιθανά όρια της $\{c_n\}$ είναι

$$c^{(1)} \ = \ \tfrac{1-\sqrt{5}}{2}\ ,\ \text{και}\ \ c^{(2)} \ = \ \tfrac{1+\sqrt{5}}{2}\ .$$

Τώρα, η $c^{(1)}$ απορίπτεται δεδομένου ότι $c_n \ \geq \ 1$, και $c^{(1)} \ < \ 0$. Αρα $c \ = \ c^{(2)}$. Εχουμε αποδείξει λοιπόν ότι

$$u^{(n)}(x) \ = \ c_n x^2 \ \to \ \tfrac{1+\sqrt{5}}{2}x^2 \ = \ 1.618x^2 \ = \ v(x)\ , \tag{24}$$

και

$$\alpha^{(n)}(x) \;=\; -\;\frac{c_n}{1+c_n}x \;\to\; \frac{-c}{1+c}\,x \;=\; -\;\frac{1+\sqrt{5}}{3+\sqrt{5}}x \;=\; \alpha^*(x)\,. \tag{25}$$

Παρατηρήσεις. *i) Μια ερμηνεία για το παραπάνω πρόβλημα είναι η εξής. Εχομε ένα σύστημα το οποίο θέλουμε να οδηγήσουμε απο μια αρχική κατάσταση $x \in [0,5]$ στην κατάσταση $\{0\}$. Πρέπει να το οδηγήσουμε "γρήγορα" στην κατάσταση 0, παράγοντας x^2 του κόστους, αλλά "τα μεγάλα βήματα κοστίζουν", παράγοντας α^2 του κόστους.*

ii) Ας υποθέσουμε ότι $x_0=5$. Τότε η βέλτιστη διαδρομή έχει την εξής μορφή

$$x_1 \;=\; 5\,,\;\; \alpha_1 \;=\; -\,3.09\,,\;\; x_2 \;=\; 1.90\,,\;\; \alpha_2 \;=\; -\,1.18\,,$$

$$x_3 \;=\; 0.72\,,\;\; \alpha_3 \;=\; -\,0.44 \;\; \text{κ.λ.π.}$$

6 Το Κριτήριο του Μέσου Κόστους.

6.1 Εισαγωγικά αποτελέσματα. *Σε τούτη την παράγραφο εξετάζουμε το πρόβλημα εύρεσης βέλτιστης πολιτικής όταν το κριτήριο που μας ενδιαφέρει είναι εκείνο του μέσου κόστους, το οποίο δεδομένης μιας αρχικής κατάστασης $x_0 = x$ και μιας πολιτικής π ορίζεται ως εξής*

$$\phi(x,\pi) = \lim_{n\to\infty} w_n(x,\pi)/n\,, \tag{1}$$

όπου

$$w_n(x,\pi) \;=\; \textstyle\sum_0^{n-1} c(x_t,\alpha_t)$$

και η τροχιά του συστήματος (x_t,α_t), $t \;=\; 0,1,2,\ldots, N-1$ καθορίζεται απο την πολιτική π (δηλαδή $\alpha_t \;=\; \pi(h_t)$, $h_{t-1} \;=\; x_0,\ldots, x_{t-1},\alpha_{t-1}, x_t$, $x_{t+1} \;=\; \overline{\alpha}_t)$.

Εάν δεν είμαστε σίγουροι (λόγω επιπλέον ιδιοτήτων ενόσπροβλήματος) για την ύπαρξη του ορίου στην (1), τότε ορίζουμε το μέσο κόστος σαν το ανώτερο όριο της ακολουθίας $\{w_n(x,\pi)/n\}$.

Μια πολιτική π_0 είναι βέλτιστη ως προς το κριτήριο του μέσου κόστους αν

$$\phi(x,\pi_0) \;=\; \min_{\pi}\,\{\phi(x,\pi)\}\;\;\forall\, x \in S\,. \tag{3}$$

Θα δείξουμε παρακάτω ότι βέλτιστες πολιτικές, (όταν υπάρχουν) είναι δυνατό να βρεθούν λύνοντας μια συναρτησιακή εξίσωση, ανάλογα με ότι εκάναμε για τη λύση του προβλήματος όταν το κριτήριο απόδοσης ήταν το ολικό εκπτωτικό κόστος Για να κατανοήσουμε όμως καλύτερα την διαδικασία, θα εξετάσουμε πρώτα την περίπτωση πεπερασμένων χώρων καταστάσεων και αποφάσεων και θα εξάγουμε την εξίσωση βελτιστοποίησης περνώντας από το κριτήριο του εκπτωτικού κόστους σε αυτό του μέσου κόστους παίρνοντας το όριο, όταν ο εκπτωτικός παράγοντας β τείνει στο 1.

Στην διαδικασία αυτή θα μας χρειαστούν τα δυο παρακάτω αποτελέσματα.

Λήμμα 1. *Εστω $\{\alpha_n\}$ μια ακολουθία πραγματικών αριθμών, και ας ορίσουμε την συνάρτηση*

$$f(s) \;=\; \sum_0^{\infty} \alpha_n s^n \tag{4}$$

Αν

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \alpha_k = \alpha < \infty , \qquad (5)$$

τότε

$$\lim_{s \to 1} (1-s) f(s) = \alpha . \qquad (6)$$

Λήμμα 2. *Εστω $P = (P(y/x))$ ένας nxn πίνακας με στοιχεία $P(y/x) \in [0,1]$ και έστω $b \in [0,1]$. Τότε*

i) $\exists (I - bP)^{-1}$ *για* $b < 1$.

ii)
$$(I - bP)^{-1} = \frac{S}{1-b} + \Delta + o(1) \quad (b \to 1) , \qquad (7)$$

όπου $o(1) \to 0$ όταν $b \to 1$, S,Δ είναι nxn πίνακες, και όλες οι γραμμές του S είναι οι ίδιες.

iii) *Οι πίνακες S, Δ ικανοποιούν τις παρακάτω σχέσεις:*

$$S = PS = SP = S^2 \qquad (8)$$

$$\Delta = (I - P + S)^{-1} - S \qquad (9)$$

$$S\Delta = 0 \qquad (10)$$

Απόδειξη. *i) Για πλήρη απόδειξη του (i) δες ( ) . Η κεντρική ιδέα της απόδειξης είναι ότι ισχύει η παρακάτω σχέση*

$$(I - bP)^{-1} = \sum_{k=0}^{\infty} b^k P^k \qquad (11)$$

και έτσι ,

$$(I - bP)^{-1} \leq \frac{1}{1-b} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \\ 1 & \cdots & 1 \end{bmatrix} . \qquad (12)$$

Ετσι, η συνάρτηση–πίνακας του b , $P(b) = (I - bP)^{-1}$ είναι ρητή και λόγω της (12) είναι δυνατό να έχει πόλους (δηλ. να μηδενίζεται ο παρανομαστής), μόνο για τιμές του $b : |b| = 1$ και όλοι οι πόλοι είναι απλοί. Ετσι αποδεικνύεται τα (i) , και (ii) και μάλιστα:

$$S = \lim_{b \uparrow 1} (1 - b) R(b) \qquad (13)$$

$$\Delta = \lim_{b \uparrow 1} \left(R(b) - \frac{S}{1-b}\right) \qquad (14)$$

Για να αποδείξουμε το (iii) παρατηρούμε ότι ισχύει η

$$bP(1 - b) R(b) = (1 - b) R(b) + (b - 1)I \qquad (15)$$

Παίρνοντας τα όρια (για $b \uparrow 1$) του αριστερού και του δεξιού μέλους της (15) και χρησιμοποιώντας την (13) βρίσκουμε ότι

$$PS = S. \tag{16}$$

Από την (16) βρίσκουμε ότι

$$P^k S = S \tag{17}$$

Οι (17) και (13) οδηγούν στην $R(b)S = \frac{1}{1-b} S$ *έτσι έχουμε* $(1-b)\, R(b)S = S$ *και παίρνοντας το όριο για* $b \uparrow 1$ *βρίσκουμε ότι:* $S^2 = S$. *Ετσι έχουμε αποδείξει την (8).*

Για την (9) παρατηρούμε ότι $(P-S)^2 = P^2 - 2PS + S^2 = P^2 - S^2$ *και με επαγωγή αποδεικνύεται ότι*

$$(P-S)^n = P^n - S \tag{18}$$

Τώρα

$$\begin{aligned}
\Delta &= \lim_{b \uparrow 1} \left(R(b) - \tfrac{S}{1-b}\right) = \\
&= \lim_{b \uparrow 1} \left(\textstyle\sum_0^\infty b^k P^k - \tfrac{S}{1-b}\right) = \\
&= \lim_{b \uparrow 1} \left(\textstyle\sum_0^\infty b^k P^k - \sum_0^\infty b^k S\right) = \\
&= \lim_{b \uparrow 1} \left(\textstyle\sum_{k=0}^\infty b^k (P^k - S)\right) \overset{(18)}{=} \lim_{b \uparrow 1} \left(\textstyle\sum_0^\infty b^k (P-S)^k\right) - S \\
&= (I - P + S)^{-1} - S .
\end{aligned}$$

Για την (10) έχουμε

$$S\Delta = \lim_{b \uparrow 1} \left(\textstyle\sum_0^\infty b^k S P^k - \tfrac{S}{1-b}\right) \overset{(9)}{=} \lim_{b \uparrow 1} \left(\textstyle\sum_0^\infty b^k S - \tfrac{S}{1-b}\right) = 0 .$$

Πόρισμα. *Για κάθε απλή πολιτική* π *έχουμε*

$$w(\pi) = \frac{\phi(\pi)}{1-b} + \delta(\pi) + o(1), \tag{19}$$

όπου τα διανύσματα ϕ *και* δ *ικανοποιούν τις σχέσεις*

$$\phi(\pi) = P(\pi)\, \phi(\pi) \tag{20}$$

$$\phi(\pi) + \delta(\pi) = c(\pi) + P(\pi)\delta(\pi) \tag{21}$$

$$S(\pi)\delta(\pi) = 0, \tag{22}$$

όπου $S(\pi) = \lim_{k \to \infty} \sum_{t=0}^{k-1} P^t(\pi)$.

Απόδειξη. *Το αποτέλεσμα έπεται εύκολα απο το λήμμα 2 αν ορίσουμε τις ποσότητες*

$$\phi(\pi) = S(\pi)c(\pi) \text{ και } \delta(\pi) = \Delta(\pi)c(\pi).$$

Παρατηρήσεις. ***1.*** *Ο πίνακας* $P(\pi)$ *ορίζεται έτσι ώστε να περιγράφει την δυναμική του συστήματος κάτω απο την απλή πολιτική* π. *Συγκεκριμένα*

$$P(\pi) = (P(y/x,\pi)) \;\; x,y \in S,$$

όπου

$$P(y/x,\pi) \;=\; 1 \;\; \alpha\nu \;\; \overline{\pi}(x) \;=\; y\,,$$

και

$$P(y/x,\pi) \;=\; 0 \;\; \alpha\nu \;\; \overline{\pi}(x) \;\neq\; y\,.$$

2. *Ολες οι γραμμές του πίνακα S (ή S(π)) παραπάνω είναι οι ίδιες (γιατί;) και τα στοιχεία του ονομάζονται οριακές πιθανότητες του συστήματος κάτω απο την δυναμική που ορίζεται απο τον πίνακα* $P(\pi)$.

3. *Αν γράψουμε κατά στοιχεία τις σχέσεις (20) , (21) , (22) παίρνουμε τις*

$$\phi(x,\pi) \;=\; \phi(\overline{\pi}(x),\pi) \tag{20.a}$$
$$\phi(x,\pi) + \delta(x,\pi) \;=\; c(x,\pi(x)) + \delta(\overline{\pi}(x),\pi) \tag{21.a}$$
$$\textstyle\sum_{x\in x} S(x,\pi)\;\phi(x,\pi) \;=\; 0\,, \tag{22.a}$$

όπου $s(x,\pi)$ *είναι το* x *στοιχείο μιας (οποιασδήποτε) γραμμής του πίνακα* $S(\pi)$.

6.2 Εξισώσεις Βελτιστοποίησης για το Πρόβλημα του Μέσου Κόστους. *Ας υποθέσουμε ότι εξετάζουμε ένα πρόβλημα σύμφωνα με το κριτήριο του εκπτωτικού μέτρου απόδοσης και ας υποθέσουμε ότι υπάρχει μια πολιτική* π^0 *που είναι βέλτιστη για κάθε τιμή του εκπτωτικού παράγοντα* $b \in (b_0,1)$. *Μια τέτοια πολιτική* π^0 *υπάρχει όταν οι χώροι καταστάσεων και αποφάσεων είναι πεπερασμένοι (βλ. Παρατήρηση 1 παρακάτω). Οι εξισώσεις βελτιστοποίησης για την* π^0 *είναι*

$$w(x,\pi^0,b) \;=\; v(x,b) \;=\; \min_{\alpha\in D(x)} \{c(x,\alpha) + bv(\overline{\alpha},b)\} \tag{23}$$

Εστω $\phi = \phi(\pi^0)$, $\delta = \delta(\pi^0)$. *Χρησιμοποιώντας το λήμμα (2) παραπάνω βρίσκουμε ότι*

1) $$\frac{\phi(x,\pi^0)}{1-b} \;+\; \delta(x,\pi^0) \;=\; \min_{\alpha\in D(x)} \Big\{c(x,\alpha) + \;b\;\Big(\; \frac{\phi(\overline{\alpha},\pi^0)}{1-b} + \delta(\overline{\alpha},\pi^0)\;\Big)\;\Big\} + \;o(1) \tag{24}$$

και από την (24) θεωρώντας τιμές του b *κοντά στο* 1 *παίρνουμε ότι*

$$\phi(x,\pi^0) \;=\; \min_{\alpha\in D(x)} \;\phi(\overline{\alpha},\pi^0) \tag{25}$$

Επίσης, λόγω της (23) έχουμε

$$\begin{aligned}
\delta(x,\pi^0) \quad &= \lim_{b\uparrow 1}\,\Big(w(x,\pi^0,b) - \;\frac{\phi(x,\pi^0)}{1-b}\Big) \;\;\overset{(23)}{\leq}\\
&\leq\;\; \lim_{b\uparrow 1}\,\Big(c(x,\alpha) + b(w(\overline{\alpha},\pi^0,b) - \;\frac{b\phi(x,\pi^0)}{1-b}\; - \;\phi(x,\pi^0)\Big)\\
&= \lim_{b\uparrow 1}\Big(c(x,\alpha) + b(w(\overline{\alpha},\pi^0,b) - \frac{\phi(\overline{\alpha},\pi^0)}{1-b}) + \frac{b}{1-b}(\phi(\overline{\alpha},\pi^0) - \phi(x,\pi^0)) - \phi(x,\pi^0)\Big)\\
&=\; c(x,\,\alpha) + \delta(\overline{\alpha},\pi^0) + \frac{b}{1-b}(\phi(\overline{\alpha},\pi^0) - \phi(x,\pi^0)) - \phi(x,\pi^0)\,.
\end{aligned}$$

Δηλαδή παίρνουμε

$$\phi(x,\pi^0) \;\leq\; c(x,\,\alpha) + \delta(\overline{\alpha},\pi^0)\,,\quad \forall x\,,\;\; \forall\alpha\in D(x)\,. \tag{26}$$

*Δηλαδή βλέπουμε (δεδομένου ότι η (26) ισχύει σαν ισότητα για $\alpha = \pi^0(x)$), ότι για να είναι μια πολιτική βέλτιστη για τιμές του b κοντά στο 1, **αρκεί** να ισχύουν οι εξής σχέσεις για την π^0*

$$\phi(x,\pi_0) = \min_{\alpha \in D(x)} \phi(\overline{\alpha},\pi^0), \quad x \in S, \tag{27}$$

και

$$\phi(x,\pi^0) \neq \delta(x,\pi^0) = \min_{\alpha \in D(x)} \{c(x,\alpha) + \delta(\overline{\alpha},\pi^0)\}, \quad x \in S. \tag{28}$$

Θα δούμε παρακάτω ότι κάτω απο κατάλληλες συνθήκες οι (27), (28) είναι αναγκαίες και ικανές για την ύπαρξη και τον καθορισμό της βέλτιστης πολιτικής για το πρόβλημα με κριτήριο το μέσο κόστος.

<u>Παρατήρηση.</u> *Το $w(x,\pi,b)$ είναι ρητή συνάρτηση του b για κάθε απλή πολιτική π, μπορεί δηλαδή να βρεθεί σαν λύση ενόςγραμμικού συστήματος. Επομένως για δύο απλές πολιτικές π_1, π_2, η διαφορά $(w(x,\pi_1,b) - w(x,\pi_2,b))$ είναι επίσης ρητή συνάρτηση του b, και έτσι μπορεί να αλλάξει πρόσημο το πολύ ένα πεπερασμένο αριθμό φορών στο διάστημα $b \in [0,1)$. Αυτό, σε συνδυασμό με το γεγονός ότι οι χώροι καταστάσεων S και αποφάσεων $D(x)$, $x \in S$, είναι πεπερασμένοι, μας οδηγεί στο συμπέρασμα ότι υπάρχει το πολύ ένας πεπερασμένος αριθμός πολιτικών.*

Για την κατανόηση των παραπάνω ας θεωρήσουμε το εξής:

<u>Εισαγωγικό Πρόβλημα.</u> *Εχουμε $S = \{1,2,3\}$, $D(1) = \{(1,2),(1,3)\}$, $D(2) = \{(2,1)\}$, $D(3) = \{(3,1),(3,3)\}$.*

1 2

3 3

Εικόνα 1

Το κόστος της κάθε απόφασης δίδεται απο τον αριθμό δίπλα στο αντίστοιχο βέλος της της εικόνας 1.

Εστω x_0 η αρχική κατάσταση. Είναι εύκολο να δούμε ότι στο εισαγωγικό πρόβλημα υπάρχουν οι εξής απλές πολιτικές

$$\pi_1 := \{(1,2),(2,1),(3,3)\}, \qquad \pi_2 := \{(1,2),(2,1),(3,2)\},$$
$$\pi_3 := \{(1,3),(2,1),(3,3)\}, \qquad \pi_4 := \{(1,3),(2,1),(3,2)\}.$$

Εστω $\phi(\pi) = (\phi(\pi,x)) = (\phi(\pi,1),\phi(\pi,2),\phi(\pi,3))$. Ανάλογα με την πολιτική που επιλέγουμε οι τιμές του κριτηρίου μέσου κόστους είναι εύκολο να βρεθούν. Ετσι έχουμε

$$\phi(\pi_1) = (4,4,3), \ \phi(\pi_2) = (4,4,4), \ \phi(\pi_3) = (3,3,3), \ \phi(\pi_4) = (2,2,2).$$

Πραγματικά για παράδειγμα

$$\phi(\pi_1,1) = \lim_{n \to \infty} \tfrac{1}{n}(5 + 3 + 5 + 3 + \ldots) =$$

$$= \lim_{n \to \infty} \left(\tfrac{1}{n}\left(\tfrac{8n}{2}\right) + o(n)\right) = 4,$$

$$\phi(\pi_1,3) = \lim_{n \to \infty} \tfrac{1}{n}(3 + 3 + \ldots) = 3.$$

Παρατηρούμε λοιπόν ότι η βέλτιστη πολιτική είναι η π_4 και $\phi(\pi_4, x) \leq \phi(\pi_k, x) \quad \forall x$, $k = 1,2,3$.

Ας θεωρήσουμε τις ανάλογες τιμές του εκπτωτικού κόστους

$$w(\pi, x, b) = \sum_0^{\infty} b^t c(x_t, \overline{\pi}(x_t)) .$$

Είναι εύκολο να υπολογίσουμε τις τιμές $w(\pi)$, π.χ. όταν έχουμε την πολιτική π_1, οι αντίστοιχες τιμές $w(\pi_1)$ βρίσκονται σαν η μοναδική λύση του παρακάτω συστήματος

$$w(\pi_1, 1, b) = 5 + b w(\pi_1, 2, b) \tag{29.α}$$

$$w(\pi_1, 2, b) = 3 + b w(\pi_1, 1, b) \tag{29.β}$$

$$w(\pi_1, 3, b) = 3 + b w(\pi_1, 3, b) \tag{29.γ}$$

Το παραπάνω σύστημα μπορεί να γραφεί σύντομα ως

$$w(\pi_1) = c(\pi_1) + bP(\pi_1) w(\pi_1) \tag{30}$$

Η λύση του (30) είναι εύκολο να βρεθεί για $b < 1$

$$w(\pi_1) = (I - bP(\pi_1))^{-1}\, c(\pi 1) = \left(\tfrac{5+3b}{1-b}, \tfrac{3+5b}{1-b}, \tfrac{3}{1-b}\right)$$

Παρατηρούμε λοιπόν, ότι σύμφωνα με το λήμμα 2, έχουμε ότι

$$w(1, \pi_1, b) = \frac{4}{1-b} + \frac{1}{2} + \left(\frac{1}{1+b} - \frac{1}{2}\right) = \frac{4}{1-b} + \frac{1}{2} + o(1)$$

$$w(2, \pi_1, b) = \frac{4}{1-b} - \frac{1}{2} + o(1)$$

$$w(3, \pi_1, b) = \frac{3}{1-b} + 0 + o(1) .$$

Δηλαδή $\delta(1) = 1/2$, $\delta(2) = -1/2$, $\delta(3) = 0$, $\phi(1, \pi_1) = \phi(2, \pi_1) = 4$, & $\phi(3, \pi_1) = 3$.

Για να δούμε αν η π_1 είναι βέλτιστη σύμφωνα με το κριτήριο του μέσου κόστους, ας εξετάσουμε τις (27), (28).

Η (25) μας δίδει τις παρακάτω ανισότητες

$$\phi(x) \leq \phi(\overline{\alpha}) \quad \forall \alpha \in D(x)$$

Δηλ.

$$\phi(1) \leq \phi(2) \tag{31}$$

$$\phi(1) \leq \phi(3) \tag{32}$$

$$\phi(2) \leq \phi(1) \tag{33}$$

$$\phi(3) \leq \phi(2) . \tag{34}$$

βλέπουμε ότι η (32) παραβιάζεται, άρα η π_1 δεν είναι βέλτιστη για τιμές του b κοντά στο 1.

Ας εξετάσουμε λοιπόν την πολιτική που στην κατάσταση 1 ορίζει την απόφαση (1,3), και κατά τις άλλες αποφάσεις παραμένει η ίδια με την π_1. Δηλαδή οδηγούμαστε στην πολιτική π_3. Για την π_3 έχουμε

$$w(1, \pi_3, b) = \frac{3}{1-b} - 2$$

$$w(2,\pi_3,b) \;=\; \frac{3}{1-b} \;-\; 2 \;+\; 2(1-b)$$

$$w(3,\pi_3,b) \;=\; \frac{3}{1-b} \;+\; 0 \;+\; 0\,.$$

Δηλαδή έχουμε : $\delta(1) \;=\; \delta(2) \;=\; -2\,,\;\; \delta(3) \;=\; 0$ *και*

$$\phi(1,\pi_3) \;=\; \phi(3,\pi_3) \;=\; 3\,.$$

Παρατηρούμε λοιπόν ότι οι σχέσεις (27) ισχύουν. Αν εξετάσουμε όμως τις σχέσεις (28), έχουμε

$$\phi(1,\pi_3) + \delta(1,\pi_3) \;\overset{;}{\leq}\; c(1,2) + \delta(2,\pi_3) \tag{35}$$

$$\phi(3,\pi_3) + \delta(3,\pi_3) \;\overset{;}{\leq}\; c(3,2) + \delta(2,\pi_3)\,, \tag{36}$$

και παρατηρούμε ότι η (35) ισχύει $(3 + (-2) < 5 + (-2))$, *ενώ η (36) δεν ισχύει* $(3 + 0 \leq 2 + (-2))$.

Αρα η π_3 *δεν είναι βέλτιστη για τιμές του* b *κοντά στο* 1.

Ας εξετάσουμε λοιπόν την πολιτική που στην κατάσταση 3 ορίζει την απόφαση (3,2) (η οποία απόφαση παραβίαζε την (36)). Παίρνομε λοιπόν την πολιτική π_4, *για την οποία είναι εύκολο να δούμε ότι οι αντίστοιχες στις (27), (28) εξισώσεις ισχύουν. Εχουμε λοιπόν ότι η πολιτική* π_4 *είναι βέλτιστη για τιμές του* b *κοντά στο* 1.

***<u>Παρατηρήσεις</u>* 1.** *Οι σχέσεις (27), (28) παραπάνω μας οδήγησαν στην βέλτιστη πολιτική* π_4 *με τρόπο ανάλογο της διαδικασίας των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών για τα προβλήματα που έχουμε ήδη εξετάσει.*

Πραγματικά, θα δούμε παρακάτω, ότι υπάρχει μια μέθοδος διαδοχικών προσεγγίσεων στο χώρο των πολιτικών για την εύρεση της βέλτιστης (των βέλτιστων) πολιτικής (πολιτικών) για το κριτήριο του μέσου κόστους, της οποίας ειδική περίπτωση είναι, ουσιαστικά, η διαδικασία που ακολουθήσαμε στο παραπάνω εισαγωγικό πρόβλημα.

2. *Ενας επιπλέον τρόπος για να κατανοήσουμε το νόημα των εξισώσεων (27), (28) είναι ο εξής. Εστω* π *μια απλή πολιτική για την οποία το όριο της ακολουθίας* $w_n(x,\pi) = \lim_{n\to\infty} w_n(x,\pi)/n$ *υπάρχει. Δηλαδή*

$$\phi(x,\pi) = \lim_{n\to\infty} w_n(x,\pi)/n \tag{37}$$

Απο την εξίσωση (37) έπεται ότι :

$$w_n(x,\pi) \;=\; \phi(x,\pi)\,n \;+\; g(x,n,\pi)\,, \tag{38}$$

όπου $g(x,n)$ *είναι μια* $o(n)$ *συνάρτηση (για* $n \to \infty$*), δηλαδή* $g(x,n)/n \to 0$ $(n \to \infty)$. *Από τον ορισμό του* $w_n(x,\pi)$ *(σαν την απόδοση της πολιτκής* π *για τον πεπερασμένο ορίζοντα* $\{0,\ldots,n\}$ *όταν η αρχική κατάσταση είναι η* x*), έχουμε ότι*

$$w_{n+1}(x,\pi) \;=\; c(\overline{\pi}(x)) + w_n(\overline{\pi}(x),\pi) \tag{39}$$

Από τις (38), (39) παίρνουμε ότι

$$(n+1)\phi(x,\pi) + o(n) \;=\; c(\overline{\pi}(x)) + n\phi(\overline{\pi}(x),x) + o(n) \tag{40}$$

και η τελευταία συνεπάγεται ότι

$$\phi(x,\pi) \;=\; \phi(\overline{\pi}(x),\pi) \tag{41}$$

Εάν ήταν δυνατό οι συναρτήσεις που εμφανίζονται στην (38) να γραφούν ως εξής

$$g(x,n,\pi) \;=\; \delta(x,\pi) + \epsilon(n,\pi)\,, \tag{42}$$

όπου $\epsilon(n,\pi)$ *είναι μια* $o(1)$ *συνάρτηση (για* $n \to \infty$*) , δηλ.* $\epsilon(n) \to \infty$ $(n \to \infty)$ *, τότε η (37) θα μπορούσε να γραφεί ως εξής*

$$(n+1)\,\phi(x,\pi) + \delta(x,\pi)\,o(1) \;=\; c(\overline{\pi}(x)) \;+\; n\phi(\overline{\pi}(x),\pi) \;+\; \delta(\overline{\pi}(x),\pi) \;+\; o(1) \tag{43}$$

Οι εξισώσεις (41) , (43) μας οδηγούν στην

$$\phi(x,\pi) \;+\; \delta(x,\pi) \;=\; c(\overline{\pi}(x)) \;+\; \delta(\overline{\pi}(x),\pi) \tag{44}$$

Για μια δεδομένη πολιτική, οι εξισώσεις (41) , (44) είναι ανάλογες των εξισώσεων (27), (28) .

Εστω $\phi(x) \;=\; \inf_{\pi} \phi(x,\pi)$ *,* $v(x) \;=\; \inf_{\pi} w_n(x,\pi)$ *. Αν ήταν δυνατό να υποθέσουμε ότι*

$$v_n(x) \;=\; n\phi(x) \;+\; g\prime(x,\,n)\,, \tag{45}$$

όπου $g\prime(x,n)$ *είναι* $o(n)$ $(n \to \infty)$ *, τότε παίρνουμε την σχέση*

$$\phi(x) \;=\; \inf_{\pi(x)\,\in\, D(x)} \phi(\overline{\pi}(x)) \;=\; \inf_{\alpha\,\in\, D(x)} \phi(\overline{\alpha})\,. \tag{46}$$

Αν επιπλέον, δεχθούμε ότι οι συναρτήσεις $g\prime(x,n)$ *έχουν μορφή ανάλογη της* $g(x,n,\pi)$ *, τότε οδηγούμαστε στην*

$$\phi(x) + \delta(x) \;=\; \inf_{a\,\in\, D(x)} \{c(\overline{\alpha}) + \delta(\overline{\alpha})\} \tag{47}$$

Η εξίσωση (42) δεν ισχύει πάντα, δεδομένου ότι είναι δυνατό οι συναρτήσεις $g(x,n,\pi)$*,* $(g\prime(x,n))$ *να περιέχουν όρους περιοδικούς ως προς* n *(π.χ. της μορφής* $\gamma(-1)^n$ *άσκηση) . Ισχύει στην περίπτωση που η δυναμική κίνηση του συστήματος δεν εμφανίζει περιοδικότητα , δηλαδή δεν υπάρχει κατάσταση* x_o *, για την οποία να υπάρχει μια σταθερή* d *, έτσι ώστε κάτω απο μια πολιτική* π *:* $x_t = x_o$ *για* $t = kd$ *,* $k = 0,1,2,\ldots$ *και* $x_t \neq x_o$ *για* $t \neq kd$ *,* $k = 0,1,\ldots$ *. Η περιοδικότητα είναι φαινόμενο που ισχύει πάντα στην περίπτωση της προσδιοριστικής δυναμικής, και είναι δυνατό να εμφανίζεται και στην περίπτωση της πιθανοθεωρητικής δυναμικής, όπως θα δούμε παρακάτω.*
Οταν δεν υπάρχει περιοδικότητα στη δυναμική του συστήματος, τότε οι (38) , (42) μας οδηγούν στην σχέση

$$w_n(x,\pi) \;=\; \phi(x,\pi)n \;+ \delta(x,\pi) + o(1)\,, \tag{48}$$

η οποία συνεπάγεται την εξής ερμηνεία για το $\delta(x,\pi)$ *: Η ποσότητα* $\delta(x,\pi)$ *είναι η* ***επίδραση*** *της αρχικής κατάστασης στο ολικό κόστος, όταν ο ορίζοντας είναι μεγάλος (δηλαδή όταν το κομάτι* $o(1)$ *παίρνει τιμές κοντά στο 0) . Ετσι για δύο διαφορετικές αρχικές καταστάσεις* $x_o, x_o\prime$ *, η διαφορά* $\delta(x_o,\pi) - \delta(x_o',\pi)$ *μας δίδει την* ***σχετική αξία*** *των καταστάσεων, π.χ. το ποσό που κάποιος λογικός παίκτης θα ήταν διατεθειμένος να “πληρώσει", έτσι ώστε να του επιτραπεί να διαλέξει σαν αρχική κατάσταση την* x_o' *αντί της* x_o *, εάν είχε σκοπό “να παίξει" για ένα μακρύ χρονικό διάστημα χρησιμοποιώντας την πολιτική* π*.*

3. *Παρατηρούμε λοιπόν από τα παραπάνω, ότι δεν πρέπει να είναι οι απόλυτες τιμές των* $\delta(x,\pi)$ *που είναι σημαντικές για τον καθορισμό μιας βέλτιστης πολιτικής, αλλά οι διαφορές :* $\delta(x,\pi) - \delta(x',\pi)$ *. Πράγματι, είναι εύκολο να δούμε ότι αν η εξίσωση (47) ισχύει για μια*

συγκεκριμένη συνάρτηση $\delta(x,\pi)$, τότε θα ισχύει για την συνάρτηση $\delta'(x,\pi) = \delta(x,\pi) + d$, για κάθε σταθερή τιμή d.
Ακολουθώντας τον αντίστροφο συλογισμό, βλέπουμε ότι είναι επιτρεπτό να ορίσουμε $\delta(x,\pi) = 0$ για όσες καταστάσεις θέλουμε, εφόσον δεν επηρεάζουμε τις **διαφορές :** $\delta(x,\pi) - \delta(x',\pi) \quad x,x' \in S$, όταν εξετάζουμε το κατά πόσο η πολιτική π είναι βέλτιστη.

6.3 Επανεξέταση του Εισαγωγικού Προβλήματος. Στη πρώτη λύση του εισαγωγικού προβλήματος εχρησιμοποιήσαμε τα αναπτύγματα των ποσοτήτων $w(x,\pi,b)$ για να υπολογίσουμε τις τιμές των $\phi(x,\pi)$, $\delta(x,\pi)$. Σε τούτο το τμήμα θέλουμε να δείξουμε ότι οι ποσότητες $\phi(x,\pi)$, $\delta(x,\pi)$ είναι δυνατό να υπολογιστούν κατ' ευθείαν από τις εξισώσεις (20), (21) και (22), πράγμα απλούστερο.

Πραγματικά, ας θεωρήσουμε την πολιτική π_1 η οποία συνεπάγεται την
!!FOOTNOTE 1:Για απλούστευση του συμβολισμού γράφουμε $\phi(x)$ αντί $\phi(x,\pi)$, όταν δεν υπάρχει φόβος σύγχυσης για την π.
δυναμική της εικόνας 2.

5

1 2

3

3 3

Εικόνα 2

Εχουμε $P(n) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Ο πίνακας $S(\pi)$ μπορεί να υπολογιστεί απο τις σχέσεις (8), και βρίσκουμε ότι :

$$S(\pi) = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Οι σχέσεις (20) $(\phi(\pi) = P(\pi)\phi)$ μας οδηγούν στις εξισώσεις[1] :

$$\phi(1) = \phi(2) \tag{1.1}$$
$$\phi(2) = \phi(1) \tag{1.2}$$
$$\phi(3) = \phi(3) \tag{1.3}$$

Οι σχέσεις (21) $(\phi(\pi) + \delta(\pi) = c(\pi) + P(\pi)\delta(\pi))$ μας οδηγούν στις :

$$\phi(1) + \delta(1) = 5 + \delta(2) \tag{2.1}$$
$$\phi(2) + \delta(2) = 3 + \delta(1) \tag{2.2}$$
$$\phi(3) + \delta(3) = 3 + \delta(3) \tag{2.3}$$

Τέλος οι σχέσεις (22) $(S(\pi)\delta(\pi) = 0)$ μας οδηγούν στις :

$$\tfrac{1}{2}\delta(1) + \tfrac{1}{2}\delta(2) = 0 \tag{3.1}$$

$$\tfrac{1}{2}\delta(1) + \tfrac{1}{2}\delta(2) = 0 \tag{3.2}$$

1

$$\delta(3) \ = \ 0 \, . \qquad (3.3)$$

Παρατηρούμε λοιπόν ότι για τον υπολογισμό των 6 μεταβλητών $\phi(1)$, $\phi(2)$, $\phi(3)$, $\delta(1)$, $\delta(2)$, $\delta(3)$ *έχουμε 9 εξισώσεις, και μάλιστα αρκετές απ' αυτές πλεονάζουν. Το πιο σημαντικό όμως είναι ότι έχουμε δύο ανεξάρτητα υποσυστήματα εξισώσεων, ένα για κάθε μια από τις δύο ανεξάρτητες ομάδες καταστάσεων* $\{1,2\}$ *και* $\{3\}$.

Συγκεκριμένα, έχουμε τα εξής δύο συστήματα (όπου γράφουμε τον ελάχιστο αριθμό εξισώσεων που είναι απαραίτητες για τον υπολογισμό των αγνώστων) :

Το σύστημα για την ομάδα $\{1,2\}$, (Σ_1)

$$\phi(1) \ = \ \phi(2) \, ,$$

$$\phi(1) + \delta(1) \ = \ 5 + \delta(2) \, ,$$

$$\phi(2) + \delta(2) \ = \ 3 + \delta(1) \, ,$$

$$\tfrac{1}{2}\delta(1) + \tfrac{1}{2}\delta(2) \ = \ 0 \ \, ,$$

και το σύστημα για την ομάδα $\{3\}$, (Σ_2):

$$\phi(3) + \delta(3) \ = \ 3 + \delta(3) \, ,$$

$$\delta(3) \ = \ 0 \, .$$

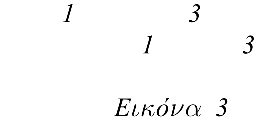
Το σύστημα (Σ_1) *έχει σαν μοναδική λύση την :* $\phi(1) = \phi(2) = 4$, $\delta(1) = 1/2$, $\delta(2) = -1/2$. *Το σύστημα* (Σ_2) *έχει σαν μοναδική λύση την:* $\phi(3) = 3$, $\delta(3) = 0$.

Οι παραπάνω τιμές συμφωνούν με εκείνες που βρήκαμε όταν πήραμε τα αναπτύγματα των $w(x,\pi,b)$, *όπως άλλωστε περιμέναμε.*

Αν αντικαταστήσουμε την τελευταία εξίσωση του συστήματος Σ_1 *με την εξίσωση* $\delta(2) = 0$, *έχουμε ότι η μόνη λύση του νέου συστήματος είναι η* $\phi'(1) = 4$, $\delta'(1) = 1$, $\delta'(2) = 0$. *Παρατηρούμε λοιπόν ότι η νέα τιμή του* $\phi(1)$ *είναι η ίδια με την προηγούμενη, ενώ η τιμή της διαφοράς* $\delta'(1) - \delta'(2) = 1$ *είναι η ίδια με την τιμή της διαφοράς* $\delta(1) - \delta(2) = 1/2 - (-1/2) = 1$, *όπως περιμέναμε από την παρατήρηση 3 παραπάνω.*

Το κύριο σημείο είναι ότι είναι δυνατό να υπολογίσουμε τις τιμές του $\phi(x)$, *και τιμές* $\delta'(x)$ *(για τις οποίες* $\delta'(x) - \delta'(x') = \delta(x) - \delta(x')$*)* , *χρησιμοποιώντας τις εξισώσεις (20) και (21) και αποφεύγοντας τις εξισώσεις (22) , οι οποίες περιέχουν τον πίνακα* $S(\pi)$, *που πρέπει να υπολογιστεί ανεξάρτητα.*

Τώρα μπορούμε να συνεχίσουμε τη διαδικασία βελτίωσης της πολιτικής την οποία χρησιμοποιήσαμε στο εισαγωγικό παράδειγμα. Ετσι έχουμε ότι η ανισότητα $\phi(1) \leq \phi(3)$ *παραβιάζεται, άρα η* π_1 *δεν είναι βέλτιστη. Ετσι οδηγούμαστε στην πολιτική* $\pi_3 := \{(1,3),(2,1),(3,3)\}$, *με δυναμική που περιγράφεται στην εικόνα 3. Οι εξισώσεις (20) είναι οι εξής*



Εικόνα 3

$$\phi(1) \ = \ \phi(3) \qquad (4.1)$$

$$\phi(2) \ = \ \phi(1) \qquad (4.2)$$

$$\phi(3) \ = \ \phi(3) \, , \qquad (4.3)$$

και οι εξισώσεις (21) είναι :

$$\phi(1) + \delta(1) \ = \ 3 \tag{5.1}$$

$$\phi(2) + \delta(2) \ = \ 3 + \delta(1) \tag{5.2}$$

$$\phi(3) + \delta(3) \ = \ 3 + \delta(3) \tag{5.3}$$

Αντί της εξίσωσης που θα παίρναμε από την απαίτηση $S(\pi_3)\delta(\pi_3) \ = \ 0$ *, ας προσθέσουμε την απαίτηση ότι :*

$$\delta(3) \ = \ 0 \tag{6}$$

στις παραπάνω εξισώσεις, δεδομένου ότι έχουμε μια μόνο ομάδα καταστάσεων. (Σε τούτη την περίπτωση είναι εύκολο να δούμε ότι :

$$S(\pi) \ = \ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

και έτσι $\delta(3) \ = \ 0$ *θα ήταν ούτως ή άλλως ο περιορισμός που θα μας επέβαλλαν οι εξισώσεις (22)) .*

Το σύστημα των εξισώσεων (4)—(6) έχει σαν μοναδική λύση την εξής :

$$\phi(1) \ = \ \phi(2) \ = \ \phi(3) \ = 3 \ , \quad \delta(1) \ = \ -\ 2 \ , \quad \delta(2) \ = \ -\ 2 \ , \quad \delta(3) \ = \ 0 \tag{7}$$

Στη λύση (7) είχαμε οδηγηθεί και με τον προηγούμενο τρόπο λύσης.

Παρατηρούμε ότι έχουμε ελευθερία ως προς την εξίσωση (6) , λόγω της παρατήρησης 3 παραπάνω. Συγκεκριμένα, αν πάρουμε αυθαίρετα :

$$\delta(3) \ = \ 8 \ , \tag{8}$$

τότε η μόνη λύση του συστήματος των (4) , (5) , (7) είναι η εξής :

$$\phi(1) = \phi(2) = \phi(3) = 3 \ , \quad \delta(1) = \ -\ 2 \ , \quad \delta(2) = \ -\ 2 \ , \quad \delta(3) = 0 \ , \tag{9}$$

και έτσι έχουμε τις ίδιες τιμές για τα $\phi(x)$ *, και τις ίδιες τιμές για τις διαφορές*
!!FOOTNOTE 2:Είναι εύκολο να δούμε ότι

$$S(\pi_4) \ = \ \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix},$$

και έτσι, η εξίσωση που αντιστοιχεί στην απαίτηση $S(\pi_4)\ \delta(\pi_4) \ = \ 0$ *οδηγεί στην εξίσωση :*

$$\delta(1) \ + \ \delta(2) \ + \ \delta(3) \ = \ 0.$$

$\delta(x) - \delta(x')$ *για τις λύσεις (7) και (9) , όπως περιμέναμε. Επίσης θα μπορούσαμε να είχαμε θέσει* $\delta(1) \ = \ 0$ *, οπότε σαν μόνη λύση θα βρίσκαμε :*

$$\delta(1) = \delta(2) = 0 \ , \ \ \delta(3) = \text{μη περιορισμένο} \ . \tag{10}$$

Οι ανισότητες (35) , (36) για τον έλεγχο της βελτιστότητας της π_3 *είναι οι ίδιες, ανεξάρτητα απο το αν χρησιμοποιήσουμε την λύση (7) , ή τις λύσεις (9) ή (10) . Ετσι όπως και πριν οδηγούμαστε στην πολιτική* π_4 *για την οποία η δυναμική δίδεται στην εικόνα 4 ,* *και οι εξισώσεις για τα* $\phi(x) = \phi(x,\pi_4)$, $\delta(x) = \delta(x,\pi_4)$, *που αντιστιχούν στις (20) , (21) είναι οι εξής :*

2
3 2
1 3
1

Εικόνα 4

$$\phi(1) = \phi(3) \quad (11.1)$$
$$\phi(3) = \phi(2) \quad (11.2)$$
$$\phi(2) = \phi(1) \quad (11.3)$$

$$\phi(1) + \delta(1) = 1 + \delta(3) \quad (12.1)$$
$$\phi(2) + \delta(2) = 3 + \delta(1) \quad (12.2)$$
$$\phi(3) + \delta(3) = 2 + \delta(2) \quad (12.3)$$

Επιπλέον, επειδή έχουμε μια μόνο κλειστή ομάδα καταστάσεων, "μας επιτρέπεται" να διαλέξουμε αυθαίρετα την (σχετική) τιμή μιας μόνο από τις μεταβλητές $\delta(1)$, $\delta(2)$, $\delta(3)$. *Εστω λοιπόν ότι διαλέγουμε :*

$$\delta(1) = 0 .\text{[2]} \quad (13.1)$$

Η μόνη λύση του (11) , (12) και (13) είναι η εξής :

$$\phi(1) = \phi(2) = \phi(3) = 2 ,\ \delta(1) = 0 ,\ \delta(2) = 1 ,\ \delta(3) = 1 . \quad (14)$$

Τώρα, οι ανισότητες της μορφής (27) ισχύουν κατά προφανή τρόπο, ενώ οι ανισότητες της μορφής (28) είναι οι εξής :

κατάσταση 1 : $\phi(1) + \delta(1) \le c(1,2) + \delta(2) \Leftrightarrow 2 \le 5 + 1 \Leftrightarrow 2 \le 6$ *ισχύει* (15.1)
κατάσταση 3 : $\phi(3) + \delta(3) \le c(3,3) + \delta(3) \Leftrightarrow 2 + 1 \le 3 + 1 \Leftrightarrow 3 \le 4$ *ισχύει* , (15.2)

άρα η π_4 *είναι βέλτιστη.*

Παρατηρήσεις. ***1.*** *Τονίζουμε ότι οι ανισότητες της μορφής (15.1) , (15.2) παραπάνω εξαρτώνται μόνο απο το μέγεθος των διαφορών* $\delta(2) - \delta(1)$, *π.χ. , και όχι από τις απόλυτες τιμές των* $\delta(x)$.

2. *Η παραπάνω μέθοδος λύσης του εισαγωγικού προβλήματος είναι η μέθοδος των διαδοχικών προσεγγίσεων στον χώρο των πολιτικών. Είναι δε εύκολο να γενικευθεί για την περίπτωση ενός αυθαίρετου προβλήματος, με πεπερασμένους χώρους καταστάσεων* S *και αποφάσεων* $D(x)$, $x \in S$.

3. *Αν εξετάσουμε προσεκτικά την λύση (14) , όπου εθέσαμε στην αρχή αυθαίρετα* $\delta(1) = 0$, *θα δούμε ότι :*

$$\delta(2) = 3 - \phi \quad (16)$$

$$\delta(3) \overset{(12.3)}{=} 2 - \phi + (2) \overset{(16)}{=} (2 - \phi) + (3 - \phi) + \delta(1) = 2 + 3 - 2\phi =$$

$$= \sum_{t=0}^{\tau-1} (c\,(x_t, \overline{\alpha}_t) - \phi) , \quad (17)$$

όπου στην (17) , $\tau = 2$ *είναι ο "χρόνος" (αριθμός βημάτων) που απαιτείται για να φθάσουμε στην κατάσταση 1 (για την οποία διαλέξαμε* $\delta(1) = 0$) , *όταν η αρχική κατάσταση είναι η 3 . Δηλαδή βρίσκουμε άλλη ερμηνεία για τις ποσότητες* $\delta(x)$, *σαν το ολικό κόστος της μετάβασης από την αρχική κατάσταση* x , *στην κατάσταση* x_o *της κλειστής ομάδας στην οποία η* x *ανήκει, για την οποία διαλέξαμε* $\delta(x_o) = 0$, *όπου σε κάθε βήμα το κόστος αναπροσαρμόζεται κατά* $\phi(x)$, *το μέσο κόστος της κλειστής ομάδας. Τούτη*

2

η ερμηνεία των σχετικών τιμών ισχύει κάτω απο τις πιο γενικές συνθήκες του προβλήματος όπως θα δούμε παρακάτω.

6.4 Η Αναγκαιότητα των Εξισώσεων Βελτιστοποίησης. *Στα προηγούμενα είδαμε ότι οι συναρτησιακές εξισώσεις (1) και (2) παρακάτω* ***αρκεί*** *να ισχύουν (είναι* ***ικανές****) για να είναι μια πολιτική π^o βέλτιστη ως προς το κριτήριο του μέσου κόστους.*

$$\phi(x,\pi^o) = \min_{\alpha \in D(x)} \phi(\overline{\alpha},\pi^o) \tag{1}$$

$$\phi(x,\pi^o) + \delta(x,\pi^o) = \min_{\alpha \in D(x)} \{c(\overline{\alpha}) + \delta(\overline{\alpha},\pi^o)\} \tag{2}$$

Οδηγηθήκαμε στις παραπάνω εξισώσεις χρησιμοποιώντας ιδιότητες του εκπτωτικού κόστους, και το γεγονός ότι για προβλήματα με πεπερασμένους χώρους καταστάσεων και αποφάσεων υπάρχουν πολιτικές βέλτιστες ως προς το δεύτερο κριτήριο, για όλες τις τιμές του εκπτωτικού παράγοντα β κοντά στο 1 (δηλαδή $\forall b \in (b_0,1)$, για κάποιο $b_0 < 1$).

Σε τούτο το τμήμα θα αποδείξουμε κατ' ευθείαν ότι οι (1) και (2) ***πρέπει*** *να ισχύουν, για να είναι η π^o μια βέλτιστη πολιτική ως προς το κριτήριο του μέσου κόστους, για την γενική περίπτωση χώρων καταστάσεων και αποφάσεων.*

Εστω λοιπόν

$$\phi^*(x) = \inf_{\pi} \left(\limsup_{n \to \infty} w_n(x,\pi)/n \right). \tag{3}$$

Το κύριο αποτέλεσμα είναι το εξής

Θεώρημα 1

1. Εάν υπάρχουν συναρτήσεις $\psi(x)$, $d(x)$ ομοιόμορφα φραγμένες (δηλ. $|\psi(x)| \le M_1$, $|d(x)| \le M_2$ $\forall x$, M_1,M_2 σταθερές) που ικανοποιούν τις εξισώσεις

$$\psi(x) = \inf_{\alpha \in D(x)} \psi(\overline{\alpha}) \tag{4}$$

$$\psi(x) + d(x) = \inf_{\alpha \in D(x)} \{c(\overline{\alpha}) + d(\overline{\alpha})\}, \tag{5}$$

τότε

$$\psi(x) \le \phi^*(x), \ \forall x \in S. \tag{6}$$

2. Εάν για μια απλή πολιτική π_o ισχύουν οι ανισότητες

$$\psi(x) \ge \psi(\phi(\overline{\pi}_o(x))) - \epsilon/2 \tag{7}$$

και

$$\psi(x) + d(x) \ge c(\overline{\pi}_o(x)) + \psi(\overline{\pi}_o(x)) - \epsilon/2, \tag{8}$$

για κάποια σταθερά $\epsilon > 0$, τότε

$$\phi(x,\pi_o) \ge \phi^*(x) - \epsilon, \tag{9}$$

οπότε η π^o είναι $\epsilon -$ βέλτιστη (δηλαδή $\phi(x,\pi^o) \in (\phi^*(x), \phi^*(x) + \epsilon)$).

3. Εαν οι (4) και (5) παραπάνω ισχύουν με min στην θέση των inf και επιπλέον για κάθε κατάσταση $x_o = x$ τα ελάχιστα επιτυγχάνονται για τις ίδιες αποφάσεις $\alpha(x) = \pi^o(x)$, τότε η πολιτική π^o είναι βέλτιστη.

Απόδειξη. Εστω π μια αυθαίρετη πολιτική, και έστω η τροχιά του συστήματος που αντιστοιχεί στην π : $x_o = x, \alpha_o\, x_1\, \alpha_1\, x_2\, \alpha_2 \dots$, όπου $x_{t+1} = \overline{\alpha}_t = \pi(h_t)$.
Εχουμε λοιπόν ότι

$$\phi(x,\pi) \;=\; -\sup_{n\to\infty} \sum_{0}^{n-1} c(\overline{\alpha}_t)/n\,. \tag{10}$$

Αρκεί λοιπόν να αποδείξουμε ότι

$$\psi(x) \;\leq\; \phi(x,\pi^o)\,, \tag{11}$$

διότι τότε, δεδομένου ότι η π^o είναι αυθαίρετη, η (11) συνεπάγεται την (6).

Εχουμε λοιπόν

$$c(\overline{\alpha}_o) + d(\overline{\alpha}_o) \overset{(5)}{\geq} \psi(x) + \delta(x)\,.$$

Επίσης

$$\begin{aligned} c(\overline{\alpha}_o) + c(\overline{\alpha}_1) + d(\overline{\alpha}_1) \quad & \overset{(5)}{\geq} c(\overline{\alpha}_o) + \psi(\overline{\alpha}_o) + d(\overline{\alpha}_o) \\ (\alpha_0 \in D(x)) \quad & \overset{(4)}{\geq} \psi(x) + c(\overline{\alpha}_o) + d(\overline{\alpha}_o) \\ & \overset{(5)}{\geq} \psi(x) + \psi(x) + d(x) \\ & = 2\psi(x) + d(x)\,, \end{aligned}$$

δηλαδή

$$2\psi(x) + d(x) \;\leq\; c(\overline{\alpha}_o) + c(\overline{\alpha}_1) + d(\overline{\alpha}_1) \;=\; c(\overline{\alpha}_o) + c(\overline{\alpha}_1) + d(x_2)\,. \tag{12}$$

Ανάλογα είναι εύκολο να αποδείξουμε ότι

$$n\psi(x) + d(x) \;\leq\; \sum_{t=0}^{n-1} c(\overline{\alpha}_t) + d(x_n)\,, \tag{13}$$

και απο την (13) έπεται ότι

$$\psi(x) \;\leq\; \frac{1}{n} \sum_{t=0}^{n-1} c(\overline{\alpha}_t) \;+\; \frac{(d(x_n) - d(x))}{n}\,. \tag{14}$$

Τώρα από την (14) και την υπόθεση ότι $|d(x)| \;\leq\; M_2 \;\; \forall x$, έπεται η (11).

(2) : Από τις (7) και (8) παίρνουμε με ανάλογο τρόπο ότι

$$n\psi(x) + d(x) \;\geq\; \sum_{t=0}^{n-1} c(\overline{\pi}^o(x_t)) + d(x_n) - n\epsilon \tag{15}$$

και η (15) συνεπάγεται ότι

$$\psi(x) \;\geq\; \phi(x,\pi^o) - \epsilon\,, \tag{16}$$

και η απόδειξη είναι πλήρης, δεδομένου του μέρους 1 του θεωρήματος, και του γεγονότος ότι, εξ ορισμού, $\phi^*(x) \le \phi(x,\pi)$, $\forall x \in S$, $\forall \pi$.

(3) : Το μέρος (3) είναι άμεση συνέπεια των κομματιών (1) και (2) του θεωρήματος.

Παρατηρήσεις. ***1.*** *Είναι δυνατό να δειχθεί ότι οι εξισώσεις βελτιστοποίησης είναι ικανές για να είναι μια πολιτική βέλτιστη ως προς το πιο ισχυρό κριτήριο, όταν μια πολιτική είναι βέλτιστη ως προς το κριτήριο του εκπτωτικού κόστους για κάθε τιμή του εκπτωτικού παράγοντα* b *σε μια περιοχή της μορφής* $(b_0,1)$.

2. *Στη περίπτωση πεπερασμένων χώρων καταστάσεων–αποφάσεων, οι συνθήκες του Θεωρήματος 1 ισχύουν για προφανείς λόγους.*

7. Ασκήσεις.

Ασκηση 1. *Δίνεται το δίκτυο του σχήματος 1 , όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι*

α) *Το συνολικό εκπτωτικό κόστος*

$$\sum_{t=0} \beta^t c(x_t,\alpha_t), \ \text{με} \ \beta = 0.5 ,$$

β) *Το μέσο κόστος*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{t=0}^{n} c(x_t,\alpha_t) ,$$

όταν $x_0 = 1$.

Σχήμα 1

Ασκηση 2. *Θεωρείστε την παρακάτω παραλλαγή του αλγόριθμου διαδοχικών προσεγγίσεων για τη λύση των εξισώσεων βελτιστοποίησης στο πρόβλημα άπειρου ορίζοντα με πεπερασμένους χώρους καταστάσεων και αποφάσεων.*

Εφαρμόζουμε τη διαδικασία διαδοχικών προσεγγίσεων στο χώρο τιμών για N *συνεχόμενα βήματα. Μετά χρησιμοποιούμε τις τιμές των* $\alpha^{(N)}(x)$ *που έχουμε βρει και μ' αυτές εφαρμόζουμε μια φορά το τμήμα υπολογισμού των* w *της διαδικασίας διαδοχικών προσεγγίσεων στο χώρο των πολιτικών. Σταματάμε αν ικανοποιούνται οι συνθήκες αυτής της μεθόδου για τα* Δw. *Διαφορετικά επαναλαμβάνουμε όλη την παραπάνω μέθοδο, με αρχική τιμή για τα* $v^{(N+1)}$ *τις τιμές των* w *που έχουμε υπολογίσει από τις διαδοχικές προσεγγίσειsν στο χώρο των πολιτικών. Εξηγήστε γιατί η μέθοδος που περιγράψαμε συγκλίνει στη λύση των εξισώσεων βελτιστοποίησης σε πεπερασμένο αριθμό βημάτων.*

Ασκηση 3. *Θεωρούμε την παρακάτω εξίσωση βελτιστοποίησης*

$$v(x) = \max_{0 \le \alpha \le x} \left\{ g(\alpha) + h(x-\alpha) + v(p\alpha + q(x-\alpha)) \right\} ,$$

όπου κάνουμε τις εξής παραδοχές για την ασυμπτωτική συμπεριφορά των g *και* h

α) $g(\alpha) \sim c_1\alpha^d$, $h(\alpha) \sim c_2\alpha^d$, $c_1,c_2,d > 0$, *όταν* $\alpha \to \infty$.

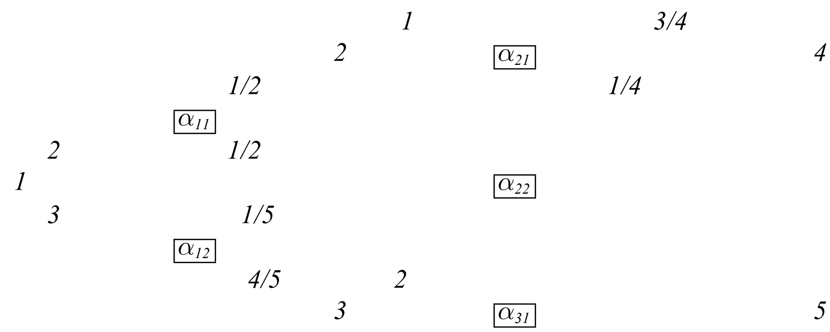
β) $g(\alpha) \sim c_1\alpha^{d_1}$, $h(\alpha) \sim c_2\alpha^{d_2}$, $c_1,c_2,d_1,d_2 > 0$, *όταν* $\alpha \to \infty$.

Και στις δύο περιπτώσεις βρείτε την ασυμπτωτική συμπεριφορά της $v(x)$, *όταν* $x \to \infty$.

ΚΕΦΑΛΑΙΟ 3
ΠΡΟΒΛΗΜΑΤΑ ΜΕ ΠΙΘΑΝΟΘΕΩΡΗΤΙΚΗ ΔΥΝΑΜΙΚΗ

1. Προβλήματα Πεπερασμένου Ορίζοντα.

1.1. Εισαγωγικό πρόβλημα. *Ας θεωρήσουμε το πρόβλημα πο`υ σχηματικά δίδεται απ`ο την εικόνα 1.*



Εικόνα 1

Η παραπάνω εικόνα είναι ανάλογη της εικόνας 1.5 , με την διαφορά ότι στην εικόνα 1 έχουμε προσθέσει τους κόμβους τυχαιότητας α_{xk} και τα διακεκομένα βέλη : $- - - >$. Η παραπάνω εικόνα περιγράφει την κίνηση μέσα στον χρόνο (ή χώρο) ενός συστήματος ως εξής:

α) *Τα σύνολα των κόμβων $S_0 = \{1\}$, $S_1 = \{2,3\}$, $S_2 = \{4,5\}$ συμβολίζουν τις δυνατές καταστάσεις του συστήματος κατά τις χρονικές στιγμές $t = 0,1,2$. Το σύνολο S_o είναι το σύνολο των αρχικών κόμβων, ενώ το σύνολο S_2 είναι το σύνολο των τερματικών κόμβων (καταστάσεων) . Τα σύνολα D_t $0 \le t \le 1$ συμβολίζουν τις δυνατές αποφάσεις στις διάφορες χρονικές περιόδους. Οι δυνατές αποφάσεις στον κόμβο $x \in S_t$ συμβολίζονται με $D(x) \subset D_t$. Ετσι έχουμε*

$$D(1) = \{\alpha_{11},\alpha_{12}\} ,\ D(2) = \{\alpha_{21},\alpha_{22}\} ,\ D(3) = \{\alpha_{31}\} .$$

β) *Η δυναμική του συστήματος ορίζεται μέσω γνωστών συναρτήσεων κατανομών πιθανότητας ως εξής : Αν την χρονική στιγμή t το σύστημα βρίσκεται στην κατάσταση $x \in S_t$ και εφαρμόσουμε την απόφαση $\alpha \in D(x) \subset D_t$, τότε κατά την επόμενη χρονική στιγμή το σύστημα θα βρίσκεται στην κατάσταση $y \in S_{t+1}$, με πιθανότητα $p(y/x,\alpha)$, όπου*

$$\sum_{y \in S_{t+1}} p(y/x,\alpha) = 1 \quad \forall\, x \in S_t ,\ \alpha \in D(x) \tag{1}$$

Δηλαδή, για κάθε κατάσταση $x \in S_t$ και για κάθε απόφαση $\alpha \in D(x)$, υπάρχει μια γνωστή συνάρτηση κατανομής (μάζα πιθανότητας στο παράδειγμά μας) $p(\cdot/x,\alpha)$ σύμφωνα με την οποία θα επιλεγεί η κατάσταση του συστήματος κατά την επόμενη χρονική στιγμή. Η δυναμική του συστήματος, λοιπόν, ορίζεται απο το σύνολο των συναρτήσεων

$$\Delta = \{p(\cdot/x,\alpha) ,\ x \in S_o \cup S_1 ,\ \alpha \in D(x)\} \tag{2}$$

Στο συγκεκριμένο παράδειγμα έχουμε

$$\Delta = \{p(\cdot/1,\alpha_{11}) , p(\cdot/1,\alpha_{12}) , p(\cdot/2,\alpha_{21}) , p(\cdot/2,\alpha_{22}) , p(\cdot/3,\alpha_{31})\}$$

όπου π.χ.

$$p(y/1,\alpha_{12}) = \begin{cases} 1/3 & y = 2 \\ 2/3 & y = 3 \end{cases}$$

Στην εικόνα 3.1 οι συναρτήσεις $p(\cdot/x,\alpha)$ ορίζονται απο τα διακεκομμένα βέλη και τους αριθμούς δίπλα σε αυτά.

γ) *Ανάλογα με την περίπτωση του προβλήματος με προσδιοριστική κίνηση, ορίζουμε*

i) την συνάρτηση τρέχοντος κόστους

$$c(x,\alpha)\,,\quad x = 0,1,\dots \quad \alpha \in D(x)\,, \tag{3}$$

και

ii) την συνάρτηση τερματικού κόστους

$$\hat{c}(x)\,,\quad x \in S_2\,. \tag{4}$$

1.2 Πολιτικές–Διαδρομές–Υπολογισμός μέτρου απόδοσης. *Ο πιο απλός τύπος πολιτικής (στρατηγικής) είναι ανάλογος με εκείνο που έχουμε θεωρήσει για το πρόβλημα με προσδιοριστική κίνηση. Ετσι, μια απλή πολιτική είναι ένας πλήρης κανόνας που αντιστοιχεί σε κάθε κατάσταση μια μοναδική απόφαση. Εστω λοιπόν π μια απλή πολιτική, π είναι μια συνάρτηση απο το σύνολο $S = S_o \cup S_1$ στο σύνολο $D_o \cup D_1$, όπου $\pi(x)$ $(\pi(x) \in D(x))$ συμβολίζει την απόφαση στην κατάσταση x.*

Η κύρια διαφορά του προβλήματος με πιθανοθεωρητική δυναμική απο εκείνο με προσδιοριστική δυναμική είναι ότι ενώ στο δεύτερο οι απλές πολιτικές προσδιορίζουν (μονοσήμαντα) τις διαδρομές του συστήματος, στο πρώτο οι διαδρομές είναι τυχαίες μεταβλητές και οι απλές πολιτικές (και όπως θα δούμε αργότερα όλες οι πολιτικές) προσδιορίζουν (μονοσήμαντα) τις κατανομές των διαδρομών.

Για παράδειγμα, ας θεωρήσουμε την απλή πολιτική π^0, που ορίζεται ως εξής : $\boldsymbol{\pi^0(1) = \alpha_{11}}$, $\boldsymbol{\pi^0(2) = \alpha_{21}}$, $\boldsymbol{\pi^0(3) = \alpha_{31}}$. *Η π^0 ορίζει τη στοχαστική διαδικασία $L(\pi^0) = X_0\, A_0\, X_1\, A_1\, X_2$ ως εξής*

(1) Οι δυνατές τιμές της $L(\pi^0)$, δηλαδή οι δειγματικές διαδρομές, είναι οι εξής

$$l_1 = 1\ \alpha_{11}\ 2\ \alpha_{21}\ 4$$

$$l_2 = 1\ \alpha_{11}\ 2\ \alpha_{21}\ 5$$

$$l_3 = 1\ \alpha_{11}\ 3\ \alpha_{31}\ 5$$

(2) Οι πιθανότητες, κάτω από την π^0, που αντιστοιχούν στις δυνατές δειγματικές διαδρομές βρίσκονται ως εξής

$$p(l_1,\pi^0) = P(X_0\, A_0\, X_1\, A_1\, X_2 = 1\ \alpha_{11}\ 2\ \alpha_{21}\ 4) =$$

$$\overset{(1)}{=} P(X_2{=}4/X_0\, A_0\, X_1\, A_1{=}1\ \alpha_{11}\ 2\ \alpha_{21}) \times P(X_0\, A_0\, X_1\, A_1{=}1\ \alpha_{11}\ 2\ \alpha_{21}) =$$

$$= P(X_2{=}4/X_0\, A_0\, X_1\, A_1{=}1\ \alpha_{11}\ 2\ \alpha_{21}) \times P(A_1{=}\alpha_{21}/X_0\, A_0\, X_1{=}1\ \alpha_{11}\ 2) \times$$

$$\times\ P(X_1 = 2\,/\,X_0\, A_0 = 1\ \alpha_{11}) \times P(A_0{=}\alpha_{11}\,/\,X_0 = 1) \times P(X_0{=}1) =$$

$$= P(X_2 = 4\,/\,X_1 = 2\,,\, A_1{=}\,\alpha_{21}) \times 1 \times P(X_1{=}\,2\,/X_0{=}1\,,\, A_0{=}\alpha_{11}) \times 1 \times 1 =$$

$$= \frac{3}{4} \times \frac{1}{2} = \frac{3}{8} .$$

Ανάλογα βρίσκουμε ότι

$$p(l_2,\pi^0) = P(X_2{=}5 / X_1{=}2 , A_1{=}\alpha_{21}) \times P(X_1{=}2 / X_0{=}1 , A_0{=}\alpha_{11}) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8},$$

$$p(l_3,\pi^0) = P(X_2{=}5 / X_1{=}3 , A_1{=}\alpha_{31}) \times P(X_1{=}3 /X_0{=}1 , A_0{=}\alpha_{11}) = 1 \times \frac{1}{2} = \frac{1}{2}.$$

Παρατηρήσεις. ***1.*** *Το πρώτο ανάπτυγμα της πιθανότητας* $p(l_1,\pi^0)$ *ισχύει κάτω από τις πιο γενικές συνθήκες για τη δυναμική του συστήματος και την πολιτική ελέγχου, δηλαδή όταν η κίνηση του συστήματος καθορίζεται από πιθανότητες της μορφής*

$$P(X_t{=}x_t / X_0\, A_0 \ldots X_{t-1}\, A_{t-1})$$

και οι αποφάσεις παίρνονται σύμφωνα με πιθανότητες της μορφής

$$P(A_t{=}\alpha_t / X_0\, A_0 \ldots X_{t-1}\, A_{t-1}\, X_t) .$$

2. *Στο δεύτερο ανάπτυγμα χρησιμοποιούμε το γεγονός ότι η δυναμική του παραδείγματός μας, όπως περιγράφεται από το σχήμα της εικ. 3.1 , είναι Μαρκοβιανή, και το ότι η πολιτική* π^0 *είναι απλή, δηλαδή:* $P(X_t{=}x_t / X_0\, A_0 \ldots X_{t-1}\, A_{t-1}) = P(X_t{=}x_t / X_{t-1}\, A_{t-1})$ *, και* $P(A_t{=}\alpha_t / X_0\, A_0 \ldots X_{t-1}\, A_{t-1}\, X_t) = 1$ *ή* 0 *, ανάλογα με το αν* $A_t = \pi^0(X_t)$ *ή* $A_t \neq \pi^0(X_t)$ *.*

3. *Γι`α να είμαστε ακριβείς στο συμβολισμό μας, είναι σωστό να γράφουμε* $X_0^\pi\ A_0^\pi \ldots X_{t-1}^\pi\ A_{t-1}^\pi$ *για τη στοχαστική διαδικασία που αντιστοιχεί στην πολιτική* π *. Για απλότητα του συμβολισμού όμως, θα παραλείπουμε το δείκτη* π *όταν δεν υπάρχει φόβος σύγχυσης.*

4. *Εστω* Λ *το σύνολο όλων των δυνατών διαδρομών, ή, ισοδύναμα, το σύνολο όλων των δυνατών τιμών (δειγματικών διαδρομών) των στοχαστικών διαδικασιών* $X_0^\pi\ A_0^\pi \ldots X_t^\pi\, A_t^\pi , \ldots$ *για όλες τις δυνατές πολιτικές* π *. Στο πρόβλημα με προσδιοριστική δυναμική μια απλή πολιτική όριζε μονοσήμαντα μια διαδρομή πάνω στο* Λ *. Στο πρόβλημα με πιθανοθεωρητική δυναμική μια απλή πολιτική ορίζει μονοσήμαντα μια διαδικασία* $X_0^\pi\ A_0^\pi \ldots X_t^\pi\, A_t^\pi , \ldots$ *, δηλαδή μια κατανομή πιθανότητας στο σύνολο* Λ*. Για παράδειγμα, η πολιτική* π^0 *ορίζει μονοσήμαντα τις πιθανότητες* $p(l_1,\pi^0) = 3/8$ *,* $p(l_2,\pi^0) = 1/8$ *,* $p(l_3,\pi^0) = 1/2$ *(παρατηρούμε ότι* $\sum_{l\in\Lambda} p(l,\pi^0) = 1$*) .*

Η ποσότητα που μας ενδιαφέρει πρωταρχικά είναι το μέτρο απόδοσης των διαφόρων πολιτικών. Ας υποθέσουμε ότι η ***αξία*** *της διαδρομής* $l = x_o\ \alpha_o\ x_1\ \alpha_1\ x_2$
ορίζεται όπως και στο πρόβλημα με προσδιοριστική κίνηση

$$w(l) = c(x_o,\alpha_o) + c(x_1,\alpha_1) + \hat{c}(x_2) . \qquad (1)$$

Ετσι βρίσκουμε ότι

$$\begin{aligned} w (l_1) &= c (1, \alpha_{11}) + c (2, \alpha_{21}) + \hat{c} (4) = 2 + 1 + 10 = 13 \\ w (l_2) &= c (1, \alpha_{11}) + c (2, \alpha_{21}) + \hat{c} (5) = 2 + 1 + 1 = 4 \\ w (l_4) &= c (1, \alpha_{11}) + c (3, \alpha_{31}) + \hat{c} (5) = 2 + 2 + 1 = 5 \end{aligned}$$

Το νόημα της ισότητας $w(l_1) = 13$ *είναι ότι, αν είχαμε τη δυνατότητα να παρατηρήσουμε την κίνηση του συστήματος μέσα στο χρονικό ορίζοντα που μας ενδιαφέρει (δηλαδή στο χρονικό διάστημα* $\{0,1,2\}$ *), και αν η διαδρομή που* ***έτυχε*** *να πραγματοποιηθεί ήταν η* l_1 *, τότε το κόστος που θα αντιστοιχούσε στην κίνηση αυτή του συστήματος θα ήταν ίσο με 13 μονάδες. Ολη η ιδέα όμως του προβλήματος με πιθανοθεωρητική δυναμική είναι ότι θέλουμε να επιδράσουμε πάνω στην κίνηση του συστήματος από την αρχή και κατά τη διάρκεια του χρονικού ορίζοντα, χωρίς να έχουμε γνώση της μελλοντικής του εξέλιξης (προσδιοριστική δυμανική).*

Στη αρχή του χρονικού ορίζοντα και δεδομένου ότι έχουμε αποφασίσει να χρησιμοποιήσουμε την πολιτική π^0, η πληροφορία που διαθέτουμε είναι ότι το μέτρο απόδοσης του συστήματος θα είναι μια τυχαία μεταβλητή $W(\pi^0,x_0)$, όπου

$$W(\pi^0,x_0) = W(L_{\pi^0}) = W(X_0^{\pi^0} A_0^{\pi^0} X_1^{\pi^0} A_1^{\pi^0} X_2^{\pi^0}) =$$

$$= c(X_0^{\pi^0} A_0^{\pi^0}) + c(X_1^{\pi^0} A_1^{\pi^0}) + \hat{c}(X_2^{\pi^0}) .$$

Εχουμε δει ότι οι δυνατές τιμές της διαδικασίας L_{π^0} είναι οι διαδρομές l_1, l_2, l_3, και έχουμε υπολογίσει τις πιθανότητες $P(L_{\pi^0}=l_i)$. Είναι εύκολο λοιπόν, να δούμε ότι η τυχαία μεταβλητή $W(L_{\pi^0})$ μπορεί να περιγραφεί ως εξής

$$W(L_{\pi^0}) = \begin{cases} w(l_1) = 13 , & \text{με πιθανότητα } P(L_\pi=l_1) = 3/8 , \\ w(l_2) = 4 , & \text{με πιθανότητα } P(L_\pi=l_2) = 1/8 , \\ w(l_3) = 5 , & \text{με πιθανότητα } P(L_\pi=l_3) = 1/2 . \end{cases}$$

Στην αρχή λοιπόν του χρονικού ορίζοντα, γνωρίζοντας την τυχαία μεταβλητή $W(x_0,\pi^0) = W(L_{\pi^0})$, μπορούμε να θεωρήσουμε σαν απλό μέτρο απόδοσης της π^0 την αναμενόμενη τιμή

$$\overline{w}(\pi^0) = 13 \times \tfrac{3}{8} + 4 \times \tfrac{1}{8} + 5 \times \tfrac{5}{8} = \tfrac{63}{8} .$$

Παρατηρούμε ότι ο παραπάνω συλλογισμός ισχύει και κατά τη διάρκεια της κίνησης του συστήματος. Για παράδειγμα, αν γνωρίζουμε ότι κατά την χρονική στιγμή $t=1$, το σύστημα βρίσκεται στην κατάσταση $X_1 = 2$, τότε το αναμενόμενο κόστος μέχρι το τέλος του ορίζοντα, κάτω από την πολιτική π^0, είναι εύκολο να υπολογιστεί. Πραγματικά, κάτω από την π^0, οι δυνατές διαδρομές με αρχικό κόμβο τον 2 και οι αντίστοιχες πιθανότητες είναι

$$l_1^{(1)} = 2\ \alpha_{21}\ 4 , \quad \text{με πιθανότητα } 3/4 ,$$

$$l_2^{(1)} = 2\ \alpha_{21}\ 5 , \quad \text{με πιθανότητα } 1/4 .$$

Ετσι έχουμε ότι

$$W(\pi^0,2) = \begin{cases} w(l_1^{(1)}) = 1+10 = 11 , & \text{με πιθανότητα } 3/4 , \\ \\ w(l_2^{(1)}) = 1+1 = 2 , & \text{με πιθανότητα } 1/4 , \end{cases}$$

και τελικά, $\overline{w}(\pi^0,2) = 11 \times 3/4 + 2 \times 1/4 = 35/4$.

Σαν δεύτερο παράδειγμα, αν γνωρίζουμε ότι κατά την χρονική στιγμή $t=1$, $X_1 = 3$, τότε είναι εύκολο να υπολογίσουμε την τιμή: $\overline{w}(3,\pi^0) = 2+1 = 3$. Τέλος, με ανάλογο τρόπο είναι εύκολο να δούμε ότι

$$\overline{w}(\pi^0,4) = 10 \quad \text{και} \quad \overline{w}(\pi^0,5) = 1 .$$

Η αξία του παραπάνω τρόπου υπολογισμού της αναμενόμενης τιμής $\overline{w}(\pi^0,x)$ βρίσκεται στο ότι μας βοηθά να κατανοήσουμε τη φύση της δυναμικής και της αναμενόμενης τιμής του κόστους σαν μέτρο απόδοσης του συστήματος, όταν η δυναμική είναι πιθανοθεωρητική. Σε υπολογισμούς όμως, καθώς και για την ανάπτυξη της θεωρίας υπολογισμού βέλτιστων πολιτικών, υπάρχει μια δεύτερη μέθοδος υπολογισμού, που είναι πολύ πιο εύχρηστη. Η μέθοδος αυτή βασίζεται στις υπό συνθήκη πιθανότητες και αναμενόμενες τιμές, και έχει ως εξής

Κατ' αρχήν, δεδομένης μιας στοχαστικής διαδρομής

$$L_{\pi^0} \ = \ X_0\, A_0\, X_1\, A_1\, X_2 \ \ldots \ ,\ \text{ορίζουμε τις διαδρομές}$$

$$L^{(1)}_{\pi^0} \ = \ X_1\, A_1\, X_2 \ \ldots \ , \quad L^{(2)}_{\pi^0} \ = \ X_2 \ \ldots \ ,\ \text{κ.λ.π.}$$

Εχουμε λοιπόν

$$\overline{w}(\pi^0,1) \ = \ E(W(L_{\pi^0})) \ =$$

$$= \ E(\ W(L_{\pi^0}) \,/\, X_1{=}2\ ) \ \times \ P(X_1{=}2 \,/\, X_0{=}1 \ ,\ \pi(1){=}\alpha_{11}) \ +$$

$$+ \ E(\ W(L_{\pi^0}) \,/\, X_1{=}3\ ) \ \times \ P(X_1{=}3 \,/\, X_0{=}1 \ ,\ \pi(1){=}\alpha_{11})\ , \tag{2}$$

αλλά

$$E(W(L_{\pi^0}) \,/\, X_1{=}2) \ = \ c(1,\alpha_{11}) \ + \ E(W(L^{(1)}_{\pi^0}) \,/\, X_1{=}2)\ , \tag{3}$$

και

$$E(W(L_{\pi^0}) \,/\, X_1{=}3) \ = \ c(1,\alpha_{11}) \ + \ E(W(L^{(1)}_{\pi^0}) \,/\, X_1{=}3)\ , \tag{4}$$

Ανάλογα έχουμε ότι

$$E(W(L^{(1)}_{\pi^0}) \,/\, X_1{=}2) \ = \ \overline{w}(2,\pi^0) \ =$$

$$= \ E(W(L^{(1)}_{\pi^0}) \,/\, X_2{=}4 \ ,\ X_1{=}2) \ \times \ P(X_2{=}4 \,/\, X_1{=}2 \ ,\ \pi^0(X_1){=}\alpha_{21}) \ +$$

$$+ \ E(W(L^{(1)}_{\pi^0}) \,/\, X_2{=}5 \ ,\ X_1{=}2) \ \times \ P(X_2{=}5 \,/\, X_1{=}2 \ ,\ \pi^0(X_1){=}\alpha_{21})\ . \tag{5}$$

Τώρα

$$E(w(L^{(1)}_{\pi^0}) \,/\, X_2{=}x_2\ ,\ \ X_1{=}x_1) \ = \ c(x_1,\pi^0(x_1)) + E(W(L^{(2)}_{\pi^0}) \,/\, X_2{=}x_2) \ =$$

$$= \ c(x_1,\pi^0(x_1)) \ + \ \overline{w}(\pi,x_2) \ \ = \ \begin{cases} 1+10 \ = \ 11\ , & \text{αν } \ x_1 \ = \ 2\ ,\ x_2 \ = \ 4 \\ 1+1 \ = \ 2\ , & \text{αν } \ x_1 \ = \ 2\ ,\ x_2 \ = \ 5 \\ 2+1 \ = \ 3\ , & \text{αν } \ x_1 \ = \ 3\ ,\ x_2 \ = \ 5 \end{cases} \tag{6}$$

Ετσι, από την (5) βρίσκουμε ότι

$$E(W(L^{(1)}_{\pi^0}) \,/\, X_1{=}2) \ = \ 11 \ \times \ \tfrac{3}{4} \ + \ 2 \ \times \ \tfrac{1}{4} \ = \ \tfrac{35}{4}\ , \tag{7}$$

$$E(W(L^{(1)}_{\pi^0}) \,/\, X_1{=}3) \ = \ 12 \ \times \ 0 \ + \ 3 \ \times \ 1 \ = \ 3\ . \tag{8}$$

Από τις (2) , (3) , (4) , χρησιμοποιώντας τις (7) , (8) βρίσκουμε

$$\overline{w}(2,\pi^0) \ = \ (2+\tfrac{35}{4}) \ \times \ \tfrac{1}{2} \ + \ (2+3) \ \times \ \tfrac{1}{2} \ = \ \tfrac{63}{8}\ . \tag{9}$$

*Παρατηρούμε λοιπόν ότι, για να υπολογίσουμε το $\overline{w}(\pi^0,2)$ με την δεύτερη μέθοδο, αρκεί να κάνουμε τους υπολογισμούς (6) , (7) , (8) , (9) , δηλαδή να φτάσουμε στον αρχικό κόμβο (κατάσταση) πηγαίνοντας **προς τα πίσω.***

1.3. Λύση του Εισαγωγικού Προβλήματος.

Μέθοδος 1. Υπολογισμός της απόδοσης όλων των απλών πολιτικών. *Ας θεωρήσουμε την πολιτική π^1, όπου $\pi^1(1) = \alpha_{12}$, $\pi^1(2) = \alpha_{21}$ και $\pi^1(3) = \alpha_{31}$. Επαναλαμβάνοντας τους παραπάνω υπολογισμούς βρίσκουμε*

$$u^{(2)}(\pi^1,4) := E(\,W(L^{(3)}_{\pi^1}) \,/\, X_2 = 4\,) = \hat{c}(4) = 10\,,$$

$$u^{(2)}(\pi^1,5) := E(\,W(L^{(3)}_{\pi^1}) \,/\, X_2 = 5\,) = \hat{c}(5) = 1\,,$$

$$u^{(1)}(\pi^1,2) := E(\,W(L^{(2)}_{\pi^1}) \,/\, X_1 = 2\,) = 1 + 10 \times \tfrac{3}{4} + 1 \times \tfrac{1}{4} = \tfrac{35}{4}\,,$$

$$u^{(1)}(\pi^1,3) := E(\,W(L^{(2)}_{\pi^1}) \,/\, X_1 = 3\,) = 2 + 10 \times 0 + 1 \times 1 = 3\,,$$

$$u^{(0)}(\pi^1,1) := \overline{w}(\pi^1,1) = 3 + \tfrac{35}{4} \times \tfrac{1}{5} + 3 \times \tfrac{4}{5} = \tfrac{143}{20}\,.$$

Παρατηρούμε ότι $\overline{w}(\pi^1,1) < \overline{w}(\pi^0,1)$.

Αλλες απλές πολιτικές είναι οι π^2, π^3, *όπου* $\pi^2(1) = \alpha_{11}$, $\pi^2(2) = \alpha_{22}$, $\pi^2(3) = \alpha_{31}$ *και* $\pi^3(1) = \alpha_{12}$, $\pi^3(2) = \alpha_{22}$, $\pi^3(3) = \alpha_{31}$. *Ανάλογα είναι εύκολο να υπολογίσουμε τις αξίες των πολιτικών* π^2, π^3 *για την αρχική κατάσταση* $x_o = 1$, *και θα βρούμε ότι η* π^1 *είναι η βέλτιστη μεταξύ τους.*

Μέθοδος 2. Εξίσωση Βελτιστοποίησης. *Ο παραπάνω τρόπος λύσης του εισαγωγικού προβλήματος είναι χρήσιμος για την κατανόηση του προβλήματος, αλλά η αξία του στη πράξη είναι περιορισμένη, λόγω του ότι απαιτεί την απαρίθμηση των αποδόσεων όλων των δυνατών πολιτικών. Είναι όμως δυνατόν να αποδείξουμε ότι για κατάλληλα ορισμένη συνάρτηση τιμών* $v(\cdot)$, *ισχύει μια εξίσωση βελτιστοποίησης ανάλογη με εκείνη του προβλήματος με προσδιοριστική δυναμική. Βέλτιστες πολιτικές είναι εκείνες για τις οποίες οι αποδόσεις τους ικανοποιούν την εξίσωση βελτιστοποίησης. Πράγματι, ας θεωρήσουμε τα υπο-προβλήματα* $P(k,x_k)$

$$(min) \qquad E\Big(\sum_{t=k}^{1} c(x_t,\alpha_t) + \hat{c}(x_2)\Big)$$

Δυναμική : $\Delta = \{p(\cdot/x_k,\alpha_k)\ \ k \in X_k\,,\ \alpha_k \in A(x_k)\,,\ k = 0,1\}$ *(1)*

Αρχική κατάσταση $x_k \in S_k$.

Εστω $v(k,x_k)$ *η ελάχιστη τιμή (υποθέτουμε ότι υπάρχει) της αντικειμενικής συνάρτησης του* $P(k,x_k)$, $k = 0,1,2$, *όπου ορίζουμε αθροίσματα της μορφής* $\sum_{2}^{1}$ *να είναι ίσα με το* 0.

Εστω $\overline{w}(\pi,(k,x_k))$ *η αξία της πολιτικής* π. *Τότε*

$$v(k,x_k) = \min_{\pi}\{\,\overline{w}(\pi,(k,x_k))\,\}\,. \tag{2}$$

Θα αποδείξουμε στην επόμενη παράγραφο τα παρακάτω

1)
$$v(k,x_k) = \min_{\alpha \in A(x_k)} \Big\{ c(x_k,\alpha) + E(v(k+1,X_{k+1}) \,/\, X_k{=}x_k,\ \alpha) \Big\} \quad \forall x_k \in X_k\,,\ k = 0,1 \tag{3}$$

$$v(2,x_2) = \hat{c}(x_2)\,. \tag{4}$$

2) Η πολιτική π^0 *είναι βέλτιστη αν*

$$\overline{w}(\pi^0,(k,x_k)) = v(k,x_k)\,,\ \forall k = 0,1,2\,,\ x_k \in S_k\,.$$

Για το εισαγωγικό πρόβλημα έχουμε

$v(2,4) \ = \ \hat{c}(4) \ = \ 10 \,,$

$v(2,5) \ = \ \hat{c}(5) \ = \ 1 \,,$

$$v(1,2) \ = \ \min \{ 1 + E(v(2,X_2)/X_1{=}2 \,,\, \alpha{=}\alpha_{21}) \ , \ 5 + E(v(2,X_2)/X_1{=}2 \,,\, \alpha{=}\alpha_{22}) \}$$

$$= \min \begin{cases} 1 \ + \ (\, v(2,4) \ \times \ 3/4 \ + \ v(2,5) \ \times \ 1/4) \\ \\ 5 \ + \ (\, v(2,4) \ \times \ 1/2 \ + \ v(2,5) \ \times \ 1/2) \end{cases} =$$

$$= \min \begin{cases} 1 \ + \ (\, 10 \ \times \ 3/4 \ + \ 1 \ \times \ 1/4) \ = \ \frac{35}{4} \\ \\ 5 \ + \ (10 \ \times \ 1/2 \ + \ 1 \ \times \ 1/2) \ = \ \frac{21}{2} \end{cases} = \ \frac{35}{4} \ , \ \alpha^*(2,1) \ = \ \alpha_{21} \,,$$

$$v(1,3) \ = \ 2 + E(v(X_2) \,/\, X_1{=}3 \,,\, \alpha{=}\alpha_{31}) \ =$$

$$= \ 2 + v(2,5) \ \times \ 1 \ = \ 3 \,, \quad \alpha^*(3,1) \ = \ \alpha_{31} \,,$$

$$v(0,1) \ = \ \min \{ \ 2 + E(v(1,X_1)/X_0{=}1 \,,\, \alpha{=}\alpha_{11}) \ , \ 3 + E(v(1,X_1)/X_0{=}1 \,,\, \alpha{=}\alpha_{12}) \ \}$$

$$= \min \begin{cases} 2 \ + \ (v(1,2) \ \times \ 1/2 \ + \ v(1,3) \ \times \ 1/2) \\ \\ 3 \ + \ (v(1,2) \ \times \ 1/5 \ + \ v(1,3) \ \times \ 4/5) \end{cases} =$$

$$= \min \begin{cases} 2 \ + \ (\frac{35}{4} \ \times \ \frac{1}{2} \ + \ 3 \ \times \ \frac{1}{2}) \ = \ \frac{63}{8} \\ \\ 3 \ + \ (\frac{35}{4} \ \times \ \frac{1}{5} \ + \ 3 \ \times \ \frac{4}{5}) \ = \ \frac{143}{20} \end{cases} = \ \frac{143}{20} \ , \ \alpha^*(1,0) \ = \ \alpha_{12} \,.$$

Σύμφωνα, λοιπόν, με τα αποτελέσματα (2) , (1) παραπάνω, η βέλτιστη πολιτική είναι η

$$\pi^1 \ : \ \pi^1(1) \ = \ \alpha_{12} \,, \ \pi^1(2) \ = \ \alpha_{21} \,, \ \pi^1(3) \ = \ \alpha_{31} \,,$$

και η αξία της είναι $\bar{w}(\pi^1,1) \ = \ \frac{143}{20}$.

2. Διατύπωση Γενικών Προβλημάτων με Πιθανοθεωρητική Δυναμική.

2.1. Πρόβλημα Πεπερασμένου Ορίζοντα *Το γενικό πρόβλημα ορίζεται ως εξής :*

1. *Χώροι καταστάσεων και αποφάσεων* $S_{N+1}, S_t, D_t, \ t \ = \ 0,1,\ldots,N$.

2. *Δυναμική του συστήματος*

$$P(X_{t+1} \leq X \,/\, X_t{=}X_t \,,\, A_t{=}\alpha) \,, \qquad \alpha \in D_t(X_t) \,, \ x \in S_t \,, \ t = 0,1,\ldots,N \,.$$

3. *Συναρτήσεις τρέχοντος και τερματικού κόστους :* $c(x,\alpha) \,, \ \hat{c}(x)$.

4. *Αρχική κατάσταση* x_0 .

Μέτρο απόδοσης.

$$\overline{w}(\pi,X_0) \ = \ E(\sum_{t=0}^{N} c(X_t,A_t) \ + \ \hat{c}(X_{N+1}))\ .$$

2.2. Προβλήματα Απειρου Ορίζοντα με Στάσιμη Δομή Κόστους και Δυναμική.

1. Χώροι αποφάσεων, καταστάσεων S, D .

2. Δυναμική του συστήματος

$$P(X_{t+1} \leq x \,/\, X_t{=}X_t \,,\, A_t{=}\alpha_t)\ ,\ \ x \in S\ ,\ \ \alpha \in D(x)\ .$$

3. Συναρτήσεις τρέχοντος κόστους $c(x,\alpha)$.

4. Αρχική κατάσταση x_0 .

Μέτρα Απόδοσης.

$$\overline{w}_1(\pi,x_0) \ = \ E(\sum_{t=0}^{\infty}\beta^t c(X_t,A_t) \,/\, X_0 = x_0)\ , \quad \beta \in [0,1]\ ,$$

$$\phi(\pi,x_0) \ = \ \limsup_{N \to \infty} \frac{1}{N+1}\ E(\sum_{t=0}^{N} c(X_t,A_t) \,/\, X_0 = x_0\,)\ .$$

Παρατήρηση. *Οπως και στά προβλήματα με προσδιοριστική δυναμική, έτσι καί τώρα υπάρχουν αναγκαίες συνθήκες κάτω απ`ο τις οποίες οι συναρτήσεις τιμών είναι καλά ορισμένες. Το μόνο επιπλέον στοιχείο γιά τά προβλήματα με πιθανοθεωρητική δυναμική έχει σχέση με συνθήκες που χρειάζονται γιά την ύπαξη της αναμενόμενης τιμής* $E(\sum_{t=0}^{N} c(X_t,A_t) \,/\, X_0 = x_0\,)$ *. Η συζήτηση όμως των συνθηκών αυτών προυποθέτει επαρκή γνώση της Θεωρίας Ολοκλήροσης σε μετρίσημους χώρους. Το βιβλίο των Dynkin and Yushkevich (1979) περιέχει μία πλήρη συζήτηση πάνω σε αυτό το θέμα.*

3. Εξισώσεις Βελτιστοποίησης για τα Γενικά Προβλήματα.

Οπως και στα αντίστοιχα προβλήματα με προσδιοριστική δυναμική, η λύση των προβλημάτων με πιθανοθεωρητική δυναμική είναι δυνατό να βρεθεί μέσω ανάλογων εξισώσεων βελτιστοποίησης. Ας θεωρήσουμε πρώτα το πρόβλημα του πεπερασμένου ορίζοντα, και έστω $P(k,x)$ *τα υποπροβλήματα*

$$(min)\ \ \overline{w} \ = \ E(\,\sum_{t=k}^{N} c(X_t,A_t) + \ \hat{c}(X_{N+1})\,)\ ,$$

αρχική κατάσταση $X_k \ = \ x$,

κάτω απο την ίδια δυναμική.

Το πρώτο αποτέλεσμα είναι το εξής.

Πρόταση 1. *Ας υποθέσουμε ότι τα προβλήματα* $P(k,x)$ *έχουν βέλτιστες λύσεις για κάθε* $x \in S_k\ ,\ \ k = 0,1,\ldots,N+1$, *και έστω* $v(k,x)$ *η τιμή του* $P(k,x)$. *Τότε*

(1) Η συνάρτηση $v(t,x)$ *είναι η μόνη λύση της εξίσωσης*

$$v(t,x) \ = \ \min_{\alpha \in D_t(x)} \{\, c(x,\alpha) + E(v(t+1,X_{t+1}) \,/\, A_t{=}\alpha\ ,\ X_t{=}x)\,\}\ . \tag{1}$$

(2) Υπάρχει μια βέλτιστη πολιτική π^0 , *που είναι Μαρκοβιανή, και που ορίζεται μέσω της εξίσωσης*

$$c(x,\pi^0(t,x)) + E(v(t+1,X_{t+1}) \,/\, X_t{=}x \,,\, A_t{=}\pi^0(t,x)) \;=\; v(t,x)\,. \tag{2}$$

Απόδειξη. *Ας παρατηρήσουμε κατ' αρχή ότι*

$$v(N{+}1,x) \;=\; \hat{c}(x)\,. \tag{3}$$

Τώρα,

$$\begin{aligned} v(N,x) \quad &= \min_{\alpha \in D_N(x)} \{c(x,\alpha) + E(v(N+1,X_{N+1}) \,/\, A_N{=}\alpha \,,\, X_N{=}x)\} \\ &= \min_{\alpha \in D_N(x)} \{c(x,\alpha) + E(\hat{c}(N+1,X_{N+1}) \,/\, A_N{=}\alpha \,,\, X_N{=}x)\}\,, \end{aligned} \tag{4}$$

Επαγωγικά έχουμε

$$\begin{aligned} v(t,x) \quad &= \min_{\pi} \left\{E\left(\sum_{k=t}^{N+1} c(X_k,A_k)/X_t{=}x\right)\right\} \\ &= \min_{a,\,\pi'} \left\{c(x,\alpha) \;+\; E\left(\sum_{k=t+1}^{N+1} c(X_k,A_k)/X_t{=}x \,,\, A_t{=}\alpha\right)\right\} \\ &= \min_{a} \left\{c(x,\alpha) \;+\; \min_{\pi'} E\left(\sum_{k=t+1}^{N+1} c(X_k,A_k)/X_t{=}x \,,\, A_t{=}\alpha\right)\right\} \\ &= \min_{a} \left\{c(x,\alpha) \;+\; E(v(X_{t+1})/X_t{=}x \,,\, A_t{=}\alpha)\right\}\,, \end{aligned} \tag{5}$$

όπου στη (5) π, π' συμβολίζουν πολιτικές στα διαστήματα $[t, t+1,\ldots,N]$ καί $[t+1, t+2,\ldots,N]$, αντίστιχα. Επίσης, στη (5) έχουμε κάνει χρήση της εοαγωγικής υπόθεσης

$$E(v(X_{t+1})/X_t{=}x \,,\, A_t{=}\alpha)\Big\} \;=\; \min_{\pi'}\left\{E\left(\sum_{k=t+1}^{N+1} c(X_k,A_k)/\, X_t{=}x \,,\, A_t{=}\alpha\right)\right\}, \tag{6}$$

Για το πρόβλημα του άπειρου ορίζοντα με μέτρο απόδοσης $w_1(\pi,x_0)$, ας ορίσουμε τα υποπροβλήματα $P(x)$ ως προς την αρχική κατάσταση $X_0 = x$, την οποία θεωρούμε μεταβλητή στο S. Αν υποθέσουμε ότι τα προβλήματα $P(x)$ έχουν πεπερασμένη τιμή

$$v(x) \;=\; \min_{\pi} \{\, E(\sum_{t=0}^{\infty}\beta^t c(X_t^\pi,A_t^\pi) \,/\, X_0{=}x) \,\}\,,$$

τότε μπορούμε να αποδείξουμε την εξής

Πρόταση 2.

(1) Η συνάρτηση $v(x)$ είναι η μόνη λύση της εξίσωσης

$$v(x) \;= \min_{\alpha \in D(x)} \{\, c(x,\alpha) + \beta E(v(X_1) \,/\, X_0{=}x \,,\, A_0{=}\alpha) \,\}\,. \tag{7}$$

(2) Υπάρχει μια βέλτιστη πολιτική π^0, που είναι απλή, και που ορίζεται μέσω της εξίσωσης

$$c(x,\pi^0(x)) \;+\; \beta\, E(v(X_1) \,/\, X_0{=}x \,,\, A_0{=}\pi^0(x)) \;=\; v(x)\,. \tag{8}$$

Απόδειξη - Ασκηση

Για το πρόβλημα του άπειρου ορίζοντα με κριτήριο βελτιστοποίησης το $\phi(\pi,x)$, το μέσο αναμενόμενο κόστος ισχύει η εξής

Πρόταση 3. ***(1)*** *Εάν υπάρχουν συναρτήσεις $\psi(x)$, $d(x)$ ομοιόμορφα φραγμένες (δηλ. $|\psi(x)| \leq M_1$, $|d(x)| \leq M_2 \ \forall x$, M_1,M_2 σταθερές) που ικανοποιούν τις εξισώσεις*

$$\psi(x) = \inf_{\alpha \in D(x)} \{ E(\psi(X_1) / X_0=x ,\ A_0=\alpha) \} \tag{9}$$

$$\psi(x) + d(x) = \inf_{\alpha \in D(x)} \{ c(x,\alpha) + E(d(X_1) / X_0=x ,\ A_0=\alpha) \} , \tag{10}$$

τότε

$$\psi(x) \leq \phi^*(x) = \inf_{\pi} \{\phi(x,\pi)\} ,\ \forall x \in S . \tag{11}$$

(2) Εάν για μια απλή πολιτική π^0 ισχύουν οι ανισότητες

$$\psi(x) \geq E(\psi(X_1) / X_0=x ,\ A_0=\pi^0(x)) - \epsilon/2 \tag{12}$$

και

$$\psi(x) + d(x) \geq c(x,\pi^0(x)) + E(d(X_1) / X_0=x ,\ A_0=\pi^0(x)) - \epsilon/2 , \tag{13}$$

για κάποια σταθερά $\epsilon > 0$, τότε

$$\phi(x,\pi^0) \geq \phi^*(x) - \epsilon , \tag{14}$$

οπότε η π^0 είναι $\epsilon -$ βέλτιστη (δηλαδή $\phi(x,\pi^0) \in (\phi^*(x),\phi^*(x) + \epsilon)$).

(3) Ε`αν οι (5) και (6) παραπάνω ισχύουν με min στην θέση των inf και επιπλέον για κάθε κατάσταση $X_0 = x$ τα ελάχιστα επιτυγχάνονται για τις ίδιες αποφάσεις $\pi^0(x)$, τότε η πολιτική π^0 είναι βέλτιστη.

Παρατηρήσεις. ***1.*** *Οταν οι χώροι καταστάσεων S_t ή S είναι πεπερασμένοι, τότε η δυναμική του συστήματος είναι δυνατό να περιγραφεί μέσω των πιθανοτήτων μεταπήδησης*

$$P(y/x,\alpha) = P(X_{t+1} = y / X_t=x ,\ A_t=\alpha) .$$

Σε τούτη την περίπτωση έχουμε

$$E(g(X_{t+1}) / X_t=x ,\ A_t=\alpha) = \sum_{y \in S} g(y)\, p(y/x,\alpha) .$$

2. *Οταν οι χώροι καταστάσεων δεν είναι πεπερασμένοι, και όταν η δυναμική του συστήματος είναι δυνατό να περιγραφεί μέσω συναρτήσεων πυκνότητας πιθανότητας $f(y/x,\alpha)$, έτσι ώστε*

$$P(X_{t+1} \leq y / X_t=x ,\ A_t=\alpha) = \int_{-\infty}^{+\infty} f(y/x,\alpha)\, dy ,$$

τότε έχουμε

$$E(g(X_{t+1}) / X_t=x ,\ A_t=\alpha) = \int_{-\infty}^{+\infty} g(y)\, f(y/x,\alpha)\, dy .$$

3. *Σε πολλά προβλήματα είναι δυνατόν η πραγματική τιμή για το τρέχον κόστος να εξαρτάται και από την τιμή της X_{t+1}, δηλαδή να έχουμε $\widetilde{c}(X_t,A_t,X_{t+1})$. Είναι εύκολο* *(άσκηση) να αποδειχθεί ότι τούτη η περίπτωση ανάγεται στην περίπτωση που έχουμε ήδη εξετάσει αν ορίσουμε την νέα συνάρτηση τρέχοντος κόστους ως εξής*

$$c(x,\alpha) \;=\; E(\tilde{c}(x,\alpha,X_{t+1}) \,/\, X_t = x \;,\; A_t = \alpha) \;=\; \int_{-\infty}^{+\infty} \tilde{c}(x,\alpha,y)\, dP(y/x,\alpha)\;.$$

4. *Ανάλογα είναι δυνατό (π.χ. πρόβλημα 3.4.2) να έχουμε ότι η πραγματική τιμή του τρέχοντος κόστους εξαρτάται απο μια τυχαία μεταβλητή* Z_t , *με γνωστή κατανομή* $G(z)$: $\tilde{c}(x,\alpha,Z_t)$. *Τότε η νέα συνάρτηση κόστους θα είναι η*

$$c(x,\alpha) \;=\; E(\tilde{c}(x,\alpha,Z_t)) \;=\; \int_{-\infty}^{+\infty} \tilde{c}(x,\alpha,z)\, dG(z)\;.$$

5. *Οι μέθοδοι λύσεις των εξισώσεων βελτιστοποίησης είναι ανάλογες με εκείνες για το πρόβλημα με προσδιοριστική δυναμική.*

4. Εφαρμογές

4.1. Απλό Πρόβλημα Παραγωγής *Μια διαδικασία παράγει όμοια κομμάτια, και η πιθανότητα κάθε ένα από αυτά να είναι ελαττωματικό, είναι* $1/2$. *Η παραγωγή γίνεται κατά ομάδες, και ο αντικειμενικός σκοπός είναι να παράγουμε ένα όχι ελαττωματικό κομμάτι. Μας επιτρέπεται να παράγουμε διαδοχικά το πολύ 3 ομάδες, και εάν κανένα κομμάτι δεν είναι όχι-ελαττωματικό, τότε έχουμε ζημιά, η οποία υπολογίζεται σε 1.600 δρχ. Το κόστος παραγωγής μιας ομάδας που περιέχει α κομμάτια είναι* $c(\alpha) \;=\; 300\,\delta(\alpha) \,+\, 100\alpha$, *όπου :* $\delta(\alpha) = 1$, *αν* $\alpha \geq 1$, *και* $\delta(\alpha) = 0$ *αν* $\alpha = 0$. *Το πρόβλημα είναι να βρούμε τα επίπεδα παραγωγής* $\alpha_1^*, \alpha_2^*, \alpha_3^*$, *έτσι ώστε να ελαχιστοποιείται το αναμενόμενο κόστος παραγωγής/ ζημιάς.*

Διατύπωση του προβλήματος.

1. Χώροι καταστάσεων αποφάσεων. *Είναι εύκολο να δούμε ότι σε κάθε "χρονική στιγμή" t η κατάσταση του συστήματος περιγράφεται από τη μεταβλητή*

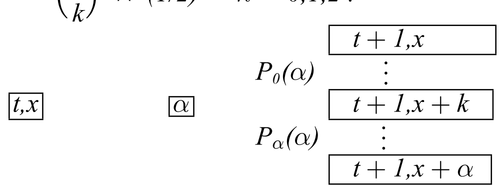
x_t: *ο αριθμός των όχι-ελαττωματικών κομματιών που έχουν παραχθεί μετά το τέλος της παραγωγής της t ομάδας ,* $t = 0,1,2,3$.

Εχουμε ότι $x_0 = 0$ *είναι η αρχική κατάσταση.*

Η απόφαση κατά την περίοδο t , α_t , *είναι ο αριθμός κομματιών που θα περιέχει η επόμενη ομάδα,* $\alpha_t \in \{0,1,2,\ldots\}$ **.**

2. Δυναμική του Συστήματος. *Η δυναμική περιγράφεται στο παραπάνω σχήμα, όπου* $P_k(\alpha)$ *συμβολίζει την πιθανότητα να έχουμε k όχι-ελαττωματικά κομμάτια σε μια ομάδα μεγέθους* α . *Εχουμε λοιπόν ότι*

$$P_k(\alpha) \;=\; \binom{q}{k} \;\times\; (1/2)^{\alpha} \quad \kappa = 0,1,2\;. \tag{1}$$



Δηλαδή, αν η παρούσα κατάσταση είναι $x_t = x$, *και αποφασίσουμε να παράγουμε μια ομάδα που περιέχει α κομμάτια, τότε η νέα κατάσταση* $x_{t+1} = x_{t+1}(\alpha)$ *είναι μια τυχαία μεταβλητή με τιμές* $(x+k)$, *στις οποίες αντιστοιχούν οι πιθανότητες* $P_k(\alpha)$, $k = 0,1,\ldots,\alpha$, $\alpha \geq 0$.

3. Συναρτήσεις κόστους.

α) Συνάρτηση τρέχοντος κόστους

$$c(x,\alpha) = \begin{cases} 1600\,, & \alpha\nu\ x = 0\ \ \kappa\alpha\iota\ \ \alpha = 0 \\ 300 + 100\alpha\,, & \alpha\nu\ \ x \geq 0\ \ \kappa\alpha\iota\ \ \alpha \geq 1\,. \end{cases} \tag{2}$$

β) Συνάρτηση τερματικού κόστους

$$\hat{c}(x) = \begin{cases} 1600\,, & \alpha\nu\ \ x = 0 \\ 0\,, & \alpha\nu\ \ x \geq 1 \end{cases} \tag{3}$$

4. Εξισώσεις Βελτιστοποίησης.

$$\begin{aligned} v(t,x) &= \min_{\alpha \geq 0} \{ 1600(1-\delta(x)) + 300\delta(\alpha) + 100\alpha + E(v(t+1,X_{t+1}) \,/\, X_t = x\,,\ A_t = \alpha) \} \\ &= \min_{\alpha \geq 0} \{ 1600(1-\delta(x)) + 300\delta(\alpha) + 100\alpha + \sum_{k=0}^{\alpha} v(t+1,x+k) \binom{\alpha}{k} (1/2)^{\alpha} \}\,, \\ & t = 0,1,2\,,\quad x \geq 0\,. \end{aligned} \tag{4.1}$$

$$v(3,x) = \hat{c}(x)\,. \tag{4.2}$$

**5. Λ`υση των Εξισ`ωσεων Βελτιστοπο`ιησης.** Παρατηρούμε κατ' αρχ`η ότι

$$v(3,x) = 0\,,\ \ \kappa\alpha\iota\ \ \alpha_3^* = \alpha^*(3,x) = 0\,,\ \ \alpha\nu\ \ x \geq 1\,, \tag{5}$$

και επαγωγικά είναι εύκολο να δούμε ότι

$$v(t,x) = 0\,,\ \ \kappa\alpha\iota\ \ \alpha_t^* = 0\,,\ \ \alpha\nu\ \ x \geq 1\,,\ \ \gamma\iota\alpha\ \ t = 0,1\,. \tag{6}$$

Αρκεί λοιπόν να υπολογίσουμε τις ποσότητες $v(t,0)$, $t = 0,1,2$. Εχουμε

$$\begin{aligned} v(2,0) &\overset{(5),(4)}{=} \min\{ 1600\,,\ 300 + 100\alpha + v(3,0)\,P_0(\alpha) \} \\ &\overset{(1),(4)}{=} \min\{ 1600\,,\ 300 + 100\alpha + 1600 \times (1/2)^{\alpha} \}\,. \end{aligned} \tag{7}$$

Εστω

$$h(\alpha) = 300 + 100\alpha + 1600 \times (1/2)^{\alpha}\,. \tag{8}$$

Εχουμε λοιπόν, ότι

$$h'(\alpha) = 100 + 1600 \times (1/2)^{\alpha} \times \ln(1/2)\,, \tag{9}$$

και

$$h''(\alpha) = 1600 \times (1/2)^{\alpha} \times \ln^2(1/2) > 0\,. \tag{10}$$

Παρατηρούμε λοιπόν ότι η $h(\alpha)$, θεωρημένη σαν συνάρτηση της ***συνεχούς*** *μεταβλητής α, είναι κυρτή. Για διακριτές τιμές της μεταβλητής α, η $h(\alpha)$ είναι δυνατό να παίρνει την ελάχιστή της τιμή για το πολύ δύο τιμές του α (εφόσον δεν είναι σταθερή).*
Δοκιμάζοντας τιμές του α γύρω από το σημείο $\alpha^*: h'(\alpha^*) = 0$, βρίσκουμε ότι οι διακριτές τιμές του α που ελαχιστοποιούν την $h(\alpha)$ είναι οι $\alpha = 3{,}4$ με $h(3) = h(4) = 800$.

Εχουμε λοιπόν

$$v(2{,}0) = 800\,, \tag{11}$$

και

$$\alpha^*(2{,}0) = 3 \text{ ή } 4\,. \tag{12}$$

Ανάλογα βρίσκουμε ότι

$$\begin{aligned} v(1{,}0) &= \min\{1600\,,\ 300\delta(\alpha) + 100\alpha + v(2{,}0) \times P_0(\alpha)\} = \\ &\overset{(11)}{=} \min\{1600\,,\ 300\delta(\alpha) + 100\alpha + 800 \times (1/2)^\alpha\} = \\ &= 700\,, \end{aligned} \tag{13}$$

και

$$\alpha^*(1{,}0) = 2 \text{ ή } 3\,. \tag{14}$$

Τέλος,

$$v(0{,}0) = \min\{1600\,,\ 300\delta(\alpha) + 100\alpha + 700 \times (1/2)^\alpha\} = 675\,, \tag{15}$$

και

$$\alpha^*(0{,}0) = 2\,. \tag{16}$$

Η λύση μπορεί να περιγραφεί ως εξής : Αρχικά θα παράγουμε μια ομάδα με 2 κομμάτια. Αν τουλάχιστον 1 είναι όχι-ελαττωματικό σταματάμε, διαφορετικά θα παράγουμε άλλη μια ομάδα που θα περιέχει 2 ή 3 κομμάτια. Αν τουλάχιστον 1 κομμάτι είναι όχι-ελαττωματικό, σταματάμε την παραγωγή. Διαφορετικά θα παράγουμε την τελευταία ομάδα που θα περιέχει 3 ή 4 κομμάτια. Αν όλα είναι ελαττωματικά τότε θα έχουμε την ζημιά των 1600 δρχ. .

4.2 Πρόβλημα Σχεδιασμού Επιπ`εδου Παραγωγής Αποθεμάτων. *Ας επανεξετάσουμε το πρόβλημα 1.4.1. Εχουμε μια πιο ρεαλιστική μορφή του προβλήματος, εάν υποθέσουμε ότι η ζήτηση για την περίοδο t έχει την εξής μορφή*

$$d_t = 3 + \epsilon_t\,, \tag{1}$$

όπου ϵ_t είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές με κατανομή που ορίζεται από

$$\epsilon_t = \begin{cases} 1\,, & \text{με πιθανότητα } 1/2 \\ 0\,, & \text{με πιθανότητα } 1/2 \end{cases} \tag{2}$$

Επιπλέον υποθέτουμε ότι όταν το απόθεμα και η παραγωγή δεν φτάνουν για να καλυφθεί η ζήτηση, τότε η επιπλέον ζήτηση χάνεται, και το γεγονός τούτο, όταν συμβεί, έχει σαν αποτέλεσμα αύξηση του κόστους κατά 2 μονάδες επί το ποσό της χαμένης ζήτησης (penalty cost). Ολα τα άλλα δεδομένα του προβλήματος παραμένουν τα ίδια. Το ζητούμενο είναι να προσδιοριστεί μια πολιτική η οποία να ελαχιστοποιεί το συνολικό αναμενόμενο κόστος.

Διατύπωση του Προβλήματος.

1) Καταστάσεις–Αποφάσεις.

X_t = απόθεμα κατά το τέλος της περιόδου t.

A_t = ποσότητα παραγωγής στην αρχή της περιόδου t.

2. Δυναμική του συστήματος.

i) $X_{t+1} = X_t + A_t - d_t$ (3)

ii) $5 \geq A_t \geq 0\,,\;\; X_t + A_t - d_t \leq 4$ (4)

3. Συναρτήσεις Κόστους.

i) Τρέχοντος κόστους

Σύμφωνα με τα δεδομένα του προβλήματος, αν $X_t = x$, $A_t = \alpha$, και $X_{t+1} = y$, έχουμε

$$\tilde{c}(x,\alpha,y) = 1 \times x + c(\alpha) + L(d_t - (x+\alpha))^+ ,$$

άρα θα χρησιμοποιήσουμε την νέα συνάρτηση τρέχοντος κόστους

$$\begin{aligned} c(x,\alpha) &= E\,(\tilde{c}\,(x,\alpha,d_t)\,/\,X_t{=}x\,,\; A_t{=}\alpha) = \\ &= 1 + c(\alpha) + L \times E(\,(dt - (x+\alpha))^+ / X_t{=}x\,,\; A_t{=}\alpha\,) = \\ &= 1 + c(\alpha) + L \times \sum_{\delta} (\delta - (x+\alpha))^+ \times P(d{=}\delta)\,, \end{aligned}$$

π.χ.

$$\begin{aligned} c(1,1) &= 1 + c(1) + L \times (\,(3-2) \times P(d{=}3) + (4-2) \times P(d{=}4)\,) = \\ &= 1 + c(1) + L \times (1 \times \tfrac{1}{2} + 2 \times 1/2)\,, \end{aligned}$$

και

$$\begin{aligned} c(2,3) &= 1 \times 2 + c(3) + L \times (\,(3-5)^+ \times P(d{=}3) + (4-5)^+ \times P(d{=}4)\,) = \\ &= 1 \times 2 + c(3) + 0\,. \end{aligned}$$

Η συνάρτηση $c(\alpha)$ δίδεται από τον αντίστοιχο πίνακα της εφαρμογής 1.4.1 .

ii) Τερματικού κόστους

$$\hat{c}(X_4) = \begin{cases} M\,,\;\; x > 0 \\ \end{cases}$$

$0, \quad x=0$

Παρατηρήσεις. 1. Η εξίσωση (19) μας δίνει την δυναμική του συστήματος στη γνωστή μορφή των πιθανοτήτων

$$
\begin{aligned}
P(y/x,\alpha) &= P(X_{t+1} = y \,/\, X_t{=}x\,,\; A_t{=}\alpha) = \\
&= P(X_t{+}A_t - d_t{=}y \,/\, X_t{=}x\,,\; A_t{=}\alpha) = \\
&= P(x+\alpha-d_t = y \,/\, X_t{=}x\,,\; A_t{=}\alpha) = \\
&= P(d_t{=}x+\alpha-\; y) = \begin{cases} 1/2\,, & \alpha\nu\;\; x+\alpha-y=4 \\ 1/2\,, & \alpha\nu\;\; x+\alpha-y=3 \\ 0\;, & \delta\iota\alpha\varphi o\rho\varepsilon\tau\iota\kappa\acute{\alpha} \end{cases} \qquad (5)
\end{aligned}
$$

Σχηματικά έχουμε

$$\boxed{t,x} \qquad \boxed{\alpha} \quad \begin{matrix} 1/2 \;\; \boxed{t+1,\, y=x+\alpha-4} \\ 1/2 \;\; \boxed{t+1,\;\, y=x+\alpha-3} \end{matrix}$$

αν $x + \alpha \;\geq\; 4$, και

$$\boxed{t,x} \qquad \boxed{\alpha} \qquad \boxed{t+1\;,\,0}$$

αν $x + \alpha \;\leq\; 3$.

2. Ο περιορισμός του επιπέδου της παραγωγής $A_t \;\leq\; 5$ είναι εύκολο να ικανοποιηθεί στη πράξη, στην αρχή κάθε περιόδου t, παράγουμε ποσότητα μικρότερη ή ίση του 5. Αντίθετα, ο περιορισμός $X_t + A_t - d_t \;\leq\; 4$ δεν είναι εύκολο να ικανοποιηθεί στην πράξη, γιατί εκφράζει μια σχέση μεταξύ των τυχαίων μεταβλητών $X_t + A_t$, που είναι όμως γνωστές στην αρχή κάθε περιόδου, και της τυχαίας μεταβλητής d_t, που δεν είναι γνωστή στην αρχή της περιόδου, αλλά στο τέλος. Υπάρχουν τρεις τρόποι αντιμετώπισης αυτής της δυσκολίας, ανάλογα με τις συνθήκες του πραγματικού προβλήματος. Ο πρώτος τρόπος συνίσταται στο να επιβάλουμε κάποιο ισχυρότερο περιορισμό. Παραδείγματος χάριν, δεδομένου ότι $d_t \;\geq\; 3$ με πιθανότητα 1, αν υποθέσουμε ότι $X_t + A_t \;\leq\; 7$, τότε ο περιορισμός $X_t + \; A_t \; - \; d_t \;\leq\; 4$ θα ισχύει με πιθανότητα 1. Μια άλλη λύση στο πρόβλημα συνίσταται στο να επιβάλουμε ο περιορισμός $X_t + \; A_t - \; d_t \;\leq\; 4$ να ισχύει με μια ορισμένη πιθανότητα. Για παράδειγμα, αν πάρουμε $X_t + \; A_t \;=\; 8$, τότε έχουμε

$$P(X_t + A_t - d_t \leq 4) \;=\; P(4 \;\leq\; d_t) \;=\; 1/2\,.$$

Τέλος, είναι δυνατό να θεωρήσουμε ότι, όταν ο περιορισμός $X_t + A_t - d_t \;\leq\; 7$ παραβιάζεται, τούτο έχει σαν αποτέλεσμα κάποιο κόστος το οποίο περιλαμβάνεται στην συνάρτηση κόστους του προβλήματος.

Στο συγκεκριμένο παράδειγμα ας ακολουθήσουμε την απλούστερη μέθοδο, δηλαδή θα υποθέσουμε ότι : $X_t + A_t \leq 7$. Οδηγούμαστε λοιπόν στα σύνολα αποφάσεων της μορφής

$$D(x) \ = \ \{ \ 0,1,\ldots,min\{7-x,\ 5\} \ \} \ .$$

4. Εξισώσεις Βελτιστοποίησης.

$$v(t,x) \ = \ min \ \{ \ c(x,\alpha) + E(v(X_{t+1}) \ / \ X_t{=}x \ , \ A_t{=}\alpha) \ \} \ =$$

$$= \begin{cases} \min\limits_{\alpha \in D(x)} \{ \ c(x,\alpha) + v(t{+}1,x+\alpha-4) \times \frac{1}{2} + v(t+1,x+\alpha-3) \times \frac{1}{2} \ , & \alpha\nu \ \ x+\alpha \geq 4 \\ \\ \min\limits_{\alpha \in D(x)} \{ c(x,\alpha) + v(t+1,0) \ \} \ , & \alpha\nu \ \ x+\alpha \leq 3 \end{cases}$$

$$t \ = \ 0,1,2,3 \ . \tag{6}$$

Τελικές συνθήκες.

$$v(4,x) \ = \ \hat{c}(x) \ . \tag{7}$$

4.3. Πρόβλημα Αντικατάστασης – Συντήρησης Μηχανημάτων. *Θεωρούμε πάλι το πρόβλημα αντικατάστασης ενός μηχανήματος που είδαμε στο κεφάλαιο 1 . Η διαφορά εδώ είναι ότι η φθορά του μηχανήματος σε κάποια χρονική περίοδο δεν είναι από πριν γνωστή, αλλά εξαρτάται από ένα μηχανισμό τυχαιότητας. Ετσι υποθέτουμε ότι στην αρχή κάθε χρονικής περιόδου t η κατάσταση του μηχανήματος όσον αφορά τη λειτουργική του ικανότητα, προσδιορίζεται από ένα ακέραιο αριθμό x_t , που γενικά δηλώνει το επίπεδο φθοράς του στοιχείου, και μπορεί να έχει διαφορετική σημασία ανάλογα με το συγκεκριμένο μοντέλο. Οι δυνατές επιλογές που έχουμε σε κάθε χρονική περίοδο είναι δύο . Η πρώτη να αφήσουμε το μηχάνημα να λειτουργήσει μια ακόμη περίοδο, και η δεύτερη να το αντικαταστήσουμε με ένα καινούργιο. Αν πάρουμε την πρώτη απόφαση, τότε στην αρχή της επόμενης περιόδου η νέα κατάσταση του μηχανήματος δεν είναι ακριβώς γνωστή, αλλά προσδιορίζεται από γνωστές πιθανότητες μετάβασης $P_{ij} = P[X_{t+1} = j \ / \ X_t = i]$, $i,j = 0,1,2,\ldots$. Αν αποφασίσουμε αντικατάσταση, τότε προκύπτει ένα σταθερό κόστος R , και στην αρχή της επόμενης περιόδου το (καινούργιο) μηχάνημα θα βρίσκεται στην κατάσταση 0 , που υποθέτουμε ότι συμβολίζει μηδενική φθορά. Επίσης, κάθε φορά που το μηχάνημα βρίσκεται στην κατάσταση i στην αρχή μιας χρονικής περιόδου, έχουμε κόστος λειτουργίας που είναι ίσο με $c(i)$. Το ζητούμενο είναι να βρεθεί η πολιτική αντικαταστάσεων που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος λειτουργίας του μηχανήματος, όπου ο εκπτωτικός παράγοντας είναι ίσος με β .*

Πριν προχωρήσουμε στη διατύπωση του προβλήματος με τη μέθοδο Δυναμικού Προγραμματισμού, είναι χρήσιμο να δούμε την εφαρμογή του παραπάνω μοντέλου σε δύο ειδικές περιπτώσεις.

1. *Οταν ο παράγοντας που καθορίζει τη φθορά του μηχανήματος είναι η ηλικία του, τότε είναι λογικό να υποθέσουμε ότι η κατάσταση x_t παριστάνει την ηλικία, δηλαδή τον αριθμό των περιόδων που λειτουργεί το μηχάνημα. Σ' αυτή την περίπτωση η μόνη δυνατή μετάβαση είναι από την κατάσταση $x_t = i$ στην $x_{t+1} = i+1$, με πιθανότητα ίση με 1 .*

2. *Στη δεύτερη περίπτωση η φθορά δεν εξαρτάται άμεσα από την ηλικία, αλλά κυρίως από τη χρήση που έχει γίνει στο μηχάνημα μέχρι τη συγκεκριμένη χρονική στιγμή. Αυτό συμβαίνει για παράδειγμα, όταν το μηχάνημα είναι φωτοτυπικό, όπου το πόσο φθείρεται το μηχάνημα εξαρτάται από το πόσα φωτοαντίγραφα έχει παράγει. Εδώ η κατάσταση θα μπορούσε να είναι η ποσότητα αντιγράφων (γενικά προιόντων) που έχει παράγει το* μηχάνημα από την αρχή της λειτουργίας του, και επομένως η επόμενη κατάσταση καθορίζεται από το πόσα αντίγραφα θα παράγει την παρούσα χρονική περίοδο, που φυσικά είναι τυχαία μεταβλητή.

Οπως φαίνεται στα παραπάνω παραδείγματα, η κατάσταση έχει ορισθεί κατά τέτοιο τρόπο, ώστε στην επόμενη χρονική περίοδο να είναι αριθμητικά μεγαλύτερη απ' ότι στην προηγούμενη. Η ιδιότητα αυτή υποθέτουμε ότι ισχύει γενικότερα για το μοντέλο που εξετάζουμε, μια και η κατάσταση συμβολίζει φθορά. Συγκεκριμένα δεχόμαστε ότι ισχύουν οι παρακάτω δύο συνθήκες.

i) Το κόστος λειτουργίας $c(i)$ ε`ιναι αύξουσα συνάρτηση της κατάστασης i.

ii) Για κάθε k, το άθροισμα $\sum_{j=k}^{\infty} P_{ij}$ είναι αύξουσα συνάρτηση ως προς i. Η συνθήκη αυτή μπορεί να ερμηνευθεί ποιοτικά και δείχνει ότι η πιθανότητα να οδηγηθεί το μηχάνημα σε μια κατάσταση από το σύνολο $[k,\infty]$ (δηλαδή μεγάλης φθοράς), είναι τόσο μεγαλύτερη όσο μεγαλύτερη είναι η παρούσα κατάσταση (φθορά).

Διατύπωση του Προβλήματος

1. Χώροι Καταστάσεων Αποφάσεων. Οπως είδαμε παραπάνω, η κατάσταση περιγράφεται από μια ακέραια μεταβλητή x_t που συμβολίζει τη φθορά του μηχανήματος. Επίσης οι χώροι αποφάσεων είναι $D(x_t) = \{1,2\}$, όπου η απόφαση 1 συμβολίζει λειτουργία, και η 2 αντικατάσταση.

2. Δυναμική. Αν $x_t = i$, και $\alpha = 1$, τότε η X_{t+1} είναι τυχαία μεταβλητή με κατανομή πιθανότητας όπως ορίσθηκε παραπάνω από τα P_{ij}.

Αν $\alpha = 2$, τ`οτε αναγαστικ`α $X_{t+1} = 0$.

3. Συναρτήσεις κόστους.

$$c(x_t,\alpha_t) = \begin{cases} c(x_t)\,, & \alpha_t = 1 \\ c(x_t) + R\,, & \alpha_t = 2\,. \end{cases} \tag{23}$$

Δεν χρειάζεται να ορίσουμε συνάρτηση τερματικού κόστους, επειδή μελετάμε το πρόβλημα άπειρου ορίζοντα.

4. Εξίσωση Βελτιστοποίησης.

$$v(i) = c(i) + min\,\{R + \beta v(0)\,,\, \beta\sum_j P_{ij}v(j)\,\}\,. \tag{24}$$

5. Λύση της Εξίσωσης Βελτιστοποίησης.

Θα αποδείξουμε επαγωγικά ότι κάτω από τις συνθήκες i και ii, το ελάχιστο κόστος $v(i)$ είναι αύξουσα συνάρτηση της κατάστασης i. Πραγματικά, αν εφαρμόσουμε τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο τιμών για τη λύση της εξίσωσης, ξέρουμε ότι η λύση $v(i)$ είναι τό όριο της ακολουθίας των διαδοχικών προσεγγ`σεων $v^{(n)}(i)$, που προκύπτουν από τις σχέσεις

$$v^{(n+1)}(i) = c(i) + min\left\{ R + \beta v^{(n)}(0)\,,\, \beta\sum_j P_{ij}v^{(n)}(j) \right\}, \quad n = 1,2,\ldots \tag{25.1}$$

$$v^{(0)}(i) \;=\; 0\,. \tag{25.2}$$

Για $n = 1$ *έχουμε* $v^{(1)}(i) \;=\; c(i)$, *που από τη συνθήκη* i *είναι αύξουσα ως προς* i.

Υποθέτουμε ότι η $v^{(n)}(i)$ *είναι αύξουσα ως προς* i. *Τότε από τη συνθήκη* ii *είναι εύκολο να δειχθεί ότι και η ποσότητα* $\sum_j P_{ij} v^{(n)}(j)$ *είναι επίσης αύξουσα ως προς* i, *και επομένως από την (25.1) βλέπουμε ότι η* $v^{(n+1)}(i)$ *είναι αύξουσα συνάρτηση του* i. *Το ίδιο λοιπόν ισχύει και για την* $v(i) = \lim_{n\to\infty} v^{(n)}(i)$.

Τώρα, η μορφή της βέλτιστης πολιτικής είναι απλή συνέπεια της ιδιότητας που μόλις αποδείξαμε. Συγκεκριμένα ισχύει η εξής

Πρόταση 1. *Αν ισχύουν οι συνθήκες* i *και* ii, *τότε υπάρχει ένας ακέραιος αριθμός* $\bar{i}$, $\bar{i} \le \infty$, *τέτοιος ώστε, σύμφωνα με τη βέλτιστη πολιτική του αναμενόμενου συνολικού εκπτωτικού κόστους, το μηχάνημα αντικαθίσταται αν η κατάστασή του είναι* $x_t \ge \bar{i}$, *και συνεχίζει να λειτουργεί αν* $x_t < \bar{i}$.

Απόδειξη. *Από την εξίσωση βελτιστοποίησης (24) προκύπτει ότι πρέπει να γίνει αντικατάσταση στην κατάσταση* i, *αν*

$$\beta \sum_j P_{ij} v(j) \;\ge\; R + \beta v(0)\,. \tag{26}$$

Αφού η $v(i)$ *είναι αύξουσα ως προς* i, *όπως είδαμε παραπάνω τοίδιο ισχύει και για την* $\sum_j P_{ij} v(j)$. *Επειδή όμως το δεύτερο μέρος είναι σταθερός αριθμός, ανεξάρτητος του* i, *το αποτέλεσμα προκύπτει αμέσως για* $\bar{i}$ *που δίνεται από τη σχέση*

$$\bar{i} \;=\; \min\,[\, i : \; \beta \sum_j P_{ij} v(j) \;\ge\; R + \beta v(0) \,]\,,$$

όπου το $\bar{i}$ *λαμβάνεται ίσο με* ∞ *αν το προηγούμενο σύνολο είναι κενό.*

4.4. Ελεγχος παροχής νερού

$$(\max)\;\; E \sum_{t=0}^{N} r_t(\alpha_t)\,.$$

Δυναμική :

$$X_{t+1} \;=\; (X_t - A_t + Y_t) \;\wedge\; M\,,\;\; t = 0,1,\ldots,N-1\,,\;\; X_t \in [0,M]\,,\;\; A_t \le Y_t\,.$$

Αρχική κατάσταση x_0,

όπου Y_t *είναι τ.μ.*, $y_t \sim F_t$, F_t *είναι γνωστές συναρτήσεις κατανομής.*

Ερμηνεία. A_t *είναι η ποσότητα νερού που χρησιμοποιείται για ύδρευση την περίοδο* t *(διάστημα* $(t,t+1)$*)*, $r_t(\alpha_t)$ *η ωφέλεια την περίοδο* t, *για ποσότητα παροχής* A_t, X_t *η ποσότητα νερού διαθέσιμη στην αρχή της περιόδου* t, M *η χωρητικότητα της διαθέσιμης δεξαμενής,* Y_t *η ποσότητα νερού που εισρέει στη δεξαμενή (π.χ. λόγω βροχοπτώσεων) κατά το διάστημα* $(t,t+1)$.

Χρησιμοποιούμε το συμβολισμό : $\alpha \wedge b \;=\; \max\{a,b\}$, $a \vee b \;=\; \min\{\alpha,b\}$. *Στο παρόν πρόβλημα η δυναμική δίδεται απο συνάρτηση της μορφής*

$$X_{t+1} \;=\; g_t(x,\alpha,y) \;=\; (x - \alpha + y) \;\wedge\; M\,.$$

4.5. Πρόβλημα κατανομής επενδύσεων.

$(max)\ \ Er(X_N)$

Δυναμική

$$X_{t+1} \;=\; (\alpha_t Y_t^1 + (1-\alpha_t)Y_t^2) \times X_t\,,\quad t = 0,1,\ldots,N-1\,,\;\; \alpha_t \in [0,1]\,.$$

Αρχικό σημείο : $X_0 \;=\; x_0$.

Ερμηνεία . $X_t :=$ *η ποσότητα χρημάτων που είναι διαθέσιμη στην αρχή της περιόδου* t .

$A_t :=$ *το ποσοστό της* X_t *που θα επενδυθεί στην πρώτη δραστηριότητα (π.χ. μετοχές).* $(1\text{-}\alpha_t)$ *πρέπει αναγκαστικά να επενδυθεί στην δεύτερη δραστηριότητα (π.χ. λογαριασμός καταθέσεων σε τράπεζα).*

$Y_t^i :=$ *η απόδοση της επένδυσης* i $(i = 1,2)$ *κατά την περίοδο* t *. Αν* $Y_t^i \geq 1$ *έχουμε κέρδος, και αν* $Y_t^i < 1$ *, έχουμε ζημιά.*

$r(x) :=$ *η αξία (για τον επενδυτή) του να βρεθεί με ποσότητα χρημάτων* x *στο τέλος του ορίζοντα. Π.χ.* $r(x) = 1$ *, αν* $x \geq h$ *, και* $r(x) = 0$ *, αν* $x < h$ *, ή* $r(x) = \beta^N x$ *,* $\beta \in (0,1)$.

Δεχόμαστε ότι το ζεύγος των τ.μ. (Y_t^1,Y_t^2) *έχει γνωστές από κοινού κατανομές* $F_t(y_1,y_2)$ *,* $t = 0,1,\ldots,N-1$.

Παρατηρούμε ότι η δυναμική δίδεται από μια εξίσωση της μορφής

$$X_{t+1} \;=\; g_t(X_t,A_t,\overrightarrow{Y}_t) \;=\; (\alpha Y_t^1 + (1-\alpha)Y_t^2) \times X_t\,,\;\; \alpha \in [0,1]\,.$$

4.6. Πρόβλημα κατανομής επενδύσεων – κατανάλωσης

$$(max)\quad E\Big(\sum_{t=0}^{N} r_t(A_t^1)\Big)$$

Δυναμική

$$X_{t+1} \;=\; (A_t^2\, Y_t^1 \;+\; (1 - A_t^2)\, Y_t^2\,)\,(X_t - A_t^1)$$

Αποφάσεις :

$$A_t \;=\; (A_t^1,A_t^2)\,,\;\; A_t^2 \in\; [0,1]\,,\;\; A_t^1 \;\leq\; x_t\,.$$

Αρχικό σημείο $X_0 \;= x_0$.

Ερμηνεία. $X_t :=$ *ποσότητα χρημάτων που είναι διαθέσιμη στην αρχή της περιόδου* t *,* $t{=}0,1,\ldots,N-1$.

$A_t^1 :=$ *ποσότητα χρημάτων που διατίθεται για κατανάλωση κατά την διάρκεια της περιόδου* $(t,t{+}1)$ *,* $t = 0,\ldots,N-1$.

$A_t^2 :=$ *το ποσοστό της* $(X_t - A_t^1)$ *(= ποσότητας κεφαλαίου που διατίθεται για επένδυση) που θα επενδυθεί στην πρώτη δραστηριότητα. Επομένως ποσό* $A_t^2\,(X_t - A_t^1)$ *θα επενδυθεί στην πρώτη δραστηριότητα , ενώ* $(1 - A_t^2)\,(X_t - A_t^1)$ *θα διατεθεί στην δεύτερη δραστηριότητα κατά την περίοδο* t .

$r_t(A) :=$ *η ωφέλεια όταν διατεθεί ποσότητα* α *για κατανάλωση κατά την χρονική περίοδο* t *, (διάστημα* $(t,t{+}1)$ *) .*

Το ζεύγος των τ.μ. (Y_t^1, Y_t^2) έχει το ίδιο νόημα με αυτό που είχε στο πρόβλημα 4.4. Παρατηρούμε λοιπόν, ότι οι αποφάσεις είναι πολυδιάστατες $\overline{A} = (A^1, A^2)$, καθώς επίσης και ο μηχανισμός της ταχύτητας $\overline{Y} = (Y^1, Y^2)$.

4.7. Πρόβλημα Σταθεροποίησης

$$(min) \quad E\left(\sum_{t=0}^{N} (X_t - A_t)^2 + c\sum_{t=0}^{N} A_t^2 \right)$$

Δυναμική

$$X_{t+1} = g(X_t, A_t, Y_t) = X_t - A_t + Y_t .$$

Αρχικό σημείο $X_0 = x_0$,
όπου Y_t είναι τυχαίες μεταβλητές με γνωστές συναρτήσεις κατανομής F_t.

5. Εφαρμογές και Προγράμματα Η.Υ.

ΕΦΑΡΜΟΓΗ 1. Θεωρούμε τον εξής τύπο προβλήματος ελαχιστοποίησης του κόστους παραγωγής και αποθεμάτων.
Ο ορίζοντας προγραμματισμού είναι άπειρος, η ζήτηση γενικά πιθανοθεωρητική, και μας ενδιαφέρει να βρούμε την πολιτική που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος. Δεχόμαστε ότι η ζήτηση δεν είναι σταθερή σε κάθε χρονική περίοδο, αλλά εμφανίζει περιοδικότητα. Αυτό σημαίνει ότι υπάρχει ένας αριθμός περιόδων N, τέτοιος ώστε η κατανομή της τυχαίας μεταβλητής που αντιστοιχεί στη ζήτηση να επαναλαμβάνεται περιοδικά με περίοδο N. Το φαινόμενο αυτό παρατηρείται για παράδειγμα σε φαρμακεία, όπου η ζήτηση αντηλιακών παρουσιάζει διακυμάνσεις κατά τη διάρκεια του χρόνου, αλλά μένει περίπου σταθερή από χρόνο σε χρόνο. (Ετσι μοντελοποιούμε την περιοδικότητα της ζήτησης).
Σε ότι αφορά την κατανομή της ζήτησης στη διάρκεια μιας χρονικής περιόδου, υποθέτουμε ότι είναι διακριτή, και από περίοδο σε περίοδο μεταβάλλονται οι δυνατές τιμές και οι αντίστοιχες πιθανότητες.
Οι συναρτήσεις κόστους παραγωγής και διατήρησης αποθεμάτων είναι, αντίστοιχα, οι εξής

$$c_t(\alpha) = \alpha^2 , \quad \text{και} \quad h_t(x) = h \cdot x ,$$

όπου τα α, x συμβολίζουν ποσότητες παραγωγής και αποθεμάτων αντίστοιχα.
Δεν είναι απαραίτητο να ικανοποιείται η ζήτηση οπωσδήποτε, αλλά όταν παρουσιαστεί έλλειψη s μονάδων προιόντος σε κάποια χρονική περίοδο, τότε λόγω διαφυγόντων κερδών, απώλειας φήμης της εταιρείας κλπ, προκύπτει το επιπλέον κόστος έλλειψης $L(x) = 5s^2$.
Με βάση τις παραπάνω παραδοχές ζητείται να βρεθεί με τη βοήθεια Η.Υ. η βέλτιστη πολιτική ως προς το αναμενόμενο συνολικό εκπτωτικό κόστος, για διάφορες περιπτώσεις κατανομής της ζήτησης και διάφορες τιμές του μοναδιαίου κόστους αποθήκευσης και του εκπτωτικού παράγοντα.

Διατύπωση του Προβλήματος.

Είναι εύκολο να δούμε ότι στο διάνυσμα κατάστασης περιλαμβάνεται η περίοδος μέσα στο διάστημα ανακύκλωσης, και το διαθέσιμο απόθεμα στην αρχή της περιόδου.
Υποθέτουμε ότι αν σε κάποια περίοδο η ζήτηση ξεπεράσει το απόθεμα, η επιπλέον ζήτηση δεν χάνεται, αλλά μένει και ικανοποιείται όταν γίνει νέα παραγωγή. Αυτό σημαίνει ότι επιτρέπουμε στο απόθεμα να παίρνει και αρνητικές τιμές.
Για λόγους προγραμματισμού είναι απαραίτητο να ορίσουμε μέγιστο και ελάχιστο επιτρεπόμενο απόθεμα M και m αντίστοιχα, ώστε οι χώροι καταστάσεων και αποφάσεων να είναι πεπερασμένοι.

Συμβολίζουμε με (k,x) το διάνυσμα κατάστασης όπως ορίστηκε παραπάνω, με D_k τη ζήτηση κατά την περίοδο k, και με $p_k(d)$ την κατανομή της ζήτησης κατά την περίοδο k, η οποία επίσης για προγραμματιστικούς λόγους υποθέτουμε ότι είναι διακριτή.

Είναι εύκολο, με βάση την προηγούμενη περιγραφή, να βρούμε την δυναμική του συστήματος. Ετσι, από την κατάσταση (k,x) *, θα πάμε στην* (k',x') *, όπου* $k' = k+1$*, αν* $k < N$ *, και* $k' = 1$ *, αν* $k = N$ *, ενώ* $x' = x + \alpha - D_k$ *.*

Το τρέχον κόστος για μια περίοδο είναι ίσο με $\alpha^2 + h \cdot (x - D_k)^+ + 5 \cdot [(D_k - x)^-]^2$ *όπου* $x^+ = \max\{x,0\}$ *, και* $x^- = \max\{-x,0\}$ *.*

Τώρα μπορούμε να γράψουμε την εξίσωση βελτιστοποίησης

$$v(k,x) = \min_{0 \le \alpha \le M-x} \Big\{ \alpha^2 + \sum_{d=0}^{x} p_k(d) \cdot h \cdot (x-d) + \sum_{d>x} p_k(d) \cdot 5 \cdot (d-x)^2 +$$

$$+ \beta \sum_d p_k(d) \cdot v(k',x+\alpha-d) \Big\} .$$

Παρακάτω δίνουμε τον κώδικα του προγράμματος σε γλώσσα Pascal.

```
{Πρόβλημα αποθεμάτων άπειρου ορίζοντα με στοχαστική ζήτηση.  Εφαρμογή της
μεθόδου διαδοχικών προσεγγίσεων στο χώρο τιμών για την εύρεση της πολιτικής
που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος.
Η εξίσωση βελτιστοποίησης είναι της μορφής :
        v(n+1)(k,x)=Tv(n)(k,x),
όπου  T  είναι ο γνωστός τελεστής
min{c(x,a)+bEv(k1,x1)}    }

program inventory;

{Δηλώσεις μεταβλητών}

const

   mininv=-50;          {Μέγιστη επιτρεπόμενη Ελλειψη}
   maxinv = 200;        {Μέγιστο Επιτρεπόμενο Απόθεμα}

   maxplan = 16;        {Περίοδος Ανακύκλωσης}

   maxdem = 10;         {Μέγιστος Αριθμός Δυνατών Τιμών Ζήτησης}

type

   cost = array [1..maxplan,mininv..maxinv] of real;

   demand  = array [1..maxplan,1..maxdem] of integer;

   prob = array [1..maxplan,1..maxdem] of real;

   dec  = array [1..maxplan,mininv..maxinv] of integer;
```

```
var

  k,i,j,hm,pl,mi,md,x,count :integer;

  v,u : cost;           {Πίνακες Διαδοχικών Προσεγγίσεων}

                        {Ο u περιέχει την παρούσα προσέγγιση , και ο v
                         την αμέσως επόμενη, που υπολογίζεται κάθε φορά}

  d : demand;           {Ζήτηση}

  p : prob;             {Πιθανότητες Αντίστοιχων Ζητήσεων}
  ac : dec;             {Πολιτική}

  ans : char;

  b,e,em : real;



function max(a,b: integer) : integer;

begin

  if b > a then max:=b else max:=a

end;



{Υπολογισμός κόστους για κάθε εναλλακτική απόφαση (k,x,a) }


function Tua(k,x,a:integer) :real;


var h,c1:real;

  kn,i:integer;


{ Κόστος Παραγωγής α κομματιών = α².

  Κόστος Αποθήκευσης x κομματιών = hm × x .
```

```
    Κόστος Ελλειψης s κομματιών = 5s²}

begin
    if k<pl then kn:=k+1 else kn:=1;    {Επόμενη Περίοδος}
    h:=0; c1:=0;
    for i:=1 to md do
    begin

        {Αναμενόμενο Κόστος μιας Περιόδου}
        h:=h+p[k,i]*(hm*max(0,x-d[k,i])+5*sqr(max(0,d[k,i]-x)));

        {Αναμενόμενο Κόστος μετά την Επόμενη Περίοδο}
        c1:=c1+b*p[k,i]*u[kn,max(-mi,x+a-d[k,i])];
    end;
    Tua:=sqr(a)+h+c1;
end;




{Υπολογισμός Ελάχιστου Κόστους για κάθε Κατάσταση}

function  Tu(k,x:integer): real;

var i,n,a,kn : integer;
    tr,t :real;
begin
    n:=mi-x;
    t:=1e8;
```

```
    for a:=0 to n do                {Δυνατές Ποσότητες Παραγωγής}
    begin
        tr:=Tua(k,x,a);
        if tr < t then
        begin
            t:=tr;
            ac[k,x]:=a;
        end;
    end;
    Tu:=t;
end;




{Κύριο Πρόγραμμα}


begin


{Εισαγωγή Στοιχείων}


    ClrScr;
    GotoXY(1,1);
    write('ΣΥΝΤΕΛΕΣΤΗΣ ΚΟΣΤΟΥΣ ΑΠΟΘΗΚΕΥΣΗΣ: ');
    readln(hm);
    write('ΠΕΡΙΟΔΟΣ ΑΝΑΚΥΚΛΩΣΗΣ: ');
    readln(pl);
    write('ΜΕΓΙΣΤΟ ΕΠΙΤΡΕΠΟΜΕΝΟ ΑΠΟΘΕΜΑ: ');
    readln(mi);
    write('ΑΡΙΘΜΟΣ ΔΥΝΑΤΩΝ ΤΙΜΩΝ ΖΗΤΗΣΗΣ: ');
```

```
readln(md);
write('ΕΚΠΤΩΤΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ: ');
readln(b);
write('ΑΚΡΙΒΕΙΑ ΔΙΑΔΟΧΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ: ');
readln(em);
for i:= 1 to pl do
begin
   writeln('ΖΗΤΗΣΗ  ',i,'ηs ΠΕΡΙΟΔΟΥ: ');
   for j:=1 to md do
   begin
      write('   ');
      read(d[i,j]);
      write(' με πιθανότητα : ');
      readln(p[i,j]);
   end;
   writeln;
end;
writeln;
write('printer ? (y/n)');
readln(ans);
ClrScr;
```

{Αρχικές Τιμές}

```
for i:=1 to pl+1 do
   for j:=-mi to mi do
   begin
      v[i,j]:=0;u[i,j]:=0;
      ac[i,j]:=0;
```

```
    end;

    e:=1e7;

    count:=1;


{Κύριος Βρόχος Επαναλήψεων}


  while e>em do                  {Εξέταση αν έχει επιτευχθεί η ακρίβεια}

  begin

    e:=0;

    for k:=1 to pl do

      for x:=-mi to mi do

      begin

        v[k,x]:=Tu(k,x);      {Εξίσωση Βελτιστοποίησης}

        {Υπολογισμός νέας ακρίβειας προσεγγίσεων}

        if e<abs(v[k,x]-u[k,i]) then e:=abs(v[k,x]-u[k,x]);

      end;

      for k:=1 to pl do

        for x:=-mi to mi do

          u[k,x]:=v[k,x];         {Υπολογισμός του νέου πίνακα προσεγγίσεων}

      count:=count+1;              {Αριθμός Βημάτων}

      GotoXY (10,10); write('Βήμα : ',count:5,'   Ακρίβεια: ',e:8:3);

  end;


{Αποτελέσματα}
```

```
   for k:=1 to pl do
      for i:=-mi to mi do
      begin
         writeln('a(',k,',',i,')=',ac[k,i],'   v(',k,',',i,')=',v[k,i]);
         if ans='y' then
         writeln(Lst,'a(',k,',',i,')=',ac[k,i]:3,'  v(',k,',',i,')=',v[k,i]:8:3);
      end;
end.
```

<u>ΕΦΑΡΜΟΓΗ 2</u> *Ο προιστάμενος του τμήματος προσωπικού της Ο.Α. αντιμετωπίζει το εξής πρόβλημα. Θέλει να αναπτύξει μια πολιτική πρόσληψης αεροσυνοδών για τα επόμενα χρόνια. Οι απαιτήσεις της εταιρείας σε ώρες πτήσης αεροσυνοδών κάθε μήνα είναι ανεξάρτητες διακριτές τυχαίες μεταβλητές D_k, με διαφορετικές γενικά κατανομές $p_k(d)$. Οι απαιτήσεις αυτές όμως υποθέτουμε ότι επαναλαμβάνονται περιοδικά ως προς την κατανομή από χρόνο σε χρόνο, όπως και στο προηγούμενο παράδειγμα.*

Η εκπαίδευση για κάθε καινούργια αεροσυνοδό διαρκεί ένα μήνα και απαιτεί 100 ώρες επί-βλεψης απο μια εκπαιδευμένη αεροσυνοδό. Ετσι η εταιρεία χάνει 100 ώρες αεροσυνοδού για την εκπαίδευση κάθε καινούργιας αεροσυνοδού.

Κάθε αεροσυνοδός μπορεί να δουλέψει το πολύ 150 ώρες το μήνα, και στο τέλος κάθε μήνα ένα ποσοστό r των αεροσυνοδών φεύγουν απο την εταιρεία για προσωπικούς λόγους. Το ποσοστό αυτό είναι επίσης τυχαία μεταβλητή, της οποίας όμως η κατανομή q(r) είναι η ίδια για όλες τις χρονικές περιόδους (μήνες).

Η εταιρεία ακολουθεί την πολιτική να μην απολύει αεροσυνοδούς, έτσι όταν ο μέγιστος αριθμός ωρών εργασίας που είναι διαθέσιμες ξεπερνά τις απαιτήσεις πτήσης και εκπαίδευσης του μήνα, οι υπάρχουσες αεροσυνοδοί απλώς εργάζονται λιγότερες ώρες.

Το μηνιαίο κόστος για την εταιρεία είναι R_1 για κάθε αεροσυνοδό και R_2 για κάθε υπό εκπαίδευση αεροσυνοδό.

Οταν παρουσιαστεί ζήτηση σε ώρες πτήσης και επίβλεψης μεγαλύτερη από τις διαθέσιμες ώρες εργασίας, τότε οι επιπλέον ώρες καλύπτονται με υπερωριακή απασχόληση που στοιχίζει στην εταιρεία ένα συγκεκριμένο ποσό s για κάθε επιπλέον ώρα.

Ζητείται να βρεθεί η πολιτική προσλήψεων που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος της εταιρείας, για μεγάλο (προσεγγιστικά άπειρο) χρονικό διάστημα, και για διάφορες κατανομές ζήτησης και διάφορες τιμές της περιόδου ανακύκλωσης και των σταθερών R_1, R_2, s.

Ζητείται επίσης να βρεθεί η πολιτική προσλήψεων στην περίπτωση που η εκπαίδευση της αεροσυνοδού διαρκεί δύο μήνες.

Διατύπωση του Προβλήματος

Στο διάνυσμα κατάστασης περιλαμβάνεται η περίοδος k μέσα στο διάστημα ανακύκλωσης (δηλαδή ο μήνας) και ο αριθμός αεροσυνοδών x στην αρχή της περιόδου. Στην περίπτωση που η εκπαίδευση διαρκεί δύο μήνες, στην κατάσταση πρέπει να περιληφθεί και ο αριθμός των αεροσυνοδών που έχουν συμπληρώσει ένα μήνα εκπαίδευσης.

Για το μέγεθος του χώρου καταστάσεων μπορούμε να κάνουμε την εξής παρατήρηση. Ο αριθμός των αεροσυνοδών x σε κάθε περίοδο δεν είναι απαραίτητο να ξεπερνάει ένα άνω φράγμα M, που εξαρτάται από το μέγιστο δυνατό αριθμό των ωρών πτήσης που μπορεί να προκύψουν κάποια περίοδο, και που πρέπει κάθε φορά να υπολογίζεται από τα δεδομένα του προβλήματος.

Κατά τα άλλα, ακολουθώντας τον συμβολισμό της εφαρμογής 1, με $N = 12$, έχουμε $x' = r \cdot (x + \alpha)$, όπου α ο αριθμός αεροσυνοδών που προσλαμβάνονται την περίοδο k. Επίσης, αφού επιτρέπεται "έλλειψη" ωρών πτήσης, ο μέγιστος αριθμός αεροσυνοδών που μπορούν να προσληφθούν καθορίζεται μόνο από τις δυνατότητες που έχουν οι ήδη υπάρχουσες αεροσυνοδοί να τις εκπαιδεύσουν. Ετσι πρέπει απλώς $0 \leq \alpha \leq [150x/100]$, όπου με $[x]$ συμβολίζουμε το ακέραιο μέρος του αιθμού x.

Η εξίσωση βελτιστοποίησης είναι

$$v(k,x) = \min_{0 \leq \alpha \leq [1.5x]} \Big\{ R_1 x + R_2 \alpha + \sum_d p_k(d) \cdot s \cdot (150x - 100\alpha - d) +$$

$$+ \beta \cdot \sum_r q(r) \cdot v(k', r \cdot (x + \alpha)) \Big\} .$$

{Πρόβλημα άπειρου ορίζοντα με στοχαστική ζήτηση. Εφαρμογή της μεθόδου διαδοχικών προσεγγίσεων στο χώρο τιμών για την εύρεση της πολιτικής που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος.
Η εξίσωση βελτιστοποίησης είναι της μορφής :
v(n+1)(k,x)=Tv(n)(k,x),
όπου T είναι ο γνωστός τελεστής
max{c(x,α)+bEv(k1,x1)}

```
program OA_stoch_version;


{Δηλώσεις μεταβλητών}


const
   maxn    = 280;          {Μέγιστος Αριθμός Αεροσυνοδών}
   maxplan = 16;           {Μέγιστη Περίοδος Ανακύκλωσης}
   maxdem  = 10;           {Μέγιστος Αριθμός Δυνατών ΤιμώνΖήτησης}

type
   cost = array [1..maxplan,0..maxn] of real;
   demand  = array [1..maxplan,1..maxdem] of integer;
   prob = array [1..maxplan,1..maxdem] of real;
   dec  = array [1..maxplan,0..maxn] of integer;
```

```
var
    k,i,j,pl,ma,md,x,count :integer;
    res : array[1..4] of real;        {Ποσοστά αεροσυνοδών που φεύγουν  (4 τιμές)}
    q   : array[1..4] of real;        { Πιθανότητες των παραπάνω ποσοστών}
    v,u : cost;                   {Πίνακες Διαδοχικών Προσεγγίσεων}

                                 {O  u  περιέχει την παρούσα προσέγγιση , και ο  v
                                   την αμέσως επόμενη, που υπολογίζεται κάθε φορά}
    d : demand;                   {Ζήτηση}
    p : prob;                     {Πιθανότητες Αντίστοιχων Ζητήσεων}
    ac : dec;                     {Πολιτική}
    ans : char;
    b,e,em,cm,rsal,edsal : real;



function max(a,b: integer) : integer;     {Μέγιστος δύο ακέραιωναριθμών}
begin
    if b > a then max:=b else max:=a
end;



function maxar(d1:demand) :integer;   {Εύρεση του μήνα με τις περισσότερες ώρες
                                          πτήσης}
var m:integer;
    i,j:integer;
begin
    m:=0;
    for i:=1 to pl do
    for j:=1 to md do
        if d1[i,j]>m then m:=d1[i,j];
    maxar:=m div 150;
end;


{Υπολογισμός κόστους για κάθε εναλλακτική απόφαση  (k,x,a) }

function Tua(k,x,a:integer) :real;
var h,c1:real;
    kn,i:integer;
    begin
        if k<pl then kn:=k+1 else kn:=1;    {Επόμενη Περίοδος}
        h:=rsal*x+edsal*a; c1:=0;

        {Αναμενόμενο Κόστος μιας Περιόδου}
        for i:=1 to md do
        h:=h+cm*p[k,i]*max(0,150*x-100*a-d[k,i]);

        {Αναμενόμενο Κόστος μετά την Επόμενη Περίοδο}
        for i:=1 to 4 do
        c1:=c1+b*q[i]*u[kn,trunc((1-res[i])*(x+a))];

        Tua:=h+c1;
    end;
```

```
{Υπολογισμός Ελάχιστου Κόστους για κάθε Κατάσταση}

function  Tu(k,x:integer): real;
var i,n,a,kn : integer;
    tr,t :real;
begin
    n:=trunc(1.5*x);     {Μέγιστος Αριθμός Αεροσυνοδών που μπορεί να προσληφθούν}
    t:=1e8;
    for a:=0 to n do
    begin
        tr:=Tua(k,x,a);
        if tr < t then
        begin
            t:=tr;
            ac[k,x]:=a;
        end;
    end;
    Tu:=t;
end;


{Κύριο Πρόγραμμα}


begin

{Εισαγωγή Δεδομένων}

    ClrScr;
    GotoXY(1,1);
    write('ΜΙΣΘΟΣ ΑΕΡΟΣΥΝΟΔΟΥ:  ');
    readln(rsal);
    write('ΜΙΣΘΟΣ ΕΚΠΑΙΔΕΥΟΜΕΝΗΣ ΑΕΡΟΣΥΝΟΔΟΥ :  ');
    readln(edsal);
    write('ΠΕΡΙΟΔΟΣ ΑΝΑΚΥΚΛΩΣΗΣ  ');
    readln(pl);
    write('ΚΟΣΤΟΣ ΥΠΕΡΩΡΙΩΝ ΑΝΑ ΩΡΑ ΠΤΗΣΗΣ :  ');
    readln(cm);
    write('ΑΡΙΘΜΟΣ  ΔΥΝΑΤΩΝ ΤΙΜΩΝ ΖΗΤΗΣΗΣ ');
    readln(md);
    write('ΕΚΠΤΩΤΙΚΟΣ  ΠΑΡΑΓΟΝΤΑΣ :  ');
    readln(b);
    write('ΑΚΡΙΒΕΙΑ ΔΙΑΔΟΧΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ  :  ');
    readln(em);

    for i:= 1 to pl do
    begin
        writeln('ΖΗΤΗΣΗ ',i,'ηs ΠΕΡΙΟΔΟΥ : ');
        for j:=1 to md do
        begin
            write('   ');
            read(d[i,j]);
            write('     ΜΕ ΠΙΘΑΝΟΤΗΤΑ : ');
            readln(p[i,j]);
        end;
        writeln;

    end;
```

```
writeln ('ΠΟΣΟΣΤΑ ΑΕΡΟΣΥΝΟΔΩΝ ΠΟΥ ΠΑΡΑΙΤΟΥΝΤΑΙ :');
for i:=1 to 4 do
begin
    write('  ',i,'.  '); read(res[i]);
    write('    ΜΕ ΠΙΘΑΝΟΤΗΤΑ : '); readln(q[i]);
end;

writeln;
write('ΕΚΤΥΠΩΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ; (y/n)');
readln(ans);
```

```
{Αρχικές Τιμές}

ma:=maxar(d);
for i:=1 to pl do
for j:=0 to ma do
begin
    v[i,j]:=0;u[i,j]:=0;
    ac[i,j]:=0;
end;
ClrScr;
e:=1e7;
count:=1;
```

```
{Κύριος Βρόχος Επαναλήψεων}

while e>em do                    {Εξέταση αν έχει επιτευχθεί η ακρίβεια}
begin
    e:=0;
    for k:=1 to pl do
    for x:=0 to ma do
    begin
        v[k,x]:=Tu(k,x);              {Εξίσωση Βελτιστοποίησης}
        {Υπολογισμός Νέας Ακρίβειας Προσεγγίσεων}
        if e<abs(v[k,x]-u[k,i]) then e:=abs(v[k,x]-u[k,x]);
    end;
    for k:=1 to pl do
    for x:=0 to ma do
    u[k,x]:=v[k,x];                      {Αλλαγή Παλιάς και Νέας Προσέγγισης}
    count:=count+1;
    GotoXY (10,10); write('ΒΗΜΑ : ',count:5,'   ΑΚΡΙΒΕΙΑ : ',e:8:3);
    end;
```

```
{Αποτελέσματα}
for k:=1 to pl do
for i:=0 to ma do
begin
    writeln('a(',k,',',i,')=',ac[k,i],'   v(',k,',',i,')=',v[k,i]);
    if ans='y' then
```

```
        writeln(Lst,'a(',k,',',i,')=',ac[k,i]:3,'      v(',k,',',i,')=',v[k,i]:8:3);
    end;
end.
```

ΕΦΑΡΜΟΓΗ 3 . *Ας θεωρ`ησουμε το εξ`ης πρ`οβλημα καθορισμο`υ β`ελτιστων πολιτικ`ων συντ`ηρησης συστημ`ατων, βλπ. Katehakis and Derman (1984), Katehakis and Melolidakis (1988), Katehakis and Derman (1989). Ενα σ`υστημα πο`υ αποτελε`ιται απ`ο* N *στοιχε`ια (π.χ υπο-συστ`ηματα) λειτουργε`ι συνεχ`ως κα`τω απ`ο την επ`ιβλεψη εν`ος πεπερασμ`ενου αριθμο`υ* R *μον`αδων συντ`ηρησης. Τα στοιχε`ια του συστ`ηματος με τη π`αροδο του χρ`ονου παθα`ινουν βλ`αβες και σταματ`α η λειτουργ`ια τους. Ο χρ`ονος λειτουργ`ιας, χωρ`ις βλ`αβη, του στοιχε`ιου* $i$ *ε`ιναι μ`ια τυχα`ια μεταβλητ`η που ακολουθε`ι εκθετικ`η κατανομ`η με γνωστ`η παρ`αμετρο* μ_i $(1 \leq i \leq N)$. *Επιπλ`εον δεχ`ομαστε `οτι ο χρ`ονος λειτουργ`ιας κ`αθε στοιχε`ιου `ειναι ανεξ`αρτητητος απ`ο την κατ`ασταση που βρ`ισκονται τα υπ`ολοιπα στοιχε`ια του συστ`ηματος. Μη λειτουργο`υντα στοιχε`ια μπορο`υν να επιδιορθωθο`υν, απ`ο τις μον`αδες συντ`ηρησης, και ο χρ`ονος που απαιτε`ιται γι`α την επιδι`ορθωση του στοιχε`ιου* $i$ *ε`ιναι μ`ια τυχα`ια μεταβλητ`η που ακολουθε`ι εκθετικ`η κατανομ`η με γνωστ`η παρ`αμετρο* λ_i $(1 \leq i \leq N)$. *Μετ`α απ`ο κ`αθε επιδι`ορθωση τα στοιχε`ια συμεριφ`ερονται σαν καινο`υργια. Τ`ελος, υποθ`ετουμε `οτι ο χρ`ονος που απαιτε`ιται γι`α την μεταφορ`α μ`ιας μον`αδας συντ`ηρησης απ`ο `ενα χαλασμ`ενο στοιχε`ιο σε `ενα `αλλο ε`ιναι αμελειτα`ιος. Οταν ο αριθμ`ος των μη λειτουργο`υντων στοιχε`ιων ε`ιναι μεγαλ`υτερος απ`ο τον αριθμ`ο των διαθ`εσιμων μον`αδων συντ`ηρησης τ`οτε πρ`επει να επιλ`εγε`ι το υποσ`υνολο των μη λειτουργο`υντων στοιχε`ιων στα οπο`ια θα διατεθο`υν οι μον`αδες συντ`ηρησης. Το ζητο`υμενο ε`ιναι να προσδιοριστε`ι `ενας καν`ονας γι`α τις αποφ`ασεις αυτε`s (μ`ια πολιτικ`η) που μεγιστοποιε`ι `ενα μ`ετρο απ`οδοσης γι`α τη λειτουργ`ια του συστ`ηματος. Ενα μ`ετρο απ`οδοσης πο`υ `εχει ν`οημα στο πρ`οβλημα αυτ`ο ε`ιναι η διαθεσιμότητα (availability) του συστήματος που ορίζεται σαν το αναμενόμενο ποσοστό του χρόνου που το σύστημα βρίσκεται σε λειτουργία (με`σος χρ`ονος λειτουργ`ιας του συστ`ηματος).*

Ζητείται να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού, και να βρεθεί προγραμματιστικά η βέλτιστη λύση.

Διατ`υπωση του προβλ`ηματος.

***1. Χ`ωροι καταστ`ασεων και αποφ`αεων.*** *Σε κ`αθε χρονικ`η στιγμ`η η κατ`ασταση του συστ`ηματος μπορε`ι να περιγραφε`ι με `ενα δι`ανυσμα* $\boldsymbol{x} = (x_1,...,x_N)$ *`οπου δεχ`ομαστε `οτι* $x_i = 1$ *or* $0$ *αν`αλογα με το κατ`α π`οσο το* i *μηχ`ανημα λειτουργε`ι `η ε`ιναι χαλασμ`ενο. Ετσι,* $S = \{0,1\}^n$ *ε`ιναι το σ`υνολο των δυνατ`ων καταστ`ασεων.*

Γι`α κ`αθε κατ`ασταση* $\boldsymbol{x} \in S$, *ας ορ`ισουμε τ`ο σ`υνολο $C_0(\boldsymbol{x}) = \{i : x_i = 0\}$. *Οι διαθ`εσιμες αποφ`ασεις στη κατ`ασταση* $\boldsymbol{x}$ *ε`ιναι τα υποσ`υναλα του* $C_0(\boldsymbol{x})$ *που αποτελο`υνται απ`ο* $R$ *το πολ`υ στοιχε`ια. Συμβολικ`α:* $D(\boldsymbol{x}) = \{\alpha \subset C_0(\boldsymbol{x}) , \|\alpha\| \leq R\}$.

2. Δυναμικ`η του συστ`ηματος. *Ας ορ`ισουμε τον εξ`ης συμβολισμ`ο*

$C_1(\boldsymbol{x}) = \{i : x_i = 1\}$, $(\delta_i,\boldsymbol{x},\theta) = ((x_1,...,x_{i-1},\delta,x_{i+1},...,x_N))$, *`οπου* $\delta = 0, 1$,

$\mu(\boldsymbol{x}) := \sum_1^N x_i\mu_i$, $\lambda(\alpha) := \sum_{j\in\alpha} \lambda_j$.

Οταν το σ`υστημα βρ`ισκεται στη κατ`ασταση* $\boldsymbol{x}$ *και επιλ`εγεται η απ`οφαση* $\alpha \in D(\boldsymbol{x})$ *τ`οτε: (α) οχρ`ονος παραμον`ης στη κατ`ασταση* $\boldsymbol{x}$ *ε`ιναι μ`ια τυχα`ια μεταβλητ`η πο`υ ακολουθε`ι την εκθετικ`η κατανομ`η με παρ`αμετρο $q(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \lambda(\alpha)$ *κα`ι (β) οι εξ`ης μεταπηδ`ησεις ε`ιναι δυνατ`ες:*

- *i) στη κατ`ασταση* $(1_i,\boldsymbol{x})$, *με πιθαν`οτητα* $\lambda_i/q(\boldsymbol{x})$, $i \in \alpha \in D(\boldsymbol{x})$,
- *ii) στη κατ`ασταση* $(0_k ,\boldsymbol{x})$, *με πιθαν`οτητα* $\mu_k/q(\boldsymbol{x})$, $k \in C_1(\boldsymbol{x})$.

3. Συναρτήσεις Απόδοσης. Δεδομένου ότι το μέτρο απόδοσης πού μας ενδιαφέρει είναι η διαθεσιμότητα του συστήματος σαν σύνολο, είναι εύκολο να δούμε ότι η πρέπουσα δομή απόδοσης ορίζεται από τη συνάρτηση δομής του συστήματος καί είναι η εξής

$$\bar{r}(\mathbf{x}) = r(\mathbf{x})/q(\mathbf{x}) \tag{1}$$

όπου στην (1) η συνάρτηση $r(\mathbf{x})$ παίρνει τις τιμές 1 ή 0 ανάλογα με το κατά πόσο στη κατάσταση $\mathbf{x}$ το σύστημα , σαν σύνολο, λειτουργεί ή όχι , έχουμε διαιρέσει με τον όρο $q(\mathbf{x})$ δεδεμένου ότι ο αναμενόμενος χρόνος παραμονής στη κατάσταση $\mathbf{x}$ είναι ίσος με $1/q(\mathbf{x})$, (βλπ. (α) παραπάνω).

4. Εξισώσεις Βελτιστοποίησης. Σύμφωνα με τα παραπάνω, οι εξισώσεις βελτιστοποίησης (βλπ. εξισώσεις (3.5), (3.6)) είναι οι εξής

$$\psi(\mathbf{x}) = \max_{\alpha \in D(\mathbf{x})} \left\{ \left(\sum_{j\in\alpha} \lambda_j \psi((1_j,\mathbf{x})) + \sum_{1}^{N} x_k \mu_k \psi((0_k,\mathbf{x})) \right)/q(\mathbf{x}) \right\}, \tag{2}$$

$$\psi(\mathbf{x})/q(x) + d(\mathbf{x}) = \max_{\alpha \in D(\mathbf{x})} \left\{ \bar{r}(\mathbf{x}) + \left(\sum_{j\in\alpha} \lambda_j d((1_j,\mathbf{x})) + \sum_{1}^{N} x_k \mu_k d((0_k,\mathbf{x})) \right)/q(\mathbf{x}) \right\}. \tag{3}$$

όπου στη (3) έχουμε διαιρέσει το $\psi(\mathbf{x})$ με τον όρο $q(\mathbf{x})$ δεδεμένου ότι ο αναμενόμενος χρόνος παραμονής στη κατάσταση $\mathbf{x}$ είναι ίσος με $1/q(\mathbf{x})$, (βλπ. (α) παραπάνω).

Οι επόμενες απλοποιήσεις - μετατροπή είναι τεχνικές που εφαρμόζονται συχνά σε προβλήματα του είδους που εξετάζουμε. Συγκεκριμένα, ας παρατηρήσουμε ότι πολιτικές που αφήνουν μονάδες συντήρησης αχρησιμοποίητες ενώ υπάρχουν μη λειτουργούντα στοιχεία δεν είναι βέλτιστες, έστω Π_0 το σύνολο που περιέχει όλες τις υπόλοιπες πολιτικές. Πολιτικές στό Π_0 έχουν την εξής ιδιότητα: "όλες οι καταστάσεις της διαδικασίας που περιγράφει τη μεταβολή της κατάστασης του συστήματος στο χρόνο επικοινωνούν με τη κατάσταση $\mathbf{1}=(1,\ldots,1)$ ". Ετσι, το μέσο κόστος του συστήματος που αντιστοιχεί σε πολιτικές στο Π_0 είναι το ίδιο για όλες τις αρχικές καταστάσεις

$$\phi(\pi,\mathbf{x}) = g(\pi)\,,\ \forall\, \mathbf{x} \in S\,,\ \forall\, \pi \in \Pi_0\,, \tag{4}$$

και η εξισώσεις βελτιστοποίησης απλοποιούνται ως εξής

$$g/q(\mathbf{x}) + d(\mathbf{x}) = \max_{\alpha \in D(\mathbf{x})} \left\{ \bar{r}(\mathbf{x}) + \left(\sum_{j\in\alpha} \lambda_j d((1_j,\mathbf{x})) + \sum_{1}^{N} x_k \mu_k d((0_k,\mathbf{x})) \right)/q(\mathbf{x}) \right\}. \tag{5}$$

Ας υποθέσουμε ότι r^{μ} είναι ένας αριθμός μεγαλύτερος από τις παραμέτρους $q(\mathbf{x})\ \ \forall\, \mathbf{x}$. Είναι εύκολο να δούμε ότι οι (5) μπορούν να γραφούν ισοδύναμα ως εξής

$$g + d(\mathbf{x}) = \max_{\alpha \in D(\mathbf{x})} \left\{ \left(r(\mathbf{x}) + (r^{\mu} - q(\mathbf{x}))d(\mathbf{x}) + \sum_{j\in\alpha} \lambda_j d((1_j,\mathbf{x})) + \sum_{1}^{N} x_k \mu_k d((0_k,\mathbf{x})) \right)/r^{\mu} \right\}. \tag{6}$$

Οι (6) ονομάζονται εξισώσεις που αντιστοιχούν στην διαδικασία με ομοιόμορφους ρυθμούς (uniformized Chain), καί χρησιμοποιήται στο παρακάτω πρόγραμμα.

Παρατηρήσεις. 1. Στο παρακάτω πρόγραμμα, θεωρούμε ότι το σύστημα έχει δομή "K από N", δηλαδή το σύστημα λειτουργεί όταν τουλάχιστον K από τα N στοιχεία του λειτουργούν. Επίσης εξετάζουμε τη περίπτωση όπου $R = 1$.

2. Στο πρόγραμμα, χρησιμοποιείται η μέθοδος των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών για την επίλυση των (6).

3. *Στο πρόγραμμα, για ευκολία προγραμματισμού συμβολίζουμε την κατάσταση του συστήματος με το δεκαδικό αριθμό που αντιστοιχεί στην κατάσταση x αν τη θεωρήσουμε σαν δυαδικό αριθμό.*

Κώδικας Προγράμματος – Γλώσσα FORTRAN

```
C       maintenance of a series system by a single repairman

C       solution of dp equations



        program series


        parameter(n1=10,nn21=2**n1)

C         n   : number of components  (each can be either in state 0 or 1)

C       nn2  : number of states


        real g

        real h(nn21)

        integer*2 pol(nn21)

        real am(n1)

        real al(n1)

        integer is(n1)

        real p(nn21,n1)
        real r(nn21)

        real x(2,nn21)

        real d(nn21)

        integer ind(nn21,n1)


C       am   :   failure rates

C       al   :   repair rates


        integer rn,ch,n,nn2
```

```
      character fname1*10,fname2*10

      common n,nn2,ru,rn,err,ch

C     DATA INPUT
      write(*,'(a,$)') ' number of components (<=10) : '
      read(*,'(i3)')   n
      write(*,'(a,$)') ' input k : '
      read(*,'(i3)')   kk
      nn2=2**n
      write(*,'(a,$)') ' filename of input data : '
      read(*,'(a10)')   fname1
      open (1,file=fname1,status='old')
      do 5 i=1,n
         read(1,1100) al(i),am(i)
5     continue
1100  format(f5.2,2x,f5.2)
      write(*,'(a,$)') ' uniformization factor : '
      read(*,'(f6.2)')  ru
      write(*,'(a,$)') ' filename for results : '
      read(*,'(a10)')   fname2
      open (2,file=fname2,status='new')
      write(*,'(a,$)') ' approximation in value determination : '
      read(*,'(f6.4)') err

C     INITIAL VALUES
      data g/0/,h/nn21*0/
      rn=0
      do 60 i=1,nn2-1
         call dectovec(n1,i-1,is)
```

```
                nk=0

                do 63 j=1,n
                    if(is(j).eq.1) nk=nk+1
63              continue
                if(nk.ge.kk) r(i)=1
                do 65 j=1,n

                    if(is(j).eq.0) then

                        pol(i)=j

                        goto 60

                    endif

65              continue

60          continue
```

```
C           MAIN  LOOP

40          ch=0

            rn=rn+1

            call vdet(n1,nn21,is,p,r,x,d,ind,g,h,pol,al,am)

            call pimp(n1,nn21,is,r,g,h,pol,al,am)

            if(ch.eq.0) then

                goto 18

            else

                goto 40

            endif
```

```
C           OUTPUT

18          write(2,'(a,$)') ' failure rates : '

            do 50 i=1,n

                write(2,1000) am(i)

50          continue
```

```
      write(2,*)
      write(2,'(a,$)') ' repair  rates : '
      do 51 i=1,n
      write(2,1000) al(i)
51    continue

      write(2,'(1x,"optimal policy  :  g=",f6.4)')  g
      write(2,*)
      write(2,'(1x,"components")')
      write(2,*)
      do 55 j=1,n
         write(2,'(1x,i1,$)')  j
55    continue
      do 70 i=1,nn2
         call dectovec(n,i-1,is)
         do 75 l=1,n
            write(2,'(1x,i1,$)') is(l)
75       continue
         write(2,1020) pol(i),h(i)
70    continue
1000  format(2x,f6.2,$)
1020  format(1x,'      a='i2,'       h=',f7.3)
      end
```

C END OF MAIN PROGRAM

C Value Determination Procedure

```
      subroutine vdet(n1,nn21,is,p,r,x,d,ind,g,h,pol,al,am)
```

```
      integer rn,ch

      common n,nn2,ru,rn,err,ch


      real g

      real h(nn21)

      integer*2 pol(nn21)

      integer new,old

      real al(n1)

      real am(n1)

      integer is(n1)

      real p(nn21,n1)
      real r(nn21)

      real x(2,nn21)

      real d(nn21)

      integer ind(nn21,n1)


      icount=0

C     Computation of the coefficients in the recursive equation  g + h = r + A h

      do 100 i=1,nn2-1


        p(i,1)=al(pol(i))/ru

        ind(i,1)=i+2**(pol(i)-1)

        d(i)=al(pol(i))

        k=2

        call dectovec(n1,i-1,is)


        do 110 j=1,n

          if(is(j).eq.1) then

            p(i,k)=am(j)/ru

            ind(i,k)=i-2**(j-1)
```

```
                d(i)=d(i)+am(j)
                k=k+1
             endif
110       continue

          do 117 j1=k,n
             p(i,j1)=0
117       continue

100    continue

       do 112 j=1,n
          p(nn2,j)=am(j)
          ind(nn2,j)=nn2-2**(j-1)
112    continue
```

C Succesive Approximations for the solution of g + h = r + A h

```
       do 120 i=1,nn2
          x(1,i)=0
          x(2,i)=0
120    continue

       old=1
       new=2
```

C Computation of the New Approximation

```
125    icon=0

       do 130 i=1,nn2-1

          x(new,i)=r(i)/ru

          do 140 j=1,n
```

```
            if((p(i,j).eq.0).or.(ind(i,j).eq.nn2)) goto 140
            x(new,i)=x(new,i)+p(i,j)*x(old,ind(i,j))
140       continue

          x(new,i)=x(new,i)-x(old,nn2)/ru+(1-d(i)/ru)*x(old,i)
```
C Convergence Test
```
          if(abs(x(new,i)-x(old,i)).ge.err) icon=1

130     continue

        x(new,nn2)=1
        do 150 j=1,n
          x(new,nn2)=x(new,nn2)+p(nn2,j)*x(old,ind(nn2,j))
150     continue
         if(abs(x(new,nn2)-x(old,nn2)).ge.err) icon=1

         if(icon.eq.1) then
          ii=old
          old=new
          new=ii
          icount=icount+1
          goto 125
        endif

        g=x(new,nn2)
        write (*,2000) rn,g,icount
2000    format(1x,'n=',i3,'     g=',f6.4,'     after ',i4,' iterations')
        h(nn2)=0
        do 160 i=1,nn2-1
          h(i)=x(new,i)
160     continue
        return
```

```
      end

C     Policy Improvement Procedure

      subroutine pimp(n1,nn21,is,r,g,h,pol,al,am)

      real g
      integer rn,ch
      real h(nn21)
      integer*2 pol(nn21)
      real al(n1)
      real am(n1)
      integer is(n1)
      common n,nn2,ru,rn,err,ch

      ch=0

      do 600 i=1,nn2-1

      cm=0
      sm=r(i)/ru
      smx=-1e30
      mx=0
      call dectovec(n,i-1,is)

      do 610 j=1,n
        if(is(j).eq.0) goto 610
        cm=cm+am(j)
        sm=sm+am(j)*h(i-2**(j-1))
610   continue
```

```
C       Computation of Maximum Right Hand Side for All Alternative Decisions
        do 620 j=1,n
          if(is(j).eq.1) goto 620
          smt=(al(j)*h(i+2**(j-1))+sm)/ru+(1-(cm+al(j))/ru)*h(i)
          if(smt.gt.smx) then
            smx=smt
            mx=j
          endif
620     continue

C       Check of Optimality  (The action is otpimal if it attains the maximum computed above.)
        if(abs(g/ru+h(i)-smx).gt.err) then
          pol(i)=mx
          ch=1
        endif

600     continue

        return
        end



C       Decimal to Vector State Conversion

        subroutine dectovec(n1,i,is)

        integer is(n1)

        common n,nn2,ru,rn,err,ch
```

```
      do 800 j=n,1,-1
         is(j)=i/2**(j-1)
         i=i-is(j)*2**(j-1)
800   continue
      return
      end
```

ΕΦΑΡΜΟΓΗ 4. *Εδώ δίνουμε ένα πρόγραμμα FORTRAN που υπολογίζει τη βέλτιστη πολιτική για το πρόβλημα του άπειρου ορίζοντα με πιθανοθεωρητική δυναμική, και κριτήριο την ελαχιστοποίηση του αναμενόμενου συνολικού εκπτωτικού κόστους, για πεπερασμένους χώρους καταστάσεων* S *και σύνολα αποφάσεων* D. *Ο υπολογισμός γίνεται με τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο τιμών. Δηλαδή έχουμε να λύσουμε με τη βοήθεια του Η.Υ. την παρακάτω εξίσωση βελτιστοποίησης*

$$v(i) \;=\; \min_{\alpha \in D(i)} \{\, r(i,\alpha) + \beta \sum_{j\in S} p(j/i,\alpha)\, v(j) \,\}\,, \quad i \in S$$

για την οποία ξέρουμε ότι

$$v(i) \;=\; \lim_{n \to \infty} v^{(n)}(i)\,, \quad \text{και} \quad \alpha^{0}(i) \;=\; \lim_{n \to \infty} \alpha^{(n)}(i)$$

όπου $\alpha^{0}(i)$ *είναι οι αποφάσεις μιας βέλτιστης πολιτικής.*

Οι διαδοχικές προσεγγίσεις είναι

$$v^{(n)}(i) \;=\; \min_{\alpha \in D(x)} \{\, r(i,\alpha) + \beta \sum_{j\in S} p(j/i,\alpha)\, v^{(n-1)}(j) \,\}\,.$$

Η εφαρμογή αποτελείται από δύο ξεχωριστά προγράμματα FORTRAN, από τα οποία το πρώτο διαβάζει τα δεδομένα που δίνει ο χρήστης ($p(j/i,\alpha)$ και $r(i,\alpha)$, $i,j \in S$, $\alpha \in D(i)$) , και τα τοποθετεί σε ένα αρχείο , ενώ το δεύτερο εφαρμόζει τον αλγοριθμο των διαδοχικών προσεγγίσεων χρησιμοποιώντας τα περιεχόμενα του προηγούμενου αρχείου.
Τονίζουμε ότι με ελάχιστες αλλαγές το δεύτερο πρόγραμμα είναι δυνατό να προσαρμοστεί για την επίλυση προβλημάτων του τύπου της πρώτης μετάβασης ή βέλτιστης διακοπής.

```
C          PROGRAM  I  :  INPUT OF DATA
C          Succesive Aproximations:  Write  P(I,J,K) ,  R(I,K)  TO DISK


           DIMENSION P(100,100,10),R(100,10)
           OPEN(3,FILE='SA.P1',STATUS='NEW',ACCESS='SEQUENTIAL')
           OPEN(4,FILE='SA.R1',STATUS='NEW',ACCESS='SEQUENTIAL')
           WRITE(*,89)
89         FORMAT(2X,'PLEASE INPUT DIMENSIONS N, K')
           READ (*,12) N,L
12         FORMAT(I3,I3)
           WRITE(*,99)
99         FORMAT(2X,'PLEASE INPUT DISCOUNT FACTOR')
           READ (*,14) D
14         FORMAT(F6.4)
           WRITE (*,13) N,L,D
13         FORMAT(2X,'N=',I3,'L=',I3,'D=',F5.3)
           WRITE (3,23) N,L,D
23         FORMAT(2X,2I3,F6.3)

C         Input  Pij(k)  for states i,j and decsions k.
           DO  102 K=1,L
              DO 103 I=1,N
                 DO 104 J=1,N
                    WRITE (*,201) I,J,K
201                 FORMAT(1X,'Input P(',I3,','I3,',',I2,')')
                    READ (*,202) P(I,J,K)
202                 FORMAT(F10.3)
                    WRITE(3,212) I,J,K,P(I,J,K)
212                 FORMAT(3I3,F10.3)
104              CONTINUE
103           CONTINUE
102        CONTINUE

C         Input  Ri(k)  for states i and decisions k
           DO 105 K=1,L
              DO 106 I=1,N
                 WRITE (*,301) I,K
301              FORMAT(1X,'Input R(',I3,',',I2,')')
                 READ (*,302) R(I,K)
302              FORMAT(F10.3)
                 WRITE(4,312) I,K,R(I,K)
312              FORMAT(2I3,F10.3)
106           CONTINUE
105        CONTINUE
           DO 402 K=1,L
              DO 403 I=1,N
                 DO 404 J=1,N
                    WRITE (*,241) I,J,K,P(I,J,K),I,K,R(I,K)
```

```
241          FORMAT(1X,'P(',I3,',',I3,',',I3,')=',F7.2,4X,'R(',I3,',',I3,')=',F7.2)
404        CONTINUE
403      CONTINUE
402    CONTINUE
       STOP
       END
```

C PROGRAM II : SOLUTION OF THE OPTIMALITY EQUATIONS

C Suceesive Approximations for the Expected Discounted Cost

```
       DIMENSION V(20),W(20),P(20,20,10),MK(20)

       REAL A,AA,R(20,10)

       OPEN(3,FILE='SA.P1',STATUS='OLD',ACCESS='SEQUENTIAL')

       OPEN(4,FILE='SA.R1',STATUS='OLD',ACCESS='SEQUENTIAL')

       OPEN(5,FILE='SB.t1',STATUS='NEW',ACCESS='SEQUENTIAL')

       REWIND 3

       REWIND 4

       IJ=0
```

C Data Input from the created files

```
       WRITE(*,89)

89     FORMAT(2X,'PLEASE INPUT MM')

       READ (*,12) M

12     FORMAT(I2)


       READ (3,23) N,L,D

23     FORMAT(2X,2I3,F6.3)


       D1=D**(M-1)

       D=D**M

       A=100.

       AA=0.
```

```
          DO  102 K=1,L
             DO 103 I=1,N
                DO 104 J=1,N
                   READ (3,212) I,J,K,P(I,J,K)
212                FORMAT (3I3,F10.3)
                   AA=D*P(I,J,K)
                   WRITE(*,111)A
111                FORMAT(F8.4)
                   IF(AA.EQ.0.) GOTO 104
                   IF(AA.LT.A) A=AA
104             CONTINUE
103          CONTINUE
102       CONTINUE



          DO 225 K=1,L
             DO 226 I=1,N
                READ(4,312) I,K,R(I,K)
                R(I,K)=R(I,K)*D1
312             FORMAT(2I3,F10.3)
226          CONTINUE
225       CONTINUE

C         Write Data in the Screen
          DO 602 K=1,L
             DO 603 I=1,N
                DO 604 J=1,N
                   WRITE (5,41) I,J,K,P(I,J,K),I,K,R(I,K)
41                 FORMAT(1X,'P(',I3,','I3,',',I3,')=',F5.2,4X,'R(',I3,',',I3,')=',F5.2)
604             CONTINUE
```

```
603         CONTINUE
602     CONTINUE


        DO 402 K=1,L
            DO 403 I=1,N
                DO 404 J=1,N
                    WRITE (*,241) I,J,K,P(I,J,K),I,K,R(I,K)
                    WRITE (5,241) I,J,K,P(I,J,K),I,K,R(I,K)
241                 FORMAT(1X,'P(',I3,','I3,',',I3,')=',F10.5,4X,'R(',I3,',',I3,')=',F10.5)
404             CONTINUE
403         CONTINUE
402     CONTINUE
```

```
C       Initial Values
        DO 113 I=1,N
            MK(I)=0
            W(I)=0.0
113     CONTINUE
```

```
C       Main Loop
        IL=0

        DO 121 II=1,200
            IF (IL.EQ.1) GOTO 121

            DO 122 I=1,N
                X2=100000000.
```

```
C       Computation of right hand side for alternative actions

            DO 123 K=1,L

              S=0.0

              DO 124 J=1,N

                S=S+P(I,J,K)*W(J)

124           CONTINUE

              S=(S*D)+R(I,K)

              IF (S.LT.X2) GOTO 331

              GOTO 123

C       New Optimal Acion for State i is k
331           X2=S

              MK(I)=K

123         CONTINUE

            V(I)=X2

            WRITE(*,401) I,V(I),MK(I)

401         FORMAT(1X,'V(',I3,')=',F18.6,4X,'MK=',I3)

            PAUSE

122       CONTINUE

          IF=0

C       Test of Convergence — Replacement of the last approximation
          DO 420 J=1,N

            IF (W(J).NE.V(J)) IF=1

            W(J)=V(J)

420       CONTINUE

          IF (IF.EQ.0) goto 555

          GOTO 121

555       IL=1
```

```
      IJ=II
121   CONTINUE

C     Output
      DO 528 I=1,N
        IF(IJ.EQ.0) IJ=-1
        WRITE(*,504) I,V(I),MK(I),IJ
        WRITE(5,504) I,V(I),MK(I),IJ
504     FORMAT(1X,'V(',I3,')=',F10.5,4X,'MK=',I3,'  #ITER.=',I5)
528   CONTINUE
      STOP
      END
```

6. Ασκήσεις

Ασκηση 1. *Ο χρόνος ζωής (σε έτη) ενός μηχανήματος είναι τυχαία μεταβλητή που μπορεί να πάρει τις τιμές 1,2,3, με αντίστοιχες πιθανότητες $p_1 = 1/4$, $p_2 = 1/2$, $p_3 = 1/4$. Στο τέλος κάθε έτους έχουμε δύο εναλλακτικές λύσεις εφόσον το μηχάνημα λειτουργεί.*

α) *Να το αφήσουμε να λειτουργήσει ένα χρόνο ακόμη, ή*

β) *να το αντικαταστήσουμε με ένα καινούργιο, παρόλο που δεν έχει χαλάσει.*

Η αντικατάσταση ενός μηχανήματος πριν χαλάσει κοστίζει 10 μονάδες. Αν το μηχάνημα χαλάσει κατά τη διάρκεια της λειτουργίας του, τότε εκτός από το κόστος αντικατάστασης με καινούργιο, έχουμε και παραγωγή ελαττωματικών προιόντων, με αποτέλεσμα το συνολικό κόστος σ' αυτήν την περίπτωση να ανέρχεται σε 24 μονάδες.

Ζητείται να βρεθεί η πολιτική που ελαχιστοποιεί το αναμενόμενο μέσο κόστος αντικαταστάσεων του μηχανήματος.

Υπόδειξη. Βολεύει από άποψη πράξεων να πάρετε αρχική πολιτική αυτήν που αντικαθιστά πάντοτε το μηχάνημα, εκτός αν αυτό είναι εντελώς καινούργιο.

Ασκηση 2. *Στο παιχνίδι "το κυνήγι του θησαυρού" οι όροι είναι οι εξής. Στην αρχή του παιχνιδιού ο παίκτης ξέρει ότι η πιθανότητα να υπάρχει ο θησαυρός είναι ίση με α. Σε καθένα από τα επόμενα βήματα, μπορεί είτε να δοκιμάσει να βρει το θησαυρό, είτε να σταματήσει το παιχνίδι. Κάθε δοκιμή έχει κόστος c. Αν υπάρχει ο θησαυρός, η πιθανότητα να βρεθεί με μια δοκιμή είναι ίση με p. Επίσης, αν βρεθεί ο θησαυρός, το κέρδος του παίκτη είναι R και το παιχνίδι τελειώνει. Αν ο παίκτης αποφασίσει να σταματήσει τις δοκιμές δεν πληρώνει τίποτε, χάνει όμως την ευκαιρία να κερδίσει τις R μονάδες.*

Το πρόβλημα είναι να βρεθεί η πολιτική που μεγιστοποιεί το αναμενόμενο συνολικό κέρδος του παίκτη.

Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

Υπόδειξη. Θεωρήστε σαν κατάσταση σε κάθε βήμα την πιθανότητα (posterior probability) να υπάρχει ο θησαυρός, και βρείτε πώς αλλάζει η πιθανότητα αυτή ύστερα από κάθε αποτυχημένη προσπάθεια, χρησιμοποιώντας τον τύπο του Bayes.

Ασκηση 3. *Θεωρούμε το δίκτυο του παρακάτω σχήματος*

Οι πλάγιοι αριθμοί συμβολίζουν τα αναμενόμενα κόστη για κάθε ζεύγος κατάστασης – απόφασης, ενώ τα τετράγωνα είναι κόμβοι πιθανότητας. Οι αριθμοί μετά τους κόμβους αυτούς συμβολίζουν τις αντίστοιχες πιθανότητες μετάβασης. Οπου δεν υπάρχει κόμβος πιθανότητας, η αντίστοιχη μετάβαση γίνεται με πιθανότητα 1.

Ζητείται η βέλτιστη πολιτική για το πρόβλημα του άπειρου ορίζοντα, όταν το κριτήριο είναι η ελαχιστοποίηση του αναμενόμενου συνολικού εκπτωτικού κόστους $E\left\{\sum_{t=0}^{\infty}\beta^t c(X_t,A_t)\right\}$*, με $\beta = 0.9$ και $X_0 = 1$.*

Ασκηση 4. *Θεωρούμε το παρακάτω στοχαστικό δυναμικό σύστημα*

$$x(t+1) = g\,x(t) + h\,\alpha(t) + z(t), \quad t = 0,1,\ldots,N-1,$$

όπου g, h σταθερές, $x(t)$ συμβολίζει την κατάσταση κατά τη χρονική περίοδο t, $\alpha(t)$ την απόφαση κατά τη χρονική περίοδο t, ενώ οι $z(t)$, $t = 0,1,\ldots,N-1$, είναι ανεξάρτητες και όμοια κατανεμημένες τυχαίες μεταβλητές με μέση τιμή 0 και διασπορά $\sigma^2(t) = \sigma^2\alpha^2(t)$, $\sigma^2 =$ σταθερό. Η παραδοχή αυτή σημαίνει ότι όσο μεγαλύτερο πάρουμε το $\alpha(t)$, τόσο μεγαλύτερη είναι η "τυχαιότητα" του συστήματος. Υποθέτουμε ότι σε κάθε χρονική περίοδο t αποφασίζουμε για την τιμή της $\alpha(t)$ πριν παρατηρήσουμε την $z(t)$, παίρνοντας όμως υπόψη την κατάσταση $x(t)$. Ζητάμε να βρούμε την πολιτική $\alpha(t)$, $t = 1,2,\ldots,N-1$, που ελαχιστοποιεί την αναμενόμενη τιμή του συνολικού κόστους

$$\sum_{t=0}^{N-1} \left\{ bx^2(t) + c\alpha^2(t) \right\} + lx^2(N).$$

Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

Ασκηση 5. *Ενα πολιτικό κόμμα σχεδιάζει οικονομική εξόρμηση και σκοπεύει να διαθέσει χρηματικό ποσό S (το πολύ) για διαφημιστική εκστρατεία διάρκειας N εβδομάδων. Από προηγούμενες στατιστικές εκτιμήσεις είναι γνωστό ότι αν στην αρχή της k εβδομάδας έχει ήδη συγκεντρωθεί ποσό x, τότε η πιθανότητα μέσα στη διάρκεια της εβδομάδας να συγκεντρωθεί επιπλέον ποσό x_k, αν διατεθεί για διαφήμιση ποσό s_k, είναι ίση με $p_k(x_k/x,s_k)$, όπου p_k, $k = 1,\ldots,N$, είναι γνωστές συναρτήσεις. Το πρόβλημα είναι να βρεθεί η διαφημιστική στρατηγική που μεγιστοποιεί το αναμενόμενο συνολικό ποσό που θα συγκεντρωθεί, συμπεριλαμβανομένων και των τυχόν αδιάθετων ποσών από το διαφημιστικό κεφάλαιο S.*

α) *Δείξτε ότι η βέλτιστη στρατηγική μπορεί να βρεθεί χρησιμοποιώντας μεθόδους Δυναμικού Προγραμματισμού.*

β) *Δείξτε πώς πρέπει να αλλάξει η μοντελοποίηση στο (α), όταν υπάρχει κεφάλαιο S στην αρχή της πρώτης εβδομάδας διαθέσιμο για διαφημίσεις, αλλά ακόμη το κόμμα είναι διατεθειμένο να διαθέσει όσα επιπλέον χρήματα μαζέψει κατα τη διάρκεια της διαφημιστικής εκστρατείας για περαιτέρω διαφήμιση μέχρι το τέλος των N εβδομάδων.*

Ασκηση 6. *Ο διευθυντής ενός σούπερ μάρκετ αντιμετωπίζει το εξής πρόβλημα σε σχέση με την εξυπηρέτηση των πελατών. Υπάρχουν στατιστικά στοιχεία σχετικά με τον αριθμό των πελατών που μπαίνουν στο κατάστημα κάθε χρονική περίοδο (ας πούμε 15 λεπτών), όπως επίσης και για τον αριθμό των πελατών που εξυπηρετούνται μέσα σε μία περίοδο. Ο τελευταίος αριθμός εξαρτάται τόσο από τον αριθμό πελατών που βρίσκονται στο κατάστημα, όσο και από τον αριθμό των ανοιχτών ταμείων. Ετσι έχει εκτιμήσει ότι $p_t(y)$ είναι η πιθανότητα y πελάτες να μπουν στο κατάστημα την περίοδο t (όπου για λόγους απλότητας υποθέτουμε ότι μπαίνουν όλοι μαζί στην αρχή της περιόδου). Επίσης $q_t(m/x,s)$ είναι η πιθανότητα να εξυπηρετηθούν m πελάτες στη διάρκεια της περιόδου t, δεδομένου ότι στην αρχή της περιόδου υπάρχουν x πελάτες στο κατάστημα και s ανοιχτά ταμεία. Εστω ότι w είναι το κόστος ανά περίοδο για κάθε ανοιχτό ταμείο. Επίσης ο διευθυντής έχει μια σχετική εκτίμηση για το κόστος που προκύπτει από την αναμονή των πελατών και αυτό είναι h_t για κάθε πελάτη που μένει στο κατάστημα στο τέλος της περιόδου t.*

α) *Να βρεθεί ένα μοντέλο Δυναμικού Προγραμματισμού για τον υπολογισμό της βέλτιστης πολιτικής κατά τις ώρες αιχμής, που υπολογίζονται σε N περιόδους, ούτως ώστε να ελαχιστοποιείται το αναμενόμενο συνολικό κόστος.*

β) *Δείξτε πώς αλλάζει η μοντελοποίηση στο (α) όταν υπάρχει πρόσθετο κόστος K (setup cost) για κάθε ταμείο που ανοίγει στην αρχή μιας περιόδου, δηλαδή έχουμε επιπλέον κόστος rK στη διάρκεια μιας περιόδου, αν r ταμεία ανοίξουν στην αρχή αυτής της περιόδου.*

Ασκηση 7. *Μια μεγάλη εταιρεία συνεργάζεται με μια τράπεζα για ευκολότερη διεκπεραίωση των οικονομικών της συναλλαγών. Οι οικονομικοί όροι της συνεργασίας αυτής είναι οι εξής.*

Οι συναλλαγές της εταιρείας με την τράπεζα (καταθέσεις και αναλήψεις) μπορούν να γίνονται στην αρχή κάθε μέρας, και για κάθε κατάθεση ή ανάληψη η τράπεζα χρεώνει ποσό ίσο με D ή W αντίστοιχα. Κατά τη διάρκεια της μέρας το υπόλοιπο του λογαριασμού της εταιρείας έχει διακυμάνσεις που οφείλονται είτε σε καταθέσεις τρίτων στο όνομα της εταιρείας, είτε σε αναλήψεις από τρίτους επιταγών που η εταιρεία έχει εκδώσει. Αν το υπόλοιπο πέσει κάτω από μηδέν, η τράπεζα δανείζει αυτόματα το λογαριασμό με τα απαιτούμενα ποσά, με επιτόκιο r ανά ημέρα.

Η εκτίμηση της εταιρείας για το "κόστος ευκαιρίας" είναι s ανά ημέρα (δηλαδή θα μπορούσε να κερδίσει ποσό s κάθε μέρα από κάθε χρηματική μονάδα που θα επένδυε αλλού αντί να την καταθέσει στην τράπεζα). Υποθέτουμε ότι τα r και s υπολογίζονται πάνω στο υπόλοιπο του λογαριασμού, όπως αυτό έχει διαμορφωθεί στο τέλος της μέρας.

Εστω Q το ποσό κατά το οποίο μεταβάλλεται το υπόλοιπο του λογαριασμού στη διάρκεια μιας μέρας (το Q μπορεί να είναι είτε θετικό είτε αρνητικό) εξ αιτίας των συναλλαγών της εταιρείας με τρίτους. Το Q είναι τυχαία μετβλητή που ακολουθεί μία γνωστή συνάρτηση κατανομής $p_t(q)$, κατά την t ημέρα.

α) *Ζητείται να βρεθεί η πολιτική καταθέσεων/αναλήψεων που ελαχιστοποιεί το αναμενόμενο συνολικό κόστος της εταιρείας, σε μια περίοδο N ημερών.*

β) *Δείξτε πώς μεταβάλλεται το μοντέλο στο (α) όταν τα κόστη συναλλαγών W και D, όπως επίσης και οι τόκοι από τυχόν δανεισμό αφαιρούνται από το λογαριασμό αντί να πληρώνονται απευθείας.*

Ασκηση 8. *Ο διευθυντής προσωπικού μιας εταιρείας θέλει να προσλάβει ένα υπάλληλο καί παίρνει συνεντεύξεις από τους υποψήφιους. Απο πρίν έχει διακρίνει τρείς κατηγορίες υποψηφίων, μέτριους (1), καλούς (2) και πολυ καλούς (3). Επίσης έχει προηγούμενες εκτιμήσεις για τα ποσοστά των υποψηφίων που ανήκουν σε κάθε κατηγορία, και επομένως άν πάρει μια συνέντευξη η πιθανότητα ο υποψήφιος να ανήκει στην κατηγορία i είναι ίση με p_i, $i = 1,2,3$. Ο διευθηντής δεν μπορεί να δεί πάνω από N υποψήφιους. Αν κρίνει κάποιον πολύ καλό φυσικά θα τον προσλάβει, ενώ αν τον κρίνει μέτριο δεν χάνει τίποτα να εξετάσει και τον επόμενο. Το πρόβλημα βρίσκεται στο τι πρέπει να κάνει αν κρίνει κάποιον καλό. Αν τον διώξει, μπορεί όλοι οι επόμενοι να είναι μέτριοι, ενώ αν τον προσλάβει χάνει την ευκαιρία να προσλάβει ένα πολύ καλό μετά. Εδώ υποθέτουμε οτι αν δέν προσλάβει ενα υποψήφιο μόλις πάρει την συνέντευξη, δεν εχει ευκαιρία να τον προσλάβει αργότερα. Αν q_i είναι η "άξία" που εχει για το διευθυντή η πρόσληψη υπαλλήλου της κατηγορίας i, όπου προφανώς $q_1 < q_2 < q_3$, να βρεθεί η πολιτική πρόσληψης που μεγιστοποιεί την αναμενόμενη αξία απ' όλη τη διαδικασία. Ζητείται να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού για τις εξής δυνατές περιπτώσεις*

α) *Υπάρχει κόστος c για κάθε συνέντευξη.*

β) *Ο αριθμός των υποψηφίων που μπορεί να δεί ο διευθυντής είναι απεριόριστος $(N = \infty)$, και $c > 0$. Ποιά είναι η βέλτιστη πολιτική σ'αυτή την περίπτωση όταν $c = 0$;*

γ) *$N = \infty$, $c > 0$, και υπάρχει και εκπτωτικός παράγοντας β, $0 \leq \beta \leq 1$.*

δ) *Γράψτε το πρωτεύον και το δυαδικό γραμμικό πρόγραμμα που αντιστοιχεί στο γ).*

ε) *Υπολογίστε τη βέλτιστη πολιτική στο γ) όταν $q_1 = 1$, $q_2 = 2$, $q_3 = 3$, $p_1 = 0.3$, $p_2 = 0.5$, $p_3 = 0.2$, $c = 0.15$, $\beta = 1$.*

στ) *Αν ο διευθυντής έχει την ευκαιρία να αναβάλει την απόφαση σχετικά μ' έναν υποψήφιο, μέχρι να δεί και άλλους, σε ποιά περίπτωση θα ήθελε να εκμεταλλευτεί αυτή τη δυνατότητα; Βρείτε πώς αλλάζει η μοντελοποίηση στο γ) με την επιπλέον υπόθεση. Επίσης απαντήστε ξανά στα δ) και ε).*

(Γενική Υπόδειξη. Σαν κατάσταση x_t στο μοντέλο δυναμικού προγραμματισμού μπορείτε να χρησιμοποιήσετε την καλύτερη κατηγορία υποψήφιου που έχει δεί ο διευθυντής στις $t-1$ προηγούμενες συνεντεύξεις).

Ασκηση 9. Κάθε μήνα η τράπεζα Ελλάδος ορίζει το ποσό πού πρέπει να δανεισθεί από τις διεθνείς χρηματαγορές, για να ικανοποιήσει τις ζητήσεις για δάνεια των πελατών της. Οι συνολικές απαιτήσεις τον μήνα t είναι D_t. Η D_t είναι τυχαία μεταβλητή που ακολουθεί κατανομή $p(d_t)$, υποθέτουμε όμως οτι η τιμή $D_t = d_t$ γίνεται γνωστή στην αρχή της περιόδου t, πρίν παρθούν οι αντίστοιχες αποφάσεις. Κάθε μήνα η τράπεζα μπορεί να δανεισθεί κεφάλαια για 1, 2 ή 3 μήνες, σε ποσότητες α_1, α_2, α_3. Απ' όσα είπαμε παραπάνω, η ποσότητα $\alpha_1 + \alpha_2 + \alpha_3$ μπορεί να χρησιμοποιειθεί για να καλύψει τη ζητηση $d_t = D_t$, η $\alpha_2 + \alpha_3$ για την D_{t+1}, και η α_3 για την D_{t+2}. Υποθέτουμε ακόμη οτι το κόστος δανεισμού ποσού α_i, $i = 1,2,3$ την περίοδο t, εξαρτάται από μια παράμετρο R_t, και δίνεται από τη συνάρτηση $c_{ti}(\alpha_i/R_t)$, όπου η R_t είναι τυχαία μεταβλητή με κατανομή $q_t(r)$. Η τιμή της $R_t = r$ γίνεται γνωστή στήν αρχή της περιόδου t πρίν ληφθούν οι αποφάσεις. (Οι μελλοντικές τιμές R_{t+1}, κ.λ.π. παραμένουν άγνωστες). Υποθέτουμε ότι όλες οι τυχαίες μεταβλητές είναι ανεξάρτητες μεταξύ τους.

α) Να διατυπωθεί ένα μοντέλο Δυναμικού Προγραμματισμού για την ελαχιστοποίηση του αναμενόμενου συνολικού κόστους, υπό τον περιορισμό οτι όλες οι απαιτήσεις για δάνεια πρέπει να ικανοποιηθούν. Ο εκπτωτικός παράγοντας είναι ίσος με β, $0 < \beta < 1$.
β) Βρείτε πως απλοποιείται (αν απλοποιείται) το (α) όταν η τιμή της D_t γίνεται γνωστή μετά την απόφαση της περιόδου t. Υποθέστε ότι όταν η D_t ξεπερνάει τα διαθέσιμα κεφάλαια της Τράπεζας της Ελλάδος προς δανεισμό, υπάρχει μοναδιαίο κόστος απώλειας ίσο με c.

Ασκηση 10. Θεωρείστε τη γενική εξίσωση βελτιστοποίησης άπειρου ορίζοντα με πιθανοθεωρητική δυναμική για το αναμενόμενο εκπτωτικό κόστος

$$v(x) = \min_{x \in D(x)} \{ c(x,\alpha) + \beta \sum_{x' \in X} p[X_{t+1} = x'/X_t = x, A_t = \alpha]\, \upsilon(x') \} .$$

Υποθέστε οτι έχετε εκτιμήσεις για τα άνω και κάτω φράγματα U_x, L_x των άγνωστων τιμών $v(x)$: $L_x < v(x) < U_x$. Δείξτε ότι δεν είναι ποτέ βέλτιστο να χρησιμοποιήσουμε την απόφαση $\alpha' \in D(x)$ αν

$$c(x,\alpha') + \beta \sum_{x' \in X} p(x'/x, \alpha') \cdot L_{x'} > \min_{x \in (D(x)} + \{ c(x,\alpha) + \sum_{x' \in X} p(x'/x,\alpha) \cdot U_{x'} \} .$$

Προτείνετε πώς μπορεί να χρησιμοποιηθεί η παραπάνω ιδιότητα για να επιταχυνθεί η σύγκλιση στους αλγόριθμους των διαδοχικών προσεγγίσεων.

Ασκηση 11. Σε κάθε στάδιο μιάς εξακολουθητικής διαδικασίας, μπορούμε να πάρουμε μία από δύο διαθέσιμες αποφάσεις. Αν πάρουμε την πρώτη απόφαση μπορούμε να κερδίσουμε μία μονάδα με πιθανότητα p_1, δύο με πιθανότητα p_2, ενώ με πιθανότητα p_3 δεν κερδίζουμε τίποτα και όλη η διαδικασία σταματά. Η δεύτερη απόφαση έχει αντίστοιχο σύνολο πιθανοτήτων p_1', p_2', p_3'.

Θέλουμε να βρούμε την ακολουθία αποφάσεων που μεγιστοποιεί την πιθανότητα να κερδίσουμε n μονάδες πρίν σταματήσει η διαδικασία. Να μοντελοποιηθεί το πρόβλημα χρησιμοποιώντας τη μέθοδο του Δυναμικού Προγραμματισμού.
(Υπόδειξη. Θεωρείστε το τρέχον κέρδος ίσο με 0 και το τερματικό κέρδος ίσο με 1 ή 0, ανάλογα με το αν έχουν συγκεντρωθεί ή οχι n μονάδες στο τέλος του παιχνιδιού).

Ασκηση 12. *Ενας άνθρωπος περιμένει να εξυπηρετηθεί σε μια γραμμή αναμονής, ενώ μπροστά του βρίσκονται* N *άτομα. Αν περιμένει μέχρι να εξυπηρετηθεί θά έχει κέρδος* r*. Ομως έχει κόστος* c *για κάθε χρονική μονάδα που περιμένει. Η πιθανότητα να τελειώσει ένα άτομο απ' αυτά που βρίσκονται μπροστά του κάθε χρονική μονάδα είναι ίση με* p*. Το πρόβλημα είναι να βρεί την πολιτική αναμονής που μεγιστοποιεί το αναμενομένο κέρδος του.*
Μοντελοποιήστε τοπρόβλημα με μεθόδους του Δυναμικού Προγραμματισμού και προσπαθείστε να λύσετε την εξίσωση βελτιστοποίησης.

Ασκηση 13. *Ενας υπάλληλος πηγαίνει με το αυτοκίνητο του στη δουλειά και θέλει να βρεί μια βέλτιστη πολιτική όσον αφορά το παρκάρισμα. Μπορεί να βάλει το αυτοκίνητο σε γκαράζ, απόφαση που έχει κόστος* B*, είτε μπορεί να προσπαθήσει να παρκάρει στο δρόμο. Σ' αυτή την περίπτωση όμως, όσο πιο μακριά από το χώρο της δουλειάς του αφήσει το αυτοκίνητο , τόσο περισσότερο περπάτημα θα πρέπει να κάνει. Εστω οτι στο δρόμο υπάρχουν* N *τετράγωνα πριν και* N *μετά το κτίριο που δουλεύει, στα οποία μπορεί να βρεί θεση να παρκάρει, και τα οποία συμβολίζουμε με* $-N, -N+1, \dots, -1, 0, 1, \dots, N-1, N$*. Το τετράγωνο με αριθμό* 0 *είναι μπροστά στο κτίριο, κι εκεί απαγορεύεται η στάθμευση. Το κόστος που έχει αν αφήσει το αυτοκίνητο στο τετράγωνο με αριθμό* n*, εκτιμά ότι είναι ίσο με* $|n|$*. Εστω* p_n*,* $n = -N, \dots, -1, 1, \dots, N$ *η πιθανότητα να βρεί κενή θέση στο* n *τετράγωνο.*

α) *Βρείτε ένα μοντέλο Δυναμικού Προγραμματισμού που δίνει τη βέλτιστη πολιτική ως πρός το αναμένομενο κόστος. Υποθέστε οτι κάθε φορά ο υπάλληλος βλέπει μόνο ένα τετράγωνο .*
β) Γράψτε το αντίστοιχο πρωτεύον και το δυικό γραμμικό πρόγραμμα αν $N = 3$*.*

(Υπόδειξη . Η Κατάσταση του συστήματος μπορεί να είναι ο αριθμός του τετραγώνου και το αν είναι ή οχι κατειλημμένο).
Ασκηση 14. *Ας περιγράψουμε ένα τυχερό παιγνίδι στο οποίο θα αναφερθούμε συχνά και από τις επόμενες ασκήσεις (15 έως 19) .*
Ενας παίκτης παίρνει πληροφορίες για τα αποτελέσματα μιάς σειράς ανεξάρτητων αγώνων μπάσκετ. Το πρόβλημα είναι ότι τις πληροφορίες τις παίρνει μέσα από ένα τηλεπικοινωνιακό κανάλι που εχει θόρυβο . Ετσι για καθε αγώνα το κανάλι έχει πιθανότητα ίση με p *να μεταδώσει το σωστό αποτέλεσμα και* $q = 1 - p$ *το αντίθετο. Αφού πάρει μιά πληροφορία, ο παίκτης έχει τη δυνατότητα να στοιχηματίσει υπέρ της μιάς ή της άλλης ομάδας όσο ποσό θέλει από όσο έχει στη διάθεση του εκείνη τη στιγμή. Υποθέτοντας ότι ο παίκτης ξεκινάει με αρχικό ποσό* x*, δείξτε χρησιμοποιώντας επιχειρήματα Δυναμικού Προγραμματισμού ότι για να μεγιστοποιήσει το αναμενόμενο κεφάλαιο του στο τέλος* N *σταδίων του παιγνιδιού, πρέπει να στοιχηματίζει ολόκληρο το κεφάλαιο του σε κάθε στάδιο, αν* $p > 1/2$ *και καθόλου αν* $p < 1/2$*.*

Ασκηση 15. *Στο παιγνίδι της άσκησης* 14*, υποθέτουμε οτι ο παίκτης θέλει να μεγιστοποιήσει την αναμενόμενη τιμή του λογαρίθμου του κεφαλαίου του μετά από* N *στάδια. Αν χρησιμοποιεί την ίδια πολιτική σε κάθε στάδιο, βρείτε το λόγο τού ποσού που στοιχηματίζει προς το συνολικό κεφάλαιο.*

Ασκηση 16. *Ξανά στην άσκηση* 14*. Ας υποθέσουμε οτι ο παίκτης θέλει να μεγιστοποιήσει την αναμενόμενη τιμή του κεφαλαίου του μετά από* N *στάδια παιχνιδιού.*
Εστω οτι η $f(x)$ *δηλώνει την αναμενόμενη τιμή που πετυχαίνει ο παίκτης όταν χρησιμοποιεί τη βέλτιστη πολιτική, ξεκινά με αρχικό ποσό* x*, και παίζει* N *φορές. Δείξτε ότι*

$$f_{N+1}(x) = \max_{0 \le y \le x} \{p \cdot f_N(x+y) + q \cdot f_N(x-y)\}, \quad \mathrm{N} = 1,2,\dots$$

$$\mathrm{f}_1(x) = \max_{0 \le y \le x} \{p \cdot \log(x+y) + q \cdot \log(x-y)\}.$$

Ασκηση 17 . *Στην άσκηση* 16 *δείξτε επαγωγικά οτι*

$$f_n(x) \;=\; logx + N\,k\,,$$

όπου

$$k = \max_{0 \le r \le 1} \{\, p\,log(1+r) \;+\; q\,log(1-r)\}\,,$$

και επομένως υπάρχει ενας αριθμός r_0 , *έτσι ώστε η βέλτιστη πολιτική σε κάθε στάδιο ορίζεται από τη σχέση* $y = r_0 x$.

Ασκηση 18 . *Θεωρείστε την περίπτωση, όπου οι πιθανότητες σωστής μετάδοσης του καναλιού εξαρτώνται απο το στάδιο του παιγνιδιού. Βρείτε τις αντίστοιχες εξισώσεις βελτιστοποίησης και προσδιορίστε τη δομή της βέλτιστης πολιτικής.*

Ασκηση 19 . *Αυξάνοντας κατά ένα ακόμη βήμα την πολυπλοκότητα του μοντέλου, θεωρούμε οτι τα διαδοχικά σήματα του καναλιού δεν είναι ανεξάρτητα. Ετσι η πθανότητα σωστής μετάδοσης στο στάδιο* k *εξαρτάται απο τη μετάδοση στο στάδιο* $k-1$, *με τρόπο που θα ορισθεί παρακάτω. Για* $x > 0$ *ορίζουμε τις ποσότητες*

$f(x)$: *αναμενόμενη τιμή του τελικού κεφαλαίου πού θα αποκτηθεί στα υπόλοιπα* k *στάδια από τα* N *της όλης διαδικασίας, αν αρχίσουμε με κεφάλαιο* x , *την πληροφορία οτι η μετάδοση του* $k-1$ *σταδίου έγινε σωστά, και χρησιμοποιώντας βέλτιστη πολιτική.*

$g_k(x)$: *η αντίστοιχη με την* $f_k(x)$ *τιμή όταν στο* $k-1$ *στάδιο η μετάδοση ήταν λαθεμένη.*

Τότε δείξτε οτι η εξίσωση βελτιστοποίησης χωρίζεται στις παρακάτω δύο εξισώσεις:

$$f_k(x) \;=\; \max_{0 \le y \le x} \{p_{N-k+1} f_{k-1}(x+y) \;+\; (1-p_{N-k+1})\, g_k(x-y)\}\,,$$

$$g_k(x) \;=\; \max_{0 \le y \le x} \{\, r_{N-k+1} f_{k-1}(x+y) + (1 - r_{N-k+1}\,)\, g_{k-1}(x-y)\,\}\,,$$

όπου

p_k : *η πιθανότητα σωστής μετάδοσης του k–οστού σήματος, δεδομένου ότι το* $k-1$ *μεταδόθηκε σωστά.*

r_k : *η πιθανότητα σωστής μετάδοσης του k–οστού σήματος, δεδομένου ότι το* $k-1$ *μεταδόθηκε λάθος.*

Δείξτε ότι $f_k(x) = logx + \alpha_k$, $g_k(x) = logx + b_k$. *Προσδιορίστε τα* α_k , b_k , *καθώς επίσης και τη δομή της βέλτιστης πολιτικής.*

ΠΑΡΑΡΤΗΜΑ Α

ΕΦΑΡΜΟΓΕΣ ΚΑΙ ΠΡΟΓΡΑΜΜΑΤΑ Η.Υ.

1η ΑΣΚΗΣΗ *Θεωρούμε το πρόβλημα ελαχιστοποίησης του κόστους παραγωγής και αποθεμάτων, με τις εξής παραλλαγές .*
Ο ορίζοντας προγραμματισμού είναι άπειρος, η ζήτηση γενικά πιθανοθεωρητική, και μας ενδιαφέρει να βρούμε την πολιτική που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικ`ο κόστος. Δεχόμαστε ότι η ζήτηση δεν είναι σταθερή σε κάθε χρονική περίοδο, αλλά εμφανίζει περιοδικότητα. Αυτό σημαίνει ότι υπάρχει ένας αριθμός περιόδων N, τέτοιος ώστε η κατανομή της τυχαίας μεταβλητής που αντιστοιχεί στη ζήτηση να επαναλαμβάνεται περιοδικά με περίοδο N. Το φαινόμενο αυτό παρατηρείται για παράδειγμα σε φαρμακεία, όπου η ζήτηση φαρμάκων για γρίππη παρουσιάζει διακυμάνσεις κατά τη διάρκεια του χρόνου, αλλά μένει περίπου σταθερή από χρόνο σε χρόνο.
Σε ότι αφορά την κατανομή της ζήτησης στη διάρκεια μιας χρονικής περιόδου, υποθέτουμε ότι είναι διακριτή, και από περίοδο σε περίοδο μεταβάλλονται οι δυνατές τιμές και οι αντίστοιχες πιθανότητες.
Οι συναρτήσεις κόστους παραγωγής και διατήρησης αποθεμάτων είναι, αντίστοιχα, οι εξής:

$$c_t(\alpha) = 2\alpha - (1+\alpha)^{-2} + 1, \quad \text{και} \quad h_t(x) = h \cdot x\,,$$

όπου τα α, x συμβολίζουν ποσότητες παραγωγής και αποθεμάτων αντίστοιχα.
Δεν είναι απαραίτητο να ικανοποιείται η ζήτηση οπωσδήποτε, αλλά όταν παρουσιαστεί έλλειψη s μονάδων προιόντος σε κάποια χρονική περίοδο, τότε λόγω διαφυγόντων κερδών, απώλειας φήμης της εταιρείας κλπ, προκύπτει επιπλέον κόστος έλλειψης $L(x) = 5s^2$.
Με βάση τις παραπάνω παραδοχές ζητείται να βρεθε`ι με τη βοήθεια Η.Υ. η βέλτιστη λύση ως προς το αναμενόμενο συνολικό εκπτωτικό κόστος, για διάφορες περιπτώσεις κατανομής της ζήτησης και διάφορες τιμές του μοναδιαίου κόστους αποθήκευσης και του εκπτωτικού παράγοντα.
Υποδείξεις *1. Στο διάνυσμα κατάστασης περιλαμβάνεται η περίοδος μέσα στο διάστημα ανακύκλωσης, και το διαθέσιμο απόθεμα στην αρχή της περιόδου.*
2. Υποθέτουμε οτι αν σε κάποια περίοδο η ζήτηση ξεπεράσει το απόθεμα, η έλλειψη μένει και ικανοποιείται όταν γίνει νέα παραγωγή. Αυτόσημαίνει ότι επιτρέπουμε στο απόθεμα να παίρνει και αρνητικές τιμές.
3. Για λόγους προγραμματισμού είναι απαραίτητο να ορίσουμε με`γιστο και ελάχιστο επιτρεπόμενο απόθεμα, ώστε οι χώροι καταστάσεων και αποφάσεων να είναι πεπερασμένοι.

2η ΑΣΚΗΣΗ *Ο προιστάμενος του τμήματος προσωπικού της Ο.Α. αντιμετωπίζει το εξής πρόβλημα. Θέλει να αναπτύξει μια πολιτική πρόσληψης αεροσυνοδών για τα επόμενα χρόνια. Οι απαιτήσεις της εταιρείας σε ώρες πτήσης αεροσυνοδών κάθε μήνα ε`ιναι ανεξάρτητες διακριτές τυχαίες μεταβλητές με διαφορετικές γενικά κατανομές. Οι απαιτήσεις αυτές όμως υποθέτουμε ότι επαναλαμβάνονται περιοδικά ως προς την κατανομή από χρόνο σε χρόνο, όπως και στο προηγούμενο παράδειγμα.*

Η εκπαίδευση για κάθε καινούργια αεροσυνοδό διαρκεί ένα μήνα και απαιτεί 100 ώρες επίβλεψης απο μια εκπαιδευμένη αεροσυνοδό. Ετσι η εταιρεία χάνει 100 ώρες αεροσυνοδού για την εκπαίδευση κάθε καινούργιας αεροσυνοδού.

Κάθε αεροσυνοδός μπορεί να δουλέψει το πολύ 150 ώρες το μήνα, και στο τέλος κάθε μήνα ένα ποσοστό των αεροσυνοδών φεύγουν απο την εταιρεία για προσωπικούς λόγους. Το ποσοστό αυτό είναι επίσης τυχαία μεταβλητή, της οποίας όμως η κατανομή είναι η ίδια για όλες τις χρονικές περιόδους (μήνες).

Η εταιρεία ακολουθεί την πολιτική να μην απολύει αεροσυνοδούς, έτσι όταν ο μέγιστος αριθμός ωρών εργασίας που είναι διαθέσιμες ξεπερνά τις απαιτήσεις πτήσης και εκπαίδευσης του μήνα, οι υπάρχουσες αεροσυνοδοί απλώς εργάζονται λιγότερες ώρες.

Το μηνιαίο κόστος για την εταιρεία είναι r *για κάθε αεροσυνοδό και* c *για κάθε υπό εκπαίδευση αεροσυνοδό.*

Οταν παρουσιαστεί ζήτηση σε ώρες πτήσης και επίβλεψης μεγαλύτερη από τις διαθέσιμες ώρες εργασίας, τότε οι επιπλέον ώρες καλύπτονται με υπερωριακή απασχόληση που στοιχίζει στην εταιρεία ένα συγκεκριμένο ποσό s *για κάθε επιπλέον ώρα.*

Ζητείται να βρεθεί η πολιτική προσλήψεων που ελαχιστοποιεί το αναμενόμενο συνολικό εκπτωτικό κόστος της εταιρείας, για μεγάλο (προσεγγιστικά άπειρο) χρονικό διάστημα, και για διάφορες κατανομές ζήτησης και διάφορες τιμές της περιόδου ανακύκλωσης και των r,c,s .

Ζητείται επίσης να βρεθεί η πολιτική προσλήψεων στην περίπτωση που η εκπαίδευση της αεροσυνοδού διαρκεί δύο μήνες.

Υποδείξεις *1. Στο διάνυσμα κατάστασης περιλαμβάνεται η περίοδος μέσα στο διάστημα ανακύκλωσης (δηλαδή ο μήνας) και ο αριθμός αεροσυνοδών στην αρχή της περιόδου.*
Στην περίπτωση που η εκπαίδευση διαρκεί δύο μήνες, στην κατάσταση πρέπει να περιληφθεί και ο αριθμός των αεροσυνοδών που έχουν συμπληρώσει ένα μήνα εκπαίδευσης.

2. Για το μέγεθος του χώρου καταστάσεων μπορούμε να κάνουμε την εξής παρατήρηση. Ο αριθμός των αεροσυνοδών σε κάθε περίοδο δεν είναι απαραίτητο να ξεπερνάει ένα άνω φράγμα, που εξαρτάται από το μέγιστο δυνατό αριθμό των ωρών πτήσης που μπορε`ι να προκύψουν κάποια περίοδο, και που πρέπει κάθε φορά να υπολογίζεται από τα δεδομένα του προβλήματος.

3η ΑΣΚΗΣΗ *Θεωρούμε το παρακάτω μοντέλο αντικατάστασης μηχανημάτων. Εχουμε* n *μηχανήματα συνδεμένα σε σειρά. Ο χρόνος λειτουργίας του μηχανήματος* i *είναι τυχαία μεταβλητή που ακολουθεί την εκθετική κατανομή με παράμετρο* μ_i , $i = 1,\ldots,n$. *Για την επισκευή των χαλασμένων μηχανημάτων υπάρχει ένας επισκευαστής, ο οποίος μπορεί να εργάζεται το πολύ σε ένα μηχάνημα κάθε χρονική στιγμή. Οταν το μηχάνημα* i *χαλάσει, ο χρόνος επιδιόρθωσης, που υπολογίζεται από τη στιγμή που θα αρχίσει ο επισκευαστής να εργάζεται πάνω σαυτό μέχρι να τελειώσει η επισκευή, είναι τυχαία μεταβλητή, που ακολουθεί την εκθετική κατανομή με παράμετρο* λ_i . *Υποθέτουμε ότι κάθε μηχάνημα που επισκευάζεται συμπεριφέρεται σα να ήταν καινούργιο, όπως επίσης και ότι όταν ένα ή περισσότερα μηχανήματα είναι χαλασμένα, με αποτέλεσμα το εν σειρά σύστημα να μη λειτουργεί, τα υπόλοιπα μηχανήματα συνεχίζουν να χαλάνε με τους ίδιους ρυθμούς, όπως και όταν λειτουργούσε το σύστημα.*

Το πρόβλημα απόφασης προκύπτει, όπως είναι φανερό, όταν περισσότερα από ένα μηχανήματα είναι χαλασμένα, οπότε θα πρέπει να αποφασίσουμε ποιο μηχάνημα θα επισκευάσει πρώτο ο επισκευαστής.

Το κριτήριο βελτιστοποίησης που μας ενδιαφέρει είναι η διαθεσιμότητα (availability) του συστήματος που ορίζεται σαν το αναμενόμενο ποσοστό του χρόνου που το σύστημα βρίσκεται σε λειτουργία.

Ζητείται να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού, και να βρεθεί προγραμματιστικά η βέλτιστη λύση.

ΠΑΡΑΡΤΗΜΑ Β
ΥΠΟΔΕΙΓΜΑΤΙΚΕΣ ΑΣΚΗΣΕΙΣ ΚΑΙ ΛΥΣΕΙΣ

1. ΠΡΟΟΔΟΣ ΑΝΟΙΞΗ '87

ΠΡΟΒΛΗΜΑ 1: ***1)*** *Να λυθεί με τις μεθόδους διαδοχικών προσεγγίσεων,*

α) στο χώρο τιμών

β) στο χώρο πολιτικών

το πρόβλημα διαδρομής ελάχιστου μήκους του σχήματος 1.

2) *Να δοθεί η διατύπωση του προβλήματος μέσω γραμμικού προγραμματισμού.*

Σχ`ημα 1

ΠΡΟΒΛΗΜΑ 2: *Εστω ένα δίκτυο (σχήμα 2) και έστω* $\delta_1 = (x_o,\dots,y,\dots z)$ *μια βέλτιστη διαδρομή απο το* x_o *στο* z *(τερματικό).*

Εστω $\delta_2 = (x_o,\dots,y)$ *και* $\delta_3 = (y,\dots,z)$ *δυο κομμάτια της διαδρομής.*

1) *Είναι η* δ_3 *βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το* y *και τελικότο* z *; δώστε απόδειξη).*

2) *Είναι η* δ_2 *βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το* x_0 *και τελικότο* y *; δώστε απόδειξη).*

3) *Αν η απάντηση στο (2) είναι καταφατική, υπάρχει εξίσωση βελτιστοποίησης που βασίζεται στην ιδιότητα (2) ; (αν ναι, να βρεθεί).*

Σχ`ημα 2

ΠΡΟΒΛΗΜΑ 3 *Εστω το πρόβλημα:*

$$(min) \quad w = \sum_{t=0}^{3} [2x^2(t) + \alpha^2(t)]$$

$$\begin{aligned} \upsilon.\pi. \quad & x(t+1) = x(t) - \alpha(t) \\ & x(t) \in X = [0,1] \quad \forall\, t, \\ & \alpha(t) \in A(x(t)) = [0,1], \\ & x(0) = 1. \end{aligned}$$

Να βρεθεί η βέλτιστη πολιτική $\pi = (\alpha_o,\alpha_1,\alpha_2)$.

ΠΡΟΒΛΗΜΑ 4: *Εστω το πρόβλημα:*

$$(max) \quad \sum_{j=1}^{n} c_j x_j$$

$$\begin{aligned} \upsilon.\pi. \quad & \alpha_1 \le x_1 \le b_1 \\ & \alpha_j x_j \le x_{j+1} \le b_j x_j, \qquad j = 1,2,\dots,n-1, \end{aligned}$$

όπου c_j,α_j,b_j είναι γνωστοί αριθμοί.
Να δοθεί η διατύπωση του προβλήματος μέσω δυναμικού προγραμματισμού.

ΛΥΣΕΙΣ ΠΡΟΟΔΟΥ 87

2 . ΕΞΕΤΑΣΤΙΚΗ ΠΕΡΙΟΔΟΣ ΙΟΥΝΙΟΥ 87

ΠΡΟΒΛΗΜΑ 1: Θεωρούμε τους αριθμούς $\alpha_1,\alpha_2,\ldots,\alpha_n$, όπου $\alpha_\kappa \geq 0$ $k = 1,2,\ldots,n$, και $\sum_{\kappa=1}^{n}\alpha_\kappa = A$. Να αποδείξετε με επιχειρήματα Δυναμικού Προγραμματισμού ότι:

$$[(\frac{1}{n}\sum_{\kappa=1}^{n}\alpha_\kappa\}]^n \geq \prod_{\kappa=1}^{n}\alpha_\kappa$$

ΠΡΟΒΛΗΜΑ 2: Δίνεται το δίκτυο του διπλανού σχήματος, όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι:

$$\sum_{t=0}\beta^t c(x_t,\alpha_t), \text{ με } \beta = 0.9 \text{ , όταν } x_o = 1.$$

ΠΡΟΒΛΗΜΑ 3: Εστω το πρόβλημα:

$$\begin{array}{ll} \max & 3x + 5y \\ \upsilon.\pi. & 2x + y \leq 10 \\ & x + 2y \leq 6 \\ & x \in \{0,1,4,6\} \\ & y \in [0,\infty] \ . \end{array}$$

Να δοθε`ι η διατ`υπωση του προβλ`ηματος με τη μ`εθοδο του Δυναμικο`υ Προγραμματισμο`υ.

ΠΡΟΒΛΗΜΑ 4: Να διατυπώσετε το γενικό πρόβλημα υπολογισμού της διαδρομής ελάχιστου μήκους σε ένα δίκτυο.

1) Να δώσετε και να αποδείξετε την εξίσωση βελτιστοποίησης του Δυναμικού Προγραμματισμού για το πρόβλημα.

2) Να περιγράψετε τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο τιμών.

3) Να περιγράψετε τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών.

ΛΥΣΕΙΣ

ΠΡΟΒΛΗΜΑ 1: *Θεωρούμε το παρακάτω πρόβλημα δυναμικού προγραμματισμού: Ζητάμε να μεγιστοποιήσουμε το γινόμενο:*

$$\prod_{\kappa=1}^{n} \alpha_\kappa \quad \text{με περιορισμούς} \sum_{\kappa=1}^{n} \alpha_\kappa = A\,,\ \ \alpha_\kappa \geq 0\ \kappa = 1,\ldots,n.$$

Αυτό είναι πρόβλημα κατανομής πόρων με μόνη διαφορά ότι το κριτήριο βελτιστοποίησης είναι πολλαπλασιαστικό και όχι αθροιστικό. Παρατηρώντας ότι $\prod_{\kappa=1}^{n} \alpha_\kappa = \prod_{\kappa=1}^{n-1} \alpha_\kappa$, *μπορούμε να γράψουμε την εξίσωση βελτιστοποίησης με τρόπο ανάλογο των σημειώσεων:*

$$v(\kappa,x) = \max_{0 \leq \alpha_\kappa \leq x} \{\alpha_\kappa v(\kappa-1, x-\alpha_\kappa)\}, \quad \kappa = 1,\ldots n$$

με $v(0,x) = 1$, *όπου ζητάμε το* $v(n,A)$ *και τα* $\alpha_i^*,\ldots,\alpha_n^*$. *Το σύμβολο κ συμβολίζει εδώ τον αριθμό των κατανομών που επομένουν ακόμη.*

Θα αποδείξουμε επαγωγικά ότι:

$$v(\kappa,x) = (\tfrac{x}{\kappa})^\kappa \ \ \text{και}\ \ \alpha_\kappa^*(x) = \tfrac{x}{\kappa}$$

$\underline{\kappa = 1}$:

$$v(1,x) = \max_{0 \leq \alpha_{\kappa+1} \leq x} \{\alpha_1 v(0, x-\alpha_1)\} = \max_{0 \leq \alpha_\kappa \leq x} \{\alpha_1\} = x = (\tfrac{x}{1})^1, \quad \alpha_1^*(x) = \tfrac{x}{1}$$

Εστω ότι για κάποιο $\kappa > 1$ *έχουμε*

$$v(\kappa,x) = (\tfrac{x}{\kappa})^\kappa, \quad \alpha_\kappa^*(x) = \tfrac{x}{\kappa}. \quad \text{Τότε}$$

$$v(\kappa+1,x) = \max_{0 \leq \alpha_{\kappa+1} \leq x} \{\alpha_{\kappa+1} v(\kappa, x-\alpha_{\kappa+1})\} = \max_{0 \leq \alpha_{\kappa+1} \leq x} \{\alpha_{\kappa+1} \cdot (\tfrac{x-\alpha_{\kappa+1}}{\kappa})^\kappa\}\,.$$

Θέτουμε $f(\alpha) = \alpha(\frac{x-\alpha}{\kappa})^\kappa,\ \alpha \in [0,x]$.

Τότε παίρνοντας την πρώτη παράγωγο της f έχουμε:

$$f'(\alpha) = (\tfrac{x-\alpha}{\kappa})^\kappa - \alpha(\tfrac{x-\alpha}{\kappa})^{\kappa-1} = (\tfrac{x-\alpha}{\kappa})^{\kappa-1} \cdot (\tfrac{x-\alpha}{\kappa} - \alpha)\,.$$

Αρα $f'(\alpha) = 0$ *όταν* $x = \alpha$, *ή* $x - \alpha = \kappa\alpha \Rightarrow \alpha = \frac{x}{\kappa+1}$, *και επομένως* $\alpha_{\kappa+1}^* = \frac{x}{\kappa+1}$, *ενώ*

$$v(\kappa+1,x) = \tfrac{x}{\kappa+1} \cdot (\tfrac{\kappa x}{\kappa(\kappa+1)})^\kappa = (\tfrac{x}{\kappa+1})^{\kappa+1}\,.$$

Ετσι συμπληρώθηκε η επαγωγική απόδειξη και επομένως έχουμε:

$$\max \prod_{\kappa=1}^{n} \alpha_\kappa = \left(\frac{A}{n}\right)^n = [(\tfrac{1}{n})\sum_{\kappa=1}^{n} \alpha_\kappa]^n \Rightarrow \prod_{\kappa=1}^{n} \alpha_\kappa \leq [(\tfrac{1}{n})\sum_{\kappa=1}^{n} \alpha_\kappa]^n\,.$$

ΠΡΟΒΛΗΜΑ 2 : *Επειδή πρόκειται για πρόβλημα άπειρου ορίζοντα, έχουμε δει ότι η βέλτιστη πολιτική δεν θα εξαρτάται απο το χρόνο* t . *Ζητάμε δηλαδή βέλτιστη πολιτική της μορφής* $\alpha^* = \alpha^*(x)$. *Παρατηρούμε ακόμη ότι :*

$$\sum_{t=0}^{\infty} \beta^t c(x_t,\alpha_t) = c(x_o,\alpha_o) + \beta \cdot \{\, c(x_1,\alpha_1) + \beta \cdot [c(x_2,\alpha_2) + \cdots]\,\}\,.$$

Σύμφωνα με τα παραπάνω, αν με $\alpha = \alpha(x)$ *συμβολίσουμε τον επόμενο κόμβο που θα επισκεφθούμε μετά τον* x *σύμφωνα με την πολιτική* α *, η εξίσωση βελτιστοποίησης έχει την παρακάτω μορφή:*

$$v(x) = \max_{x \in A(x)} \{c(x,\alpha) + \beta v(\alpha)\} \;, \quad x = 1,2,3,4 \;,$$

όπου $A(1) = \{2,3\}$, $A(2) = \{1,4\}$, $A(3) = \{1\}$, $A(4) = \{2,4\}$.

Αναλυτικά οι προηγούμενες εξισώσεις γράφονται :

$$v(1) = \max\{4 + \beta v(2),\ 3 + \beta v(3)\}$$
$$v(2) = \max\{2 + \beta v(1),\ 10 + \beta v(4)\}$$
$$v(3) = 2 + \beta v(1)$$
$$v(4) = \max\{4 + \beta v(2),\ 1 + \beta v(4)\} \;.$$

Επειδή έχουμε πρόβλημα άπειρου ορίζοντα, η μέθοδος δ.π. στο χώρο τιμών μπορεί να μας δώσει μόνο προσεγγιστικά τις τιμές των $v(x)$ *, με τόσο καλύτερη προσέγγιση ότο μεγαλύτερο είναι το* n *στη σχέση :*

$$v^{(n)}(x) = \max_{x \in A(x)} \{c(x,\alpha) + \beta v^{(n-1)}(\alpha)\} \;.$$

Αντίθετα με τη μέθοδο δ.π. στο χώρο των πολιτικών, μπορούμε να βρούμε την ακριβή λύση του προβλήματος σε πεπερασμένο αριθμό βημάτων.
Ετσι ξεκινάμε με μια αυθαίρετη πολιτική $\alpha^{(o)}(1) = 2$, $\alpha^{(o)}(2) = 1$, $\alpha^{(o)}(3) = 1$, $\alpha^{(o)}(4) = 4$. *Γι' αυτήν την πολιτική υπολογίζουμε τα κόστη :*

$$w^{(o)}(x) = c(x,\alpha^{(o)}(x)) \;+\; \beta w^{(o)}(\alpha) \;.$$

$$w^{(o)}(1) \;=\; 4 + \;0.9\, w^{(o)}(2)$$
$$w^{(o)}(2) \;=\; 2 + \;0.9\, w^{(o)}(1)$$
$$w^{(o)}(3) \;=\; 2 + \;0.9\, w^{(o)}(1)$$
$$w^{(o)}(4) \;=\; 1 + \;0.9\, w^{(o)}(4) \quad .$$

Απ`ο την τελευταία σχέση παίρνουμε $w^{(o)}(4) = 10$. *Επίσης οι δυο πρώτες σχέσεις αποτελούν σύστημα δυο εξισώσεων με δυο αγνώστους, που απο τη λύση του παίρνουμε* $w^{(o)}(1) = 30.526$ *και* $w^{(o)}(2) = 29.474$. *Τέλος αντικαθιστώντας στην τρίτη σχέση βρίσκουμε* $w^{(o)}(3) = 29.474$.

Υπολογίζουμε τώρα τις διαφορές $\Delta w^{(o)}(x,\alpha) = c(x,\alpha) + \beta w^{(o)}(\alpha) - w^{(o)}(x)$ *για κε`ινες τις αποφάσεις σε κάθε κόμβο που δεν αντιστοιχούν στη βέλτιστη πολιτική.*

$$\Delta w^{(o)}(1,3) = c(1,3) + 0.9 w^{(o)}(3) - w^{(o)}(1) \;=\; 3 + 0.9 \cdot 29.474 - 30.526 = \;-\,0.999 \;<\; 0$$
$$\Delta w^{(o)}(2,4) = c(2,4) + 0.9 w^{(o)}(4) - w^{(o)}(2) \;=\; 10 + 0.9 \cdot 10 - 29.474 = \;-\,10.474 \;<\; 0$$
$$\Delta w^{(o)}(4,2) = c(4,2) + 0.9 w^{(o)}(2) - w^{(o)}(4) \;=\; 4 + 0.9 \cdot 29.474 - 10 = 20.527 \;>\; 0 \;\;.$$

Επειδή έχουμε δ`υο διαφορές αρνητικές η πολιτική $\alpha^{(o)}$ *δεν είναι βέλτιστη, και γι' αυτό προχωρούμε στη βελτίωσή της με τον τρόπο που ξέρουμε :*

$$\alpha^{(1)}(1) = 3, \quad \alpha^{(1)}(2) = 4, \;\; \alpha^{(1)}(3) = 1, \;\; \alpha^{((1)}(4) = 4.$$

Ακόμη υπολογίζουμε τα νέα κόστη :

$$w^{(1)}(1) = 3 + 0.9 w^{(1)}(3)$$
$$w^{(1)}(2) = 10 + 0.9 w^{(1)}(4)$$
$$w^{(1)}(3) = 2 + 0.9 w^{(1)}(1)$$

$$w^{(1)}(4) = 1 + 0.9w^{(1)}(4) \ .$$

Τώρα βρίσκουμε πάλι απ`ο την τέταρτη σχέση* $w^{(1)}(4) = 10$, *και απ`ο αντικατάσταση στη δεύτερη $w^{(1)}(2) = 19$, *ενώ απ`ο το σύστημα της πρώτης και τρίτης* $w^{(1)}(1) = 25.263$, $w^{(1)}(3) = 24.736$. *Οι νέες διαφορές είναι :*

$$\Delta w^{(1)}(1,2) = c(1,2) + 0.9w^{(1)}(2) - w^{(1)}(1) \ = \ 4 + 0.9 \cdot 19 - 25.263 = \ -4.163 \ < \ 0$$
$$\Delta w^{(1)}(2,1) = c(2,1) + 0.9w^{(1)}(1) - w^{(1)}(2) \ = \ 2 + 0.9 \cdot 25.273 - 19 = 5.737 \ > \ 0$$
$$\Delta w^{(1)}(4,2) = c(4,2) + 0.9w^{(1)}(2) - w^{(1)}(4) \ = \ 4 + 0.9 \cdot 19 - 10 = 11.1 > 0 \ .$$

Επειδή κι εδώ $\Delta w^{(1)}(1,2) < 0$, *συνεχίζουμε τη μέθοδο παίρνοντας τη νέα πολιτική :*

$$\alpha^{(2)}(1) = 2. \ \ \alpha^{(2)} = 4, \ \ \alpha^{(2)}(3) = 1, \ \ \alpha^{(2)}(4) = 4 \ \ .$$

$$w^{(2)}(1) = 4 + 0.9w^{(2)}(2)$$
$$w^{(2)}(2) = 10 + 0.9w^{(2)}(4)$$
$$w^{(2)}(3) = 2 + 0.9w^{(2)}(1)$$
$$w^{(2)}(4) = 1 + 0.9w^{(2)}(4) \ \ .$$

Τώρα με απλές αντικαταστάσεις παίρνουμε :

$$w^{(2)}(4) = 10 \ , \ \ w^{(2)}(2) = 19 \ , \ \ w^{(2)}(1) = 21.1 \ , \ \ w^{(2)}(3) = 20.99 \ .$$

$$\Delta w^{(2)}(1,3) = c(1,3) + 0.9w^{(2)}(3) - w^{(2)}(1) \ = \ 3 + 0.9 \cdot 20.99 - 21.1 = 0.791 \ > \ 0$$
$$\Delta w^{(2)}(2,1) = c(2,1) + 0.9w^{(2)}(1) - w^{(2)}(2) \ = \ 2 + 0.9 \cdot 21.1 - 19 = 1.99 \ > \ 0$$
$$\Delta w^{(2)}(4,2) = c(4,2) + 0.9w^{(2)}(2) - w^{(2)}(4) \ = \ 4 + 0.9 \cdot 19 - 10 = 11.1 \ > \ 0 \ \ .$$

Επειδή όλες οι διαφορές προκύπτουν θετικές, ο αλγόριθμος σταματάει και η βέλτιστη πολιτική είναι η $\alpha^{(2)}$.

ΠΡΟΒΛΗΜΑ 3: *Ζητάμε να μεγιστοποιήσουμε την παράσταση* $z = 3x + 4y$ *με περιορισμούς :*

$$2x + y \leq 10$$
$$x + 2y \leq 6$$
$$x \ \in \ \{0,1,4,6\}$$
$$y \ \in \ (0,\infty) \ .$$

Είναι φανερό ότι πρόκειται για πρόβλημα κατανομής πόρων με δ`υο μεταβλητές απόφασης και δ`υο περιορισμούς. Επειδή έχουμε δ`υο μεταβλητές απόφασης, ο ορίζοντας προγραμματισμού θα είναι* $N = 2$, *ενώ επειδή έχουμε δ`υο περιορισμούς (πόρων) η κατάσταση σε κάθε βήμα θα περιγράφεται απο 2 μεταβλητές B_1 *και* B_2, *που εκφράζουν τις ποσότητες πόρων που απομένουν για τον υπόλοιπο ορίζοντα.*
Υποθέτονας ότι στο πρώτο βήμα θα βρούμε την τιμή του x *και στο δεύτερο του* y, *θέτουμε* $\alpha^{(1)} = x$, $\alpha^{(2)} = y$.
Ο χώρος αποφάσεων για το $\alpha^{(2)}$ *προκύπτει απο τις απαιτήσεις:*

$$\alpha^{(2)} \leq B_1^{(2)}, \ \ 2\alpha^{(2)} \leq B_2^{(2)}, \quad \alpha^{(1)} \in \ [0,\infty] \ \ \Rightarrow \ \ A(\alpha^{(2)}) \ = \ [\ 0 \ , min\{B_1^{(2)}, \ \tfrac{B_2^{(2)}}{2}\} \ ] \ ,$$

ενώ για το $\alpha^{(1)}$ *απο τις :*

$$2\alpha^{(1)} \leq B_1^{(1)}, \ \ \alpha^{(1)} \leq B_2^{(1)}, \ \ \alpha^{(1)} \in \ \{0,1,4,6\} \ \Rightarrow$$
$$\Rightarrow \ \ A(\alpha^{(1)}) \ \ = \ \{0,1,4,6\} \ \bigcap \ [0, min\{ \ \tfrac{B_1^{(1)}}{2} \ , B_2^{(1)}\}] \ .$$

Η συνάρτηση τρέχοντος κόστους για το βήμα 1 είναι $R_1(\alpha^{(1)}) = 3\alpha^{(1)}$, *ενώ η συνάρτηση τερματικού κόστους για το βήμα 2 :*

$$\hat{c}(B_1^{(2)},\ B_2^{(2)}) = \max_{\alpha^{(2)} \in A(\alpha^{(2)})} \{\alpha^{(2)}\} \ = \ 5 \min\{B_1^{(2)},\ \tfrac{B_2^{(2)}}{2}\} .$$

Η δυναμική του συστήματος είναι κατά τα γνωστά:

$$B_1^{(2)} = B_1^{(1)} - 2\alpha^{(1)} \ ,\ \ B_2^{(2)} = B_2^{(1)} - \alpha^{(1)} \ .$$

Σύμφωνα με τα παραπάνω μπορούμε τώρα να γράψουμε την εξίσωση βελτιστοποίησης του δυναμικού προγραμματισμού για το πρόβλημα .

$$v(1,B_1^{(1)},\ B_2^{(1)}) = \max_{\alpha^{(1)} \in A(\alpha^{(1)})} \{\ 3\alpha^{(1)} \ +\ v(2,\ B_1^{(1)} - 2\alpha^{(1)},\ B_2^{(1)} - \alpha^{(1)})\ \} ,$$

$$v(2,B_1^{(2)},\ B_2^{(2)}) = \ 5 \min\{\ B_1^{(2)},\ \tfrac{B_2^{(2)}}{2}\ \} ,$$

όπου ζητάμε να βρούμε το $v(1,10,6)$, *δηλ.* $B_1^{(1)} = 10$, $B_2^{(1)} = 6$.

ΠΡΟΒΛΗΜΑ 4: *Οι απαντήσεις βρίσκονται όλες μέσα στις σημειώσεις.*

3 . ΕΞΕΤΑΣΤΙΚΗ ΠΕΡΙΟΔΟΣ ΣΕΠΤΕΜΒΡΙΟΥ 87

ΠΡΟΒΛΗΜΑ 1: *Δίνεται το δίκτυο:*
α) Να βρεθεί η διαδρομή ελάχιστου μήκους απο το (1) στο (6)
i) Με τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο τιμών.
β) Να δοθεί η διατύπωση του προβλήματος σαν πρόβλημα γραμμικού προγραμματισμού.

ΠΡΟΒΛΗΜΑ 2: *Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες* $E_1, E_2, \ldots E_n$. *Ο χρόνος εκτέλεσης της* E_i *είναι γνωστή σταθερά* T_i, $i = 1,2,\ldots,n$. *Μπορούμε να εκτελούμε το πολύ μια εργασία (υπάρχει ένας μόνο σταθμός εξυπηρέτησης) και αν η εργασία* E_i *τελειώσει τη χρονική στιγμή t έχουμε ένα κέρδος* $R_i(t)$.
Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που μεγιστοποιεί το συνολικό κέρδος.
Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

ΠΡΟΒΛΗΜΑ 3: *Εστω το πρόβλημα:*

$$\max_{t=1}^{n} (x_t)^2$$

$$\max_{t=1}^{n} (x_t) = 1$$

$$x_t \geq 0$$

Να λυθεί με τη μέθοδο του Δυναμικού Προγραμματισμού.

ΠΡΟΒΛΗΜΑ 4: *Ας θεωρήσουμε το γενικό πρόβλημα της διαδρομής ελάχιστου μήκους σε ένα δίκτυο. Να βρεθεί μια ικανή συνθήκη κάτω απο την οποία η βέλτιστη πολιτική είναι η μυωπική.*

ΛΥΣΕΙΣ

4 . ΠΡΟΟΔΟΣ ΑΠΡΙΛΙΟΣ 88

ΛΥΣΕΙΣ ΔΙΑΓΩΝΙΣΜΑΤΟΣ ΔΥΝΑΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ ΤΗΣ 27/4/88

ΠΡΟΒΛΗΜΑ 1: ***α) i)*** *Σύμφωνα με τη μέθοδο δ.π.τ. θεωρούμε μια αρχική προσέγγιση* $v^{(0)}(x) = 0$ *, για τα κόστη της βέλτιστης διαδρομής, και εφαρμόζουμε διαδοχικ`α τις σχέσεις*

$$v^{(n)}(x) = \min_{\alpha \in A(x)} \{c(x,\alpha) + v^{(n-1)}(\alpha)\} ,$$

μέχρις ότου βρο`υμε $v^{(n)}(x) = v^{(n-1)}(x)\ \forall x$ *, για κάποια τιμή του n . Ολη η διαδικασία φαίνεται συνοπτικά στον παρακάτω πίνακα* ($v^{(n)}(6) = 0$) .

x	$u^{(1)}(x)$	$\alpha^{(1)}(x)$	$u^{(2)}(x)$	$\alpha^{(2)}(x)$	$u^{(3)}(x)$	$\alpha^{(3)}(x)$	$u^{(4)}(x)$	$\alpha^{(4)}(x)$	$u^{(5)}(x)$	$\alpha^{(1)}(x)$
1	2	2	5	2	8	2	13	2ή3	13	3
2	3	3	6	3	11	3	12	3ή4	12	3ή4
3	3	4	8	4	9	5	9	5	9	5
4	5	5	7	5	7	5	7	5	7	5
5	2	6	2	6	2	6	2	6	6	2

Επομένως η βέλτιστη διαδρομή είναι η 1,3,5,6 με κόστος 13 .

ii) *Παίρνουμε μια αρχική προσέγγιση για την βέλτιστη πολιτική:*

x	:	1	2	3	4	5	6
$\alpha^{o}(x)$	:	2	3	5	5	6	$-$

Γι' αυτή την πολιτική υπολογίζουμε τα κόστη $w^{(o)}(x)$ *με βάση τους τύπους :*

$w^{(0)}(x) = c(x,\alpha^{(0)}(x) + w^{(o)}(\alpha^{(o)}(x))$, *και βρίσκουμε*

$$w^{(o)}(6) = 0 , \quad w^{(o)}(5) = 2 + w^{(o)}(6) = 2 , \quad w^{(o)}(4) = 5 + w^{(o)}(5) = 7 ,$$
$$w^{(o)}(3) = 7 + w^{(o)}(5) = 9 , \quad w^{(o)}(2) + 3 + w^{(o)}(3) = 12 , \quad w^{(o)}(1) = 2 + w^{(o)}(2) = 14 .$$

Εξετάζουμε τώρα τις διαφορές :

$$\Delta w^{(o)}(x,\alpha) = w^{(o)}(x) - [c(x,\alpha) + w^{(o)}(\alpha)] , \quad \alpha \in A(x) - \{\alpha^{(o)}(x)\} .$$

$$\Delta w^{(o)}(1,3) = w^{(o)}(1) - [c(1,3) + w^{(o)}(3)] = 14 - [4 + 9] = 1 > 0$$
$$\Delta w^{(o)}(2,4) = w^{(o)}(2) - [c(2,4) + w^{(o)}(4)] = 12 - [5 + 7] = 0$$
$$\Delta w^{(o)}(3,4) = w^{(o)}(3) - [c(3,4) + w^{(o)}(4)] = 9 - [3 + 7] = -1 < 0$$
$$\Delta w^{(o)}(4,6) = w^{(o)}(4) - [c(4,6) + w^{(o)}(6)] = 7 - 8 = -1 < 0 .$$

Η διαφορά $\Delta w^{(o)}$ *είναι θετική μόνο για τον κόμβο 1 και την απόφαση 3 , έτσι αλλάζουμε την πολιτική σ' αυτόν μόνο τον κόμβο και παίρνουμε :*

$$\alpha^{(1)} = 0\,,\quad \alpha^{(1)}(2) = 3\,,\quad \alpha^{(2)}(3) = 5\,,\quad \alpha^{(4)}4 = 5\,,\quad \alpha^{(1)}(5) = 6\,.$$

Για την $\alpha^{(1)}$ *έχουμε :*

$$w^{(1)}(6) = 0\,,\ w^{(1)}(5) = 2 + 0 = 2\,,\ \ w^{(1)}(4) = 5 + 2 = 7\,,\ \ w^{(1)}(3) = 7 + 2 = 9\,,$$
$$w^{(1)}(2) = 3 + 9 = 12\,,\ w^{(1)}(1) = 4 + 9 = 13\,.$$

Παίρνουμε πάλι τις διαφορές $\Delta w^{(1)}$

$$\Delta w^{(1)}(1,2) = w^{(1)}(1) - [c(1,2) + w^{(1)}(2)] = 13 - [2 + 12] = \ -1\ <\ 0$$
$$\Delta w^{(1)}(2,4) = 12 - [5 + 7] = \ 0$$
$$\Delta w^{(1)}(3,4) = 9 - [3 + 7]\ =\ -1\ <\ 0$$
$$\Delta w^{(1)}(4,6) = 7 - [8 + 0]\ =\ -1\ <\ 0\,.$$

Επειδή όλες οι διαφορές $\Delta w^{(1)} \leq 0$, *η* $\alpha^{(1)}$ *είναι βέλτιστη.*

β) Το γραμμικό πρόγραμμα είναι:

$$(max)\quad \sum_{i=1}^{6} \lambda_i u_i$$

$$
\text{υ.π.}\quad
\begin{aligned}
& u_1 - u_2 \leq 2\\
& u_1 - u_3 \leq 4\\
& u_2 - u_3 \leq 3\\
& u_2 - u_4 \leq 5\\
& u_3 - u_4 \leq 3\\
& u_3 - u_5 \leq 7\\
& u_4 - u_5 \leq 5\\
& u_4 - u_6 \leq 8\\
& u_5 - u_6 \leq 2\\
& u_6 = 0\ \ .
\end{aligned}
$$

ΠΡΟΒΛΗΜΑ 2: *Με* α_t, *διαλέγουμε να συμβολίζουμε τον κόμβο προορισμό (και όχι το βέλος) που αντιστοιχεί στην πολιτική.*

i) Εκπτωτικό κόστος: *Αν συμβολίσουμε*

$$v(x) = \sum_{t=0}^{\infty} \beta^t c(x_t,\alpha_t)\qquad \text{όπου}\ \ x_o = x\,,$$

παίρνουμε

$$v(x_o) = c(x_o,\alpha_o) + \sum_{t=1}^{\infty} \beta^t c(x_t,\alpha_t) = c(x_o,\alpha_c) + \beta \sum_{t=0}^{\infty} \beta^t c(x_o^{'},\alpha_o^{'}) = c(x_o,\alpha_o) + \beta v(\alpha_o)\,.$$

Απ`ο αυτή την αναδρομική σχέση και σύμφωνα με όσα αναφέρονται στις σημειώσεις σχετικά με το πρόβλημα του άπειρου ορίζοντα, δημιουργούμε το παρακάτω ισοδύναμο πρόβλημα δυναμικού προγραμματισμού:

$$v(x) = \min_{\alpha \in A(x)} \{c(x,\alpha) + \beta v(\alpha)\}\,,$$

όπου $\quad A(1) = \{1,3\},\ \ A(2) = \{1\},\ \ A(3) = \{2,3\}\,.$

Για τη λύση του προβλήματος χρησιμοποιούμε τη μέθοδο δ.π.π. που, όπως ξέρουμε δίνει τη βέλτιστη πολιτική σε πεπερασμένο αριθμό προσεγγίσεων. Παίρνουμε σαν αρχική πολιτική:

$$\alpha^{(o)}(1) = 3\,, \quad \alpha^{(o)}2 = 1\,, \quad \alpha^{(o)}(3) = 3\,.$$

Τα κόστη $w^{(o)}(x)$ δίνονται απο τις παρακάτω σχέσεις:

$$\left.\begin{aligned} w^{(o)}(1) &= 10 + \beta w^{(o)}(3) = 10 + 0.5w^{(o)}(3) \\ w^{(o)}(2) &= 4 + 0.5w^{(o)}(1) \\ w^{(o)}(3) &= 1 + \beta w^{(o)}(3) = 1 + 0.5w^{(o)}(3) \end{aligned}\right\} \Rightarrow \quad \begin{aligned} w^{(o)}(3) &= \tfrac{1}{0.5} = 2 \\ w^{(o)}(1) &= 10 + 0.5.2 = 11 \\ w^{(o)}(2) &= 4 + 0.5.11 = 9{,}5 \end{aligned}$$

Εξετάζουμε τις διαφορές $\Delta w^{(o)}$:

$$\begin{aligned} \Delta w^{(o)}(1,1) &= w^{(o)}(1) - [c(1,1) + \beta w^{(o)}(1)] = 11 - [2 + 0.5 \cdot 11] \;=\; 3.5 \;>\; 0 \\ \Delta w^{(o)}(3,2) &= w^{(o)}(3) - [c(3,2) + \beta w^{(o)}(2)] = 2 - [3 + 0.5 \cdot 9.5] \;<\; 0\,. \end{aligned}$$

Σημ.: Για τον κόμβο (2) δεν εξετάζουμε διαφορές επειδή σ' αυτόν υπάρχει μια μόνο δυνατή απόφαση.
Τώρα, επειδή $\Delta w^{(o)}(1,1) < 0$, αναπροσαρμόζουμε τηνπολιτική:

$$\alpha^{(1)}(1) = 1,\;\; \alpha^{(1)}(2) = 1,\;\; \alpha^{(1)}(3) = 3$$

$$\left.\begin{aligned} w^{(1)}(1) &= 2 + 0.5w^{(1)}(1) \\ w^{(1)}(2) &= 4 + 0.5w^{(1)}(1) \\ w^{(1)}(3) &= 1 + 0.5w^{(1)}(3) \end{aligned}\right\} \Rightarrow \quad \begin{aligned} w^{(1)}(1) &= \tfrac{2}{0.5} = 4 \\ w^{(1)}(2) &= 4 + 0.5.4 = 6 \\ w^{(1)}(3) &= \tfrac{1}{0.5} = 2 \end{aligned}$$

$$\begin{aligned} \Delta w^{(1)}(1,3) &= w^{(1)}(1) - [c(1,3) + \beta w^{(1)}(3)] = 4 - (10 + 0.5 \cdot 2) \;=\; -7 \;<\; 0 \\ \Delta w^{(1)}(3,2) &= w^{(1)}(3) - [c(3,2) + \beta w^{(2)}(2)] = 2 - (3 + 0.5 \cdot 6) \;=\; -4 \;<\; 0\,. \end{aligned}$$

Επειδή $\Delta w^{(1)} < 0$, η πολιτική $\alpha^{(1)}$ είναι βέλτιστη, και επομένως με αρχική κατάσταση $x_o = 1$, έχουμε ότι για να ελαχιστοποιηθεί το εκπτωτικό κόστος πρέπει να μείνουμε μόνιμα στον κόμβο 1. Σχηματικά, οι πολιτικές $\alpha^{(o)}$, $\alpha^{(1)}$ δίνονται :

ii) Μέσο κόστος: Για το κριτήριο του μέσου κόστους απο τις σχέσεις **** της σελ. *** παίρνουμε ότι για να είναι βέλτιστη μια πολιτική π^0 πρέπει να ικανοποιεί το εξής σύστημα εξισώσεων:

$$\phi(x,\pi^0) = \min_{\alpha \in A(x)} \phi(\overline{\alpha},\pi^0) \quad \forall x \in X \tag{1}$$

$$\phi(x,\pi^0) + \delta(x,\pi^0) = \min_{\alpha \in A(x)} \{c(x,\alpha) + \delta(\overline{\alpha},\pi^0)\} \quad \forall x \in X \tag{2}$$

όπου εδώ με το α συμβολίζεται το βέλος, και με το $\overline{\alpha}$ ο επόμενος κόμβος. Για τη λύση του προβλήματος εφαρμόζουμε τη διαδικασία δ.π.π. ανάλογα με τη λύση του εισαγωγικού προβλήματος, της σελ. 2.34. Ετσι ξεκινάμε με μια αρχική πολιτική

$\pi^0 : \ \alpha^{(0)}(1) = 1 , \ \alpha^{(0)}(2) = 1, \ \alpha^{(0)}(3) = 3.$

Η πολιτική π^0 συνεπάγεται τη δυναμική του διπλανού σχήματος.

*Στην π^0 αντιστοιχούν οι εξής εξισώσεις υπολογισμού των ϕ, δ (ανάλογες των ****).*

$$\begin{aligned}
&\phi(1) = \phi(1)\\
&\phi(2) = \phi(1)\\
&\phi(3) = \phi(3)\\
&\phi(1) + \delta(1) = 2 + \delta(1)\\
&\phi(2) + \delta(2) = 4 + \delta(1)\\
&\phi(3) + \delta(3) = 1 + \delta(3)\\
&\delta(3) = 0 \text{ (``αυθαίρετα'')}\\
&\delta(1) = 0 \text{ (``αυθαίρετα'')}
\end{aligned}$$

Απο τις εξισώσεις αυτές παίρνομε:

$$\begin{aligned}
&\phi(1) = \phi(2) = 2,\\
&\phi(3) = 1,\\
&\delta(1) = 0\\
&\delta(2) = 2\\
&\delta(3) = 3
\end{aligned}$$

Εξετάζουμε τώρα αν η πολιτική π^o ικανοποιεί τις εξισώσεις 1,2. Συγκεκριμένα πρέπει:

$$\begin{aligned}
(1)\quad &\phi(1) \le \phi(3)\\
&\phi(1) \le \phi(1)\\
&\phi(2) \le \phi(1)\\
&\phi(3) \le \phi(3)\\
&\phi(3) \le \phi(2)\\
&\phi(3) \le \phi(2)\\
(2)\quad &\phi(1) + \delta(1) \le 2 + \delta(1)\\
&\phi(1) + \delta(1) \le 10 + \delta(3)\\
&\phi(3) + \delta(3) \le 1 + \delta(3)\\
&\phi(3) + \delta(3) \le 3 + \delta(2)
\end{aligned}$$

Απο τις σχέσεις αυτές δεν ισχύει μόνο η πρώτη ($\phi(1) \le \phi(3)$) Γι` αυτό αναπροσαρμόζουμε την πολιτική παίρνοντας $\alpha(1) = 3$, και αφήνοντας τις υπόλοιπες αποφάσεις ίδιες. Ετσι παίρνουμε την πολιτική π^1. $\alpha^1(1) = 3, \ \alpha^{(1)}(2) = 1, \ \alpha^{(1)}(3) = 3$, της οποίας η δυναμική φαίνεται παρακάτω:

Οι εξισώσεις που αντιστοιχούν στην π^4 είναι:

$$\begin{aligned}
&\phi(1) = \phi(3)\\
&\phi(3) = \phi(3)\\
&\phi(2) = \phi(1)\\
&\phi(1) + \delta(1) = 10 + \delta(3)\\
&\phi(2) + \delta(2) = 4 + \delta(1)\\
&\phi(3) + \delta(3) = 1 + \delta(3)\\
&\delta(3) = 0
\end{aligned}$$

Απ' αυτές παίρνουμε : $\phi(1) = \phi(2) = \phi(3) = 1, \ \delta(1) = 9, \ \delta(2) = 12, \ \delta(3) = 0$.

Εξετάζουμε πάλι τις συνθήκες βελτιστότητας. Οι ανισώσεις της ομάδας (1) ισχύουν προφανώς σαν εξισώσεις. Ετσι εξετ`αζουμε τις (2).

$$\begin{aligned}
&\phi(1) + \delta(1) \le 2 + \delta(1) \Rightarrow 1 + 9 \le 2 + 9\\
&\phi(1) + \delta(1) \le 10 + \delta(3) \Rightarrow 1 + 9 \le 10 + 0\\
&\phi(3) + \delta(3) \le 1 + \delta(3) \Rightarrow 1 \le 1\\
&\phi(3) + \delta(3) \le 3 + \delta(2) \Rightarrow 1 \le 3 + 12
\end{aligned}$$

Επειδή ισχύουν και οι εξισώσεις της ομάδας 2 , η π^1 είναι η βέλτιστη πολιτική για το πρόβλημα μέσου κόστους.

Επομένως, με $x_o = 1$, πρέπει να πάμε πρώτα στην 3 και μετά να παραμείνουμε μόνιμα εκεί.

ΠΡΟΒΛΗΜΑ 3: Για να μοντελοποιήσουμε το πρόβλημα με την βοήθεια του Δυναμικο`υ Προγραμματισμο`υ, πρέπει να προσδιορίσουμε τους χώρους καταστάσεων και αποφάσεων, τη δυναμική, τη συνάρτηση κόστους/κέρδους και τις αρχικές και τερματικές συνθήκες.

Επειδή δεν επιτρέπονται διακοπές εργασιών που εξυπηρετούνται, θα έχουμε αποφάσεις μόνο στις περιόδους που κάποιος σταθμός εργασίας δεν δουλεύει, έχει τελειώσει δηλ. γενικά την εξυπηρέτηση τις προηγούμενης εργασίας.

Το κόστος που θέλουμε να ελαχιστοποιήσουμε είναι ο συνολικός χρόνος μέχρι να τελειώσουν όλες οι εργασίες. Αυτό είναι ισοδύναμο με την εξής συνάρτηση τρέχοντος κόστους.

$$c(x,\alpha) = \begin{cases} 1, \text{ αν υπάρχει τουλάχιστον μια εργασία στο σύστημα} \\ 0, \text{ διαφορετικά} \end{cases}$$

Για τον προσδιορισμό της κατάστασης του συστήματος πρέπει να ξέρουμε ποιές ακριβώς εργασίες περιμένουν να εξυπηρετηθούν, ενώ όπως θα δούμε, δεν χρειάζεται να ξέρουμε ποιές εξυπηρετούνται αυτή τη στιγμή.

Πραγματικά για να προσδιορίσουμε την αρχική χρονική στιγμή που θα χρειαστεί να πάρουμε απόφαση, πρέπει να ξέρουμε μετά απο πόση ώρα θα ελευθερωθεί κάποιος σταθμός. Αυτό το βρίσκουμε αν ξέρουμε μετά απο πόσο χρόνο θα ελευθερωθεί <u>κάθε</u> σταθμός. Επομένως, πρέπει να προσθέσουμε κι αυτές τις πληροφορίες στην κατάσταση του συστήματος. Ετσι συνολικά η κατάσταση θα προσδιορίζεται απ`ο το διάνυσμα $(t_1,t_2,\ldots,t_m,A)$, όπου $t_1,\ldots,t_m$ είναι οι χρόνοι που ορίστηκαν παραπάνω, δηλ. t_i συμβολίζει το χρονικό διάστημα που θα είναι ακόμα δεσμευμένος ο σταθμός i, ενώ το A είναι <u>το σύνολο</u> των εργασιών των οποίων η εξυπηρέτηση δεν έχει αρχίσει. Ο χώρος αποφάσεων είναι προφανώς το σύνολο A.

Αν τώρα με $v(t_1,t_2,\ldots t_m,A)$ συμβολίζουμε <u>τον ελάχιστο</u> χρόνο που απαιτείται για να τελειώσουν όλες οι εργασίες του συνόλου A, μπορούμε να εφαρμόσουμε την αρχή της βελτιστότητας για να βρούμε τις εξισώσεις του Δυναμικο`υ Προγραμματισμο`υ. Συγκεκριμένα, αν $t_1,\ldots,t_m \neq 0$, δηλ. αν όλοι οι σταθμοί είναι απασχολημένοι, δεν παίρνουμε απόφαση τώρα, αλλά μετά απ`ο διάστημα ίσο με $t = min(t_1,\ldots,t_m)$, οπότε θα έχουμε φθάσει στην κατάσταση $(t_1 - t, t_2 - t, \ldots, t_m - t, A)$ όπου ένα τουλάχιστον απ`ο τα $t_i - t$ είναι `ισο με μηδ`εν, ενώ θα έχουμε πληρώσει κόστος ίσο με t, δηλ. με το χρονικό διάστημα που πέρασε.

Επομένως όταν $t_i \neq 0,\ i = 1,\ldots,m$, έχουμε :

$$u(t_n,\ldots,t_m,A) \ = \ t + 0(t_1 - t,\ldots,t_m t,A)\ ,\ \text{όπου}\ t = min(t_1,\ldots,t_m)\ .$$

Οταν τώρα ένα ή περισσότερα $t_l = 0$, πρέπει να διαλέξουμε ισάριθμες εργασίες απ`ο το σύνολο A, για να τις αναθέσουμε στους αδρανείς σταθμούς.

Εστω, χωρίς βλάβη της γενικότητας, ότι οι κ πρώτοι σταθμοί είναι αδρανείς. Τότε αν διαλέξουμε κ εργασίες από το A, η επόμενη κατάσταση, απο $(0,\ldots,0,t_{\kappa+1},\ldots,\alpha)$ θα είναι ($T_{i_1}, T_{i_2}, \ldots, T_{i_\kappa}, t_{\kappa+1}, t_m, A - \{E_{i_1}, E_{i_2},\ldots, E_{i_\kappa}\})$, δηλαδ`η οι κ πρώτοι σταθμοί θα έχουν "δουλειά" ο καθένας για διάστημα ίσο με τη διάρκεια της εργασίας που του ανατέθηκε, ενώ απο το σύνολο A θα αφαιρεθούν οι εργασίες αυτές. Ετσι η εξίσωση θα είναι :

$$v(0,\ldots,0,t_{\kappa+1},\ldots,t_m) \ = \min_{(\alpha_1,\ldots,\alpha_\kappa) \in A} \left\{ v(T_{\alpha_1},\ldots,T_{\alpha_\kappa},\ t_{\kappa+1}\ ,\ t_m,\ A - \{\alpha_1,\ldots,\alpha_\kappa\}) \right\}.$$

Τερματική συνθήκη $v(t_1,\ldots,t_m,\emptyset) = max(t_1,\ldots,t_m)$.

Ζητάμε το $v(0,\ldots,0,\{E_1,\ldots,E_n\})$.

ΠΡΟΒΛΗΜΑ 4: *Είναι πρόβλημα κατανομής πόρων με 2 πόρους (2 εξισώσεις) και 2 περιόδους (2 μεταβλητές). Ακριβώς ανάλογα με τα μοντέλα που υπάρχουν στις σημειώσεις, το μοντέλο Δυναμικο`υ Προγραμματισμο`υ του προβλήματος είναι :*

$$v(2,B_1,B_2) = \hat{c}(B_1,B_2) = 10 \min\left\{\tfrac{B_1}{2},\tfrac{B_2}{3}\right\},$$
$$v(1,B_1,B_2) = \max_{x \in \{0,1,2,4,7\}}\{5x + 10v(2,B_1 - 3x, B_2 - x)\},$$

όπου ζητάμε το $v(1,12,8)$.

5 . ΕΞΕΤΑΣΤΙΚΗ ΠΕΡΙΟΔΟΣ ΙΟΥΝΙΟΥ 88

ΛΥΣΕΙΣ

ΠΡΟΒΛΗΜΑ 1: *Είναι φανερό ότι στην κατάσταση του συστήματος θα πρέπει να περιλαμβάνονται οι εργασίες που δεν έχουν εκτελεσθεί. Συμβολίζουμε με* X *το σύνολο αυτών των εργασιών.*

Ετσι η κατάσταση είναι το σ`υνολο X. *Εύκολα επίσης προκύπτει ότι ο χώρος αποφάσεων* A *στην κατάσταση* X *ταυτίζεται με το σύνολο* X. *Με* $v(X)$ *συμβολίζουμε το ελάχιστο κόστος απο τη στιγμή* t *ως το τέλος της διαδικασίας, όταν το σύνολο των εργασιών που απομένουν είναι* X.

Οταν βρισκ`ομαστε στην κατ`ασταση X, *εκτ`ος απ`ο τις εργασ`ιες που δεν `εχουν εκτελεστε`ι, ξ`ερουμε και ποιες `εχουν εκτελεστε`ι, και επομ`ενως μπορο`υμε να βρο`υμε το χρ`ονο που `εχει περ`ασει απ`ο την αρχ`η της διαδικασ`ιας. Αυτ`ος ε`ιναι `ισος με* $t(X) = \sum_{i \notin X} T_i$ *, γιατ`ι υποθ`ετουμε `οτι δεν υπ`αρχουν καθυστερ`ησεις στη διαδοχ`η εκτ`ελεσης των εργασι`ων. Αν στην κατάσταση* X *πάρουμε την απόφαση να εκτελέσουμε την εργασία* $E_j \in X$, *τότε αυτή θα τελειώσει τη χρονική στιγμή* $t(X) + T_j$ *, και επομένως η συνεισφορά της στο συνολικό χρόνο ροής θα είναι* $t + T_j$ *. Οσον αφορά στη δυναμική, το σύνολο* X' *θα είναι* $X - \{E_j\}$ *, επειδή η* E_j *θα έχει εκτελεστεί. Ετσι είναι εύκολο να βρούμε την εξίσωση βελτιστοποίησης :*

$$v(X) = \min_{E_j \in X}\{t(X) + T_j + v(X - \{E_j\})\} .$$

Οριακή συνθήκη στο τέλος της διαδικασίας : $v(\emptyset) = 0$.

Ζητάμε να βρούμε το $v(\{E_1,E_2,\dots,E_n\})$.

ΠΡΟΒΛΗΜΑ 2: *Παίρνουμε σαν κατάσταση την ηλικία ενός μηχανήματος στο τέλος της περιόδου που εξετάζουμε. Απόφαση για την συνέχιση ή την αντικατάσταση μπορούμε να πάρουμε αν το μηχάνημα δουλεύει, ενώ διαφορετικά το αντικαθιστούμε οπωσδήποτε με καινούργιο.*

Συμβολίζουμε με κ *την ηλικία, όπως ορίστηκε πιο πάνω. Είναι φανερό ότι, αν* $\kappa = 0$ *, δεν κάνουμε αντικατάσταση, ενώ αν* $\kappa = 3$ *κάνουμε οπωσδήποτε.*
Αν με X *συμβολίσουμε την κατάσταση όπου το μηχάνημα δεν δουλεύει, τότε αν πάρουμε απόφαση να το αφήσουμε να δουλέψει μια περίοδο, απ`ο την κατάσταση* κ *θα πάμε με πιθανότητα* $1 - r_\kappa$ *στην κατάσταση* $\kappa + 1$ *και με πιθανότητα* r_κ *στην* X*:*

όπου $\kappa = 0,1,2$ *και* r_κ *η πιθανότητα να χαλάσει το μηχάνημα όταν έχει ηλικία* κ .

Είναι εύκολο να δούμε ότι αυτή η πιθανότητα r_κ *δίνεται απ`ο :*

$$r_\kappa = P\left(\frac{Z=\kappa}{Z\geq\kappa}\right) = \frac{P(Z=\kappa, Z\geq\kappa)}{P(Z\geq\kappa)} = \frac{P(Z=\kappa)}{P(Z\geq\kappa)} = \frac{P_\kappa}{P_\kappa+\cdots+P_3}\ ,\ \kappa = 0,1,2\ ,$$

όπου Z *η ζωή του μηχανήματος, και θέτουμε* $p_0 = 0$ *, αφού ένα καινούργιο μηχάνημα έχουμε δεχθεί ότι δεν χαλάει τον πρώτο χρόνο της ζωής του. Ετσι*

$$r_0 = 0\ ,\quad r_1 = \frac{P_1}{P_1+P_2+P_3} = P_1\ ,\quad r_2 = \frac{P_2}{P_2+P_3}\ .$$

Ας δούμε τώρα τη συνάρτηση κόστους και τη δυναμική. Υποθέτουμε ότι η αντικατάσταση ενός μηχανήματος γίνεται ακαριαία. Επίσης υποθέτουμε ότι αν το μηχάνημα αφεθεί να λειτουργήσει και χαλάσει, αυτή η βλάβη θα συμβεί αμέσως, δηλ. στην αρχή της επόμενης περιόδου, οπότε θα αντικατασταθεί αμέσως με καινούργιο. Ετσι έχουμε το παρακάτω διάγραμμα μεταπηδήσεων: (με $\alpha = 0$ *συμβολ`ιζεται η απόφαση να κρατηθεί το μηχάνημα, και με* $\alpha = 1$ *η απόφαση να αντικατασταθεί. Κατά τα άλλα ακολουθείται ο συμβολισμός των σημειώσεων, όπου δίπλα σε κάθε πιθανότητα μεταπήδησης σημειώνεται και το αντίστοιχο κόστος, όταν αυτό υπάρχει).*

Με d *συμβολίζουμε το κόστος αντικατάστασης χαλασμένου μηχανήματος, και με* c *το κόστος αντικατάστασης μηχανήματος που λειτουργεί. Οι μεταπηδήσεις για* $\alpha = 1$ *γίνονται κατ' ευθείαν στην κατάσταση* 1 *και όχι στην* 0 *, επειδή έχουμε υποθέσει ότι η αντικατάσταση γίνεται αμέσως, και έτσι μετά απο μια χρονική περίοδο το μηχάνημα θα βρίσκεται πια στην κατάσταση* 1 .
Με βάση το παραπάνω σχήμα είναι εύκολο να γράψουμε τις εξισώσεις βελτιστοποίησης του μέσου αναμενόμενου κόστους :
Επειδή όλες οι δυνατές πολιτικές δημιουργούν μόνο μια αλυσίδα, οι εξισώσεις της μορφής $\phi(i) = \min\limits_{\alpha\in Ai}\{\sum\limits_j P_{ij}(\alpha)\phi(j)\}$ *ισχύουν ταυτολογικά :* $\phi(1) = \phi(2) = \phi(3) = \phi$.

Ετσι μένει μόνο το δεύτερο σύνολο εξισώσεων.

$$\phi + \delta(1) = \min\{\, r_1 d + r_1\delta(1) + (1 - r_1)\delta(2)\ ,\ c + \delta(1)\,\}$$
$$\phi + \delta(2) = \min\{\, r_2 d + r_2\delta(1) + (1 - r_2)\delta(3)\ ,\ c + \delta(1)\,\}$$
$$\phi + \delta(3) = c + \delta(1)\ .$$

Στο παράδειγμά μας έχουμε :

$$P_1 = \tfrac{1}{4}\ ,\ P_2 = \tfrac{1}{2}\ ,\ P_3 = \tfrac{1}{4}\ \Rightarrow\ r_1 = P_1 = \tfrac{1}{4}\ ,\ r_2 = \tfrac{P_2}{P_2+P_3} = \tfrac{2}{3}\ ,\ d = 24\ ,\ c = 10\ .$$

Ετσι οι εξισώσεις γίνονται :

$$\phi + \delta(1) = min\{\ 6 + \tfrac{1}{4}\delta(1) + \tfrac{3}{4}\delta(2)\ ,\ 10 + \delta(1)\ \}$$
$$\phi + \delta(2) = min\{\ 16 + \tfrac{2}{3}\delta(1) + \tfrac{1}{3}\delta(3)\ ,\ 10 + \delta(1)\ \}$$
$$\phi + \delta(3) = 10 + \delta(1)\ .$$

Για τη λύση του προβλήματος εφαρμόζουμε τη μέθοδο Δ.Π.Τ. Ξεκινάμε με την πολιτική που αντικαθιστά πάντοτε, δηλ. $\alpha^{(0)}(1) = \alpha^{(0)}(2) = \alpha^{(0)}(3) = 1$. *Γι' αυτήν έχουμε :*

$$\left.\begin{array}{l} \phi^{(0)} + \delta^{(0)}(1) = 10 + \delta^{(0)}(1) \\ \phi^{(0)} + \delta^{(0)}(2) = 10 + \delta^{(0)}(1) \\ \phi^{(0)} + \delta^{(0)}(3) = 10 + \delta^{(0)}(1) \end{array}\right\} \Rightarrow \quad \begin{array}{l} \phi^{(0)} = 10 \\ \delta^{(0)}(1) = \delta^{(0)}(2) = \delta^{(0)}(3) = 0\ . \end{array}$$

Θέτουμε $\delta^{(0)}(1) = 0$.

Εξετάζουμε αν ικανοποιούνται οι εξισώσεις βελτιστοποίησης στις καταστάσεις 1,2 όπου έχουμε εναλλακτικές αποφάσεις.

$$\phi^{(0)} + \delta^{(0)}\ \overset{?}{\leq}\ 6 + \tfrac{1}{4}\delta^{(0)}(1) + \tfrac{3}{4}\delta^{(0)}(2)\ \Leftrightarrow\ 10 \overset{?}{\leq} 6 \quad \underline{\text{όχι}}$$

$$\phi^{(0)} + \delta^{(0)}(2)\ \overset{?}{\leq}\ 16 + \tfrac{2}{3}\delta^{-(0)}(1) + \tfrac{1}{3}\delta^{(0)}(2)\ \Leftrightarrow\ 10 \overset{?}{\leq} 16 \quad \underline{\text{ναι}}$$

Βλέπουμε ότι η βελτιστότητα παραβιάζεται στην κατάσταση 1 , και επομένως παίρνουμε την καινούργια πολιτική.

$$\alpha^{(1)}(1) = 0,\ \ \alpha^{(1)}(2) = \alpha^{(1)}(3) = 1$$

$$\left.\begin{array}{l} \phi^{(1)} + \delta^{(1)}(1) = 6 + \tfrac{1}{4}\delta^{(1)} + \tfrac{3}{4}\delta^{(1)}(2) \\ \phi^{(1)} + \delta^{(1)}(2) = 10 + \delta^{(1)} \\ \phi^{(1)} + \delta^{(1)}(3) = 10 + \delta^{(1)}(1) \end{array}\right\} \Rightarrow$$

Θέτουμε $\delta^{(1)}(1) = 0$

$$\delta^{(1)}(2)\ =\ \delta^{(1)}(3)\ =\ 10 - \phi^{(1)}\ ,$$
$$\phi^{(1)}\ =\ 6 + \tfrac{3}{4}(10 - \phi^{(1)})\ =\ 6 + 7.5 - \tfrac{3}{4}\phi^{(1)}\ \Rightarrow\ \tfrac{7}{4}\phi^{(1)} =\ \tfrac{27}{2}\ \Rightarrow\ \phi^{(1)} =\ \tfrac{54}{7}\ =\ 7.7\ ,$$
$$\delta^{(1)}(2)\ \ \delta^{(1)}(1) = 0,\ \ \delta^{(1)}(2) = \delta^{(1)}(3) = 2.3\ \ .$$

Εξετάζουμε πάλι τη βελτιστότητα :

$$\phi^{(1)} + \delta^{(1)}(1)\ \overset{?}{\leq}\ 10 + \delta^{(1)}(1) \quad \Rightarrow\ \phi^{(1)} \overset{?}{\leq}\ 10 \quad \underline{\text{ναι}}$$

$$\phi^{(1)} + \delta^{(1)}(2)\ \overset{?}{\leq}\ 16 + \tfrac{2}{3}\delta^{(1)}(1) + \tfrac{1}{3}\delta^{(1)}(2)\ \Rightarrow\ 10\ \overset{?}{\leq}\ 16 + \tfrac{1}{3} \cdot\ 2.3 \quad \underline{\text{ναι}}$$

Βλέπουμε ότι οι εξισώσεις βελτιστοποίησης ικανοποιούνται και επομένως η βέλτιστη πολιτική είναι η π, με $\alpha^*(1) = 0$, $\alpha^*(2) = 1$, δηλ. η πολιτική που αντικαθιστ`α το μηχάνημα μόλις συμπληρώσει 2 χρόνια ομαλής λειτουργίας, ή βέβαια όταν χαλάσει ενδιάμεσα.
<u>Σημείωση:</u> Αν κάναμε την παραδοχή ότι η αντικατάσταση του μηχανήματος κρατάει μια χρονική περ`ιοδο, τότε θα είχαμε και μεταπήδηση στην κατάσταση 0 , επειδή στην αρχή της επόμενης περιόδου μετά την κατάσταση το μηχάνημα θα ήταν καινούργιο. Η επέκταση των εξισ`ωσεων σ' αυτή την περίπτωση είναι εύκολο να γίνει.

<u>ΠΡΟΒΛΗΜΑ 3:</u> *Επειδή δεν υπάρχει περιορισμός στον αριθμό δοκιμών, και θεωρητικά μπορεί ο παίκτης να παίξει άπειρες φορές, το πρόβλημα ε`ιναι άπειρου ορίζοντα με πιθανοθεωρητική δυναμική. Ας συμβολίσουμε με x την πιθανότητα να υπάρχει ο θησαυρός, σε κάποιο βήμα του παιχνιδιού.*
Δηλαδή $x = P\{Y\}$ όπου με Y συμβολίζουμε το ενδεχέμενο να υπάρχει ο θησαυρός και με B το ενδεχόμενο να βρεθεί (άσχετα απο το αν υπάρχει ή όχι). Τότε, αν ο παίκτης αποφασίσει να δοκιμάσει μια φορά, η πιθανότητα να βρει τον θησαυρό είναι ίση με :

$$P(B) = P(B/Y) \cdot P(Y) + P(B/\text{όχι } Y) \cdot P(\text{όχι } Y) .$$

Ομως απο τα δεδομένα του προβλήματος $P(B/Y) = P$, ενώ είναι φανερό ότι $P(B/\text{όχι } Y) = 0$.
Ετσι $P(B) = px$.

Δηλαδή ο παίκτης `η θα βρει το θησαυρό με πιθανότητα px , οπότε θα κερδίσει R μονάδες και θα τελειώσει το παιχνίδι, ή δεν θα τον βρει με πιθανότητα $1 - px$. Στην τελευταία αυτή περίπτωση ο παίκτης αποκτά περισσότερη πληροφόρηση για την πιθανότητα ύπαρξης του θησαυρού. Μετά απο μια αποτυχημένη προσπάθεια η πιθανότητα αυτή γίνεται, σύμφωνα με τον τύπο του Bayes :

$$x' = P(Y/\text{όχι } B) = \frac{P(Y \text{ και } (\text{όχι } B))}{P(\text{όχι } B)} =$$

$$= \frac{P(\text{όχι } B/Y) \cdot P(Y)}{P(\text{όχι } B)} = \frac{[1 - P(B/Y)] \cdot P(Y)}{1 - P(B)} = \frac{(1-P)\,x}{1-Px} .$$

Σύμφωνα λοιπόν με όσα είπαμε παραπάνω το διάγραμμα μεταπηδήσεων του προβλήματος είναι :

Με $\alpha = 0$ συμβολίζουμε την απόφαση για μια ακόμη δοκιμή, και με $\alpha = 1$ την απόφαση να σταματήσει. Στην απόφαση $\alpha = 0$ αντιστοιχεί "κέρδος" $-c$, δηλ. το κόστος μιας δοκιμής, ενώ υπάρχει και επιπλέον κέρδος R με πιθανότητα px .

Επομένως οι εξισώσεις βελτιστοποίησης που αντιστοιχούν στο συνολικό κέρδος είναι :

$$v(x) = \max\{ -c + pxR + pxv(\text{τέλος}) + (1 - px)\, v(x') ,\ 0 + v(\text{τέλος}) \}$$
Τερματική συνθήκη $v(\text{τέλος}) = 0$,

$$\Rightarrow\ v(x) = \max\{ -c + pxR + (1 - px)v\left(\frac{(1-p)x}{1-px}\right) , 0 \} , \quad 0 \le x \le 1 .$$

Επειδή αρχικά ξέρουμε ότι $P(Y) = \alpha$, ζητάμε την τιμή $v(\alpha)$.

<u>ΠΡΟΒΛΗΜΑ 4:</u> *Ξεκινάμε με την πολιτική* π^0 : $\alpha^{(0)}(1) = 1$, $\alpha^{(0)}(2) = 1$, $\alpha^{(0)}(3) = 2$.
Εχουμε :

$$\begin{aligned}
&\phi^{(0)}(1) = \phi^{(0)}(1)\\
&\phi^{(0)}(2) = \phi^{(0)}(1)\\
&\phi^{(0)}(3) = \phi^{(0)}(2)\\
&\phi^{(0)}(1) + \delta^{(0)}(1) = 5 + \delta^{(0)}(1)\\
&\phi^{(0)}(2) + \delta^{(0)}(2) = 4 + \delta^{(0)}(1)\\
&\phi^{(0)}(3) + \delta^{(0)}(3) = 2 + \delta^{(0)}(2)
\end{aligned}$$

Θέτουμε $\delta^{(0)}(1) = 0$

$$\begin{aligned}
&\phi^{(0)}(1) = \phi^{(0)}(2) = \phi^{(0)}(3) = \phi^{(0)}\\
&\phi^{(0)} = 5\\
&\phi^{(0)} + \delta^{(0)}(2) = 4\\
&\phi^{(0)} + \delta^{(0)}(3) = 2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow\ &\phi^{(0)}(1) = \phi^{(0)}(2) = \phi^{(0)}(3) = 5\\
&\delta^{(0)}(1) = 0,\quad \delta^{(0)}(2) = -1,\quad \delta^{(0)}(3) = -3
\end{aligned}$$

Ελέγχουμε τη βελτιστότητα:

$$\phi^{(0)}(1) + \delta^{(0)}(1) \overset{;}{\leq} 7 + \delta^{(0)}(3) \ \Leftrightarrow\ 5 \overset{;}{\leq} 7 - 3 \ \Leftrightarrow\ 5 \overset{;}{\leq} 4 \quad \underline{\textit{όχι}}$$

$$\phi^{(0)}(1) + \delta^{(0)}(1) \overset{;}{\leq} 8 + \delta^{(0)}(2) \ \Leftrightarrow\ 5 \overset{;}{\leq} 8 - 1 \ \Leftrightarrow\ 5 \overset{;}{\leq} 7 \quad \underline{\textit{ναι}}$$

$$\phi^{(0)}(2) + \delta^{(0)}(2) \overset{;}{\leq} 6 + \delta^{(0)}(3) \ \Leftrightarrow\ 5 - 1 \overset{;}{\leq} 6 - 3 \ \Leftrightarrow\ 4 \overset{;}{\leq} 3 \quad \underline{\textit{όχι}}$$

Η βελτιστότητα παραβιάζεται στις καταστάσεις 1 και 2. Η νέα πολιτική π^1 *είναι :* $\alpha^{(1)}(1) = 3$, $\alpha^{(1)}(2) = 3$, $\alpha^{(1)}(3) = 2$.

$$\begin{aligned}
&\phi^{(0)}(1) = \phi^{(1)}(3)\\
&\phi^{(1)}(2) = \phi^{(1)}(3)\\
&\phi^{(1)}(3) = \phi^{(1)}(2)\\
&\phi^{(1)}(1) + \delta^{(1)}(1) = 7 + \delta^{(1)}(3)\\
&\phi^{(1)}(2) + \delta^{(0)}(2) = 6 + \delta^{(1)}(3)\\
&\phi^{(1)}(3) + \delta^{(1)}(3) = 2 + \delta^{(1)}(2)
\end{aligned}$$

Θέτουμε $\delta^{(1)}(3) = 0$

$$\phi^{(1)}(1) = \phi^{(1)}(2) = \phi^{(1)}(3) = \phi^{(1)}$$

$$\left.\begin{aligned}
&\phi^{(1)} + \delta^{(1)}(1) = 7\\
&\phi^{(1)} + \delta^{(1)}(2) = 6\\
&\phi^{(1)} = 2 + \delta^{(1)}(2)
\end{aligned}\right\} \Rightarrow$$

$$\left.\begin{aligned}
&\phi^{(1)} + \delta^{(1)}(1) = 7\\
&2 + 2\delta^{(1)}(2) = 6\\
&\phi^{(1)} = 2 + \delta^{(1)}(2)
\end{aligned}\right\} \Rightarrow$$

$$\Rightarrow \phi^{(1)} = 4\ ,\quad \delta^{(1)} = 3\ ,\ \delta^{(1)}(2) = 2\ ,\ \delta^{(1)}(3) = 0\ .$$

Εξετάζουμε τη βελτιστότητα :

$$\phi^{(1)} + \delta^{(1)}(1)\ \overset{;}{\leq}\ 5 + \delta^{(1)}(1) \ \Leftrightarrow\ 7 \overset{;}{\leq} 5 + 3 \quad \underline{\textit{ναι}}$$

$$\phi^{(1)} + \delta^{(1)}(1)\ \overset{;}{\leq}\ 8 + \delta^{(1)}(2) \ \Leftrightarrow\ 7 \overset{;}{\leq} 8 + 2 \quad \underline{\textit{ναι}}$$

$$\phi^{(1)} + \delta^{(1)}(2) \overset{;}{\leq} 8 + \delta^{(1)}(1) \Leftrightarrow 6 \overset{;}{\leq} 8 + 3 \quad \underline{\nu\alpha\iota}$$

Αφού ικανοποιούνται οι συνθήκες βελτιστότητας, η βέλτιστη πολιτική είναι η π^1 *με* $\alpha^*(1) = 3$, $\alpha^*(2) = 3$, $\alpha^*(3) = 2$, *και* $\phi^* = 4$.

ΠΡΟΒΛΗΜΑ 5: *Γράφουμε το πρόβλημα σε ανεπτυγμένη μορφή:*

$$\begin{array}{ll} (max) & z = c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ \upsilon.\pi. & \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = b_1 \\ & \alpha_{21}x_1 + \ldots + \alpha_{2n}x_n = b_2 \\ & \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n = b_m \\ & x_i \geq 0, \; x_i \in \mathbb{Z} \end{array}$$

Πρόκειται για πρόβλημα κατανομής πόρων, όπου έχουμε m *είδη πόρων και χρονικό ορίζοντα μήκους* n .

Οπως είναι γνωστό απο τις σημειώσεις και τα παραδείγματα που έχουν λυθεί, στην κατάσταση σε κάθε βήμα θεωρούμε τις ποσότητες των πόρων που απομένουν. Στην κατάσταση $(B_1^{(\kappa)}, B_2^{(\kappa)}, \ldots, B_m^{(\kappa)})$ *όπου βρισκόμαστε στο κ βήμα, η μεταβλητή* x_κ *μπορεί να πάρει τιμές στο σύνολο* $A(\kappa) = \{0,1,2,\ldots,[\min_i B_i^{(\kappa)}]\}$, *όπου με* $[x]$ *συμβολίζουμε το ακέραιο μέρος του* x . *Ετσι η εξίσωση βελτιστοποίησης είναι:*

$$v(\kappa,B_1^{(\kappa)},\ldots,B_m^{(\kappa)}) = \max_{x \in A(\kappa)} \{c_\kappa x + v(\kappa+1,B_1^{(\kappa)} - \alpha_{1\kappa}x,\ldots,B_m^{(\kappa)} - \alpha_{m\kappa}x)\}$$

με $v(n+1,B_1^{(n+1)},\ldots,B_m^{(n+1)}) = 0$,

όπου ζητάμε το $v(1,b_1,b_2,\ldots b_m)$.

6. ΕΞΕΤΑΣΤΙΚΗ ΠΕΡΙΟΔΟΣ ΣΕΠΤΕΜΒΡΙΟΥ 88

ΛΥΣΕΙΣ

ΠΡΟΒΛΗΜΑ 1: *Θα λύσουμε το πρόβλημα με τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο των πολιτικών. Οι εξισώσεις βελτιστοποίησης είναι :*

$$v(1) = min\{2 + \beta v(1),\ 6 + \beta v(2)\}$$
$$v(2) = min\{2 + \beta v(1), 4 + \beta v(3)\}$$
$$v(3) = 1 + \beta v(3)\ .$$

$0^{\underline{o}}$ βήμα : $\pi^0 = \{ (1,1)\ ,\ (2,1)\ ,\ (3,3) \}$

$$w^{(0)}(1) \ = \ 2 + 0.9w^{(0)}(1) \ \Rightarrow \ w^{(0)}(1) = \tfrac{2}{0.1} = 20$$
$$w^{(0)}(2) \ = \ 2 + 0.9w^{(0)}(1) \ = \ 2 + 0.9.20 \ = \ 20$$
$$w^{(0)}(3) \ = \ 1 + 0.9w^{(0)}(3) \ \Rightarrow \ w^{(0)}(3) = \tfrac{1}{0.1} = 10\ .$$

Ελεγχος: $\Delta w^{(0)}(1,2) = [6 + 0.9w^{(0)}(2)] - w^{(0)}(1) \ = \ 24 - 20 = 4 \ > \ 0$

$\Delta w^{(0)}(2,3) = [4 + 0.9w^{(0)}(3)] - w^{(0)}(2) \ = \ 13 - 20 = \ -7 \ < \ 0 \quad \leftarrow$ *αλλαγή*

$1^{\underline{o}}$ βήμα : $\pi^1 = \{ (1,1)\ ,\ (2,3)\ ,\ (3,3) \}$

$$w^{(1)}(1) = 2 + 0.9w^{(1)}(1) \ \Rightarrow \ w^{(1)}(1) = 20$$
$$w^{(1)}(2) = 4 + 0.9w^{(1)}(3)$$
$$w^{(1)}(3) = 1 + 0.9w^{(1)}(3) \ \Rightarrow \ w^{(1)}(3) = 10 \ \Rightarrow \ w^{(1)}(2) = 13\ .$$

Ελεγχος: $\Delta w^{(1)}(1,2) = [6 + 0.9w^{(1)}(2)] - w^{(1)}(1) \ = \ 17.7 - 20 = \ -2{,}3 \quad \leftarrow$ *αλλαγή*

$\Delta w^{(1)}(2,1) = [2 + 0.9w^{(1)}(1)] - w^{(1)}(2) \ = \ 20 - 13 = 7 \ > \ 0\ .$

$2^{\underline{o}}$ βήμα : $\pi^2 = \{ (1,2)\ ,\ (2,3)\ ,\ (3,3) \}$

$$w^{(2)}(1) = 6 + 0.9w^{(2)}(2)$$
$$w^{(2)}(2) = 4 + 0.9w^{(2)}(3)$$
$$w^{(2)}(3) = 1 + 0.9w^{(2)}(3) \ \Rightarrow \ w^{(2)}(3) = 10\ ,\ w^{(2)}(2) = 13\ ,\ \ w^{(2)}(1) = 17.7\ .$$

Ελεγχος: $\Delta w^{(2)}(1,1) = [2 + 0.9w^{(2)}(1)] - w^{(2)}(1) \ = \ 0.23 \ > \ 0$

$\Delta w^{(2)}(2,1) = [2 + 0.9w^{(2)}(1)] - w^{(2)}(2) \ = \ 4.93 \ > \ 0\ .$

Αρα η βέλτιστη πολιτική είναι η π^2 *με* $v(1) = 17.7$, $v(2) = 13$, $u(3) = 10$, *και ξεκινώντας με* $x_0 = 1$, *πρέπει να περάσουμε απ`ο την 2 και να φθάσουμε στην 3 , όπου θα μείνουμε μόνιμα.*

ΠΡΟΒΛΗΜΑ 2: *Οι εξισώσεις βελτιστοποίησης για το αναμενόμενο συνολικό εκπτωτικό κόστος άπειρου ορίζοντα είναι :*

$$v(1) = min\{\ 3 + \beta v(3)\ ,\ 8 + \beta(\tfrac{1}{2}v(4))\ ,\ 4 + \beta(\tfrac{3}{4}v(2) + \tfrac{1}{4}v(3))\ \}$$
$$v(2) = min\{\ 1 + \beta v(1)\ ,\ 5 + \beta v(4)\ \}$$
$$v(3) = 2 + \beta v(1)$$
$$v(4) = 2 + \beta v(4) \ \Rightarrow \ v(4) = \tfrac{2}{1-\beta}\ .$$

Εφαρμόζουμε κι εδώ διαδοχικές προσεγγίσεις στο χώρο των πολιτικώ, ακριβώς όπως και στα προβλήματα με προσδιοριστική δυναμική.

Βήμα − 0 : $\pi^0 = \{ (1,\alpha_{11}) , (2,4) , (3,1) , (4,4) \}$

$$w^{(0)}(4) = \tfrac{2}{0.1} = 20$$
$$w^{(0)}(2) = 5 + 0.9w^{(0)}(4) = 5 + 18 = 23$$
$$w^{(0)}(1) = 8 + 0.9\tfrac{w^{(0)}(2)+w^{(0)}(4)}{2} = 8 + \tfrac{0.9}{2}[23 + 20] = 27.35$$
$$w^{(0)}(3) = 2 + 0.9w^{(0)}(1) = 26.62$$

Ελεγχος:
$$\Delta w^{(0)}(1,\alpha_{12}) = 4 + 0.9(\tfrac{3}{4}w^{(0)}(2) + \tfrac{1}{4}w^{(0)}(3)) - w^{(0)}(1) = -1.84 < 0 \quad \leftarrow \text{αλλαγή}$$
$$\Delta w^{(0)}(1,3) = 3 + 0.9w^{(0)}(3) - w^{(0)}(1) = -0.39 < 0$$
$$\Delta w^{(0)}(2,1) = 1 + 0.9w^{(0)}(1) - w^{(0)}(2) = 2.62 > 0$$

Βήμα 1 : $\pi^1 = \{ (1,\alpha_{12}) , (2,4) , (3,1) , (4,4) \}$

$$w^{(1)}(4) = 20$$
$$w^{(1)}(2) = 5 + 0.9w^{(0)}(4) = 23$$
$$w^{(1)}(3) = 2 + 0.9w^{(1)}(1)$$
$$w^{(1)}(1) = 4 + 0.9 \cdot [\tfrac{3}{4}w^{(1)}(2) + \tfrac{1}{4}w^{(1)}(3)] = 4 + 0.9 \cdot [\tfrac{3}{4} \cdot 23 + \tfrac{1}{4} \cdot (2 + 0.9w^{(1)}(1)] =$$
$$= 19.975 + 0.203\, w^{(1)}(1) \Rightarrow w^{(1)} = \tfrac{19.975}{0.797} = 25.06 .$$

Ελεγχος:
$$\Delta w^{(1)}(1,\alpha_{11}) = 8 + 0.9(\tfrac{1}{2}w^{(1)}(2) + \tfrac{1}{2}w^{(1)}(4)) - w^{(1)}(1) = 2.29 > 0$$
$$\Delta w^{(1)}(1,3) = 3 + 0.9w^{(1)}(3) - w^{(1)}(1) = 0.05 > 0$$
$$\Delta w^{(1)}(2,1) = 1 + 0.9w^{(1)}(1) - w^{(1)}(2) = 0.55 > 0$$

Αρα η π^1 είναι βέλτιστη. Ξεκινώντας με $X_0 = 1$ παίρνουμε την απόφαση α_{12} . Απ`ο κει αν πάμε στη 2 , προχωρούμε στην 4 όπου μένουμε μ`ονιμα, ενώ αν πάμε στην 3 , επιστρέφουμε πάλι στην 1 και συνεχίζουμε με τον ίδιο τρόπο.

ΠΡΟΒΛΗΜΑ 3: *Εστω $\alpha_1,\alpha_2,\alpha_3$ τα βάρη απο τα 1,2,3 που θα μεταφέρει το αεροπλάνο. Το πρόβλημα τώρα σε μαθηματική μορφή μπορεί να γραφτεί ως εξής :*

$$\begin{aligned} &(max) && \sum_{i=1}^{3} R_i(\alpha_i) \\ &\upsilon.\pi. && \alpha_1 + \alpha_2 + \alpha_3 \leq 10 \\ & && \alpha_1,\alpha_2,\alpha_3 \geq 0 . \end{aligned}$$

Είναι φανερό ότι πρόκειται για πρόβλημα κατανομής πόρων που μπορεί να λυθεί με Δυναμικό Προγραμματισμό. Συγκεκριμένα έχουμε $N = 3$ και μια κατάσταση x_κ, που εκφράζειτην ποσότητα του πόρου που απομένει για διάθεση στο βήμα κ . Η εξίσωση βελτιστοποίησης προκύπτει εύκολα :

$$v(\kappa,x_\kappa) = \max_{0 \leq \alpha_\kappa \leq x_\kappa} \{R_\kappa(\alpha_\kappa) + v(\kappa + 1,x_\kappa - \alpha_\kappa)\} , \ \kappa = 1,2,3$$

με τερματικη συνθήκη $v(4,x) = 0$, και αρχική $x_1 = 10$. Για να λύσουμε το πρόβλημα ξεκινάμε απ' το τέλος :

$$\kappa = 3 : \ v(3,x_3) = \max_{0 \leq \alpha_3 \leq x_3} \{R_3(\alpha_3)\} = \max_{0 \leq \alpha_3 \leq x_3} (2\alpha_3) .$$

Είναι φανερό οτι το max προκύπτει για $\alpha_3 = x_3$, οπότε :

$$\alpha_3^*(x_3) = x_3, \ v(3,x_3) = 2x_3 .$$

$$\kappa = 2 : \ v(2,x_2) = \max_{0 \leq \alpha_2 \leq x_2} \{R_2(\alpha_2) + v(3,x_2 - \alpha_2)\} =$$

$$= \max_{0 \le \alpha_2 \le x_2} \{\alpha_2^2 + 2\alpha_2 + 2(x_2 - \alpha_2)\} = \max_{o \le \alpha_2 \le x_2} \{\alpha_2^2 + 2x_2\}$$
$$= 2x_2 + \max_{0 \le \alpha_2 \le x_2} \{\alpha_2^2\}$$

Είναι πάλι πολύ εύκολο να δούμε ότι η α_2^2 *μεγιστοποιείται για* $\alpha_2 = x_2$ *, οπότε :*

$$\alpha_2^*(x_2) = x_2 \,,\; v(2,x_2) = 2x_2 + x_2^2 \;.$$

$\kappa = 1 :\;$ $v(1,x_1) = \max_{o \le \alpha_1 \le x_1} \{R_1(\alpha_1) + v(2,x_1 - \alpha_1)\} =$

$$= \max_{o \le \alpha_1 \le x_1} \{\alpha_1^2 + 2(x_1 - \alpha_1) + (x_1 - \alpha_1)^2\} = max\{\alpha_1^2 + 2x_1 - 2\alpha_1 + x_1^2 - 2\alpha_1 x_1 + \alpha_1^2\} =$$
$$= x_1^2 + 2x_1 + \max_{0 \le \alpha_1 \le x_1} \{2\alpha_1^2 - 2\alpha_1 - 2\alpha_1 x_1\} =$$
$$= x_1^2 + 2x_1 + 2\max_{0 \le \alpha_1 \le x_1} \{\alpha_1^2 - (1 + x_1)\alpha_1\} \,. \tag{1}$$

Η συνάρτηση $f(\alpha_1) = \alpha_1^2 - (1 + x_1)\alpha_1$ *είναι κυρτή (γιατί* $f''(\alpha_1) = 2 > 0$*) επομένως μεγιστοποιείται σ' ένα απ' τα άκρα του διαστήματος* $[0,x_1]$ *.*

$$f(0) = 0$$
$$f(x_1) = x_1^2 - (1 + x_1)x_1 = -x_1 \le 0 \;\; (\alpha\varphi o\acute{\upsilon}\; x_1 > 0) \,.$$

Επομένως $f(0) \ge f(x_1)$ *, και έτσι* $\max_{0 \le \alpha_1 \le x_1} f(\alpha_1) = f(0)$ *.*

Δηλαδή το max της (1) προκύπτει για $\alpha_1 = 0$ *, οπότε :*

$$\alpha_1^*(x_1) = 0 \,,\; v(1,x_1) = x_1^2 + 2x_1 + 2 \cdot 0 = x_1^2 + 2x_1 \,.$$

Αφού βρήκαμε $\alpha_1^* = 0,\; \alpha_2^*(x_2) = x_2,\; \alpha_3^*(x_3) = x_3,$ *έχουμε :*

$$x_1 = 10 \;\Rightarrow x_2 = x_1 - \alpha_1 = 10 \,,\quad \alpha_2^* = 10 \;\Rightarrow\; x_3 = x_2 - \alpha_2 = 0 \,,\quad \alpha_3^* = 0 \,.$$

Δηλαδή όλη η ποσότητα του πόρου κατανέμεται στη δραστηριότητα 2 , που σημαίνει ότι θα μεταφερθούν 10 τόννοι απο το υλικό 2 , και η συνολική αξία των μεταφερθέντων υλικών θα είναι :

$$v(1,10) = R_2(10) = 10^2 + 2 \cdot 10 = 120 \,.$$

<u>ΠΡΟΒΛΗΜΑ 4:</u> *Εστω ότι ακολουθούμε μια διαδρομή* $\delta_1 = x_1 = 1,\alpha_1,x_2,\alpha_2,\ldots,x_\kappa = N$*. Τότε η πιθανότητα το μήνυμα να φθάσει απο τον 1 στον N μέσω της* δ_1 *είναι :*

$$p(\delta_1) = p(x_1,\alpha_1) \cdot (x_2,\alpha_2) \cdots p(x_{\kappa-1},\alpha_{\kappa-1}) \,.$$

Βλέπουμε δηλαδή ότι μπορούμε να εφαρμόσουμε τις γνωστές έννοιες για την ελαχιστοποίηση του κόστους (ή ισοδύναμα τη μεγιστοποίηση του κέρδους σ' ένα δίκτυο) με μόνη τη διαφορά ότι εδώ το κέρδος ή η αξία μιας διαδρομής είναι το γινόμενο και όχι το άθροισμα των επιμέρους κερδών. Αυτή η "δυσκολία" μπορεί να αντιμετωπισθεί με δυο τρόπους:

<u>1ος τρόπος:</u> *Αφού όλες οι επιμέρους πιθανότητες είναι μη αρνητικές συμβολίζουμε με v(x) την πιθανότητα να φτάσει στο μήνυμα απο τον κόμβο x στον N, ακολουθώντας τη βέλτιστη διαδρομή. Τότε είναι φανερό ότι η εξίσωση βελτιστοποίησης, ακριβώς ανάλογα με τις σημειώσεις μπορεί να γραφτεί:*

$$v(x) = \max_{\alpha \in A(x)} \{p(x,\alpha) \cdot v(\overline{\alpha})\} \;\; x = 1,2,\ldots,N-1.$$

Επειδή έχουμε γινόμενο, η τερματική συνθήκη θα είναι $v(N) = 1$.

2os τρόπος: *Μπορούμε να χρησιμοποιήσουμε την ιδιότητα ότι ο λογάριθμος μιας συνάρτησης μεγιστοποιείται στα ίδια σημεία που μεγιστοποιείται η συνάρτηση. Ετσι αν πάρουμε σαν αξία της διαδρομής δ_1:*

$$w(\delta_1) = \ln p(\delta_1) = \sum_{i=1}^{\kappa} \ln p(x_i,\alpha_i)\ ,$$

τότε το μέτρο απόδοσης είναι προσθετικό, που ξέρουμε εύκολα να το αντιμετωπίσουμε. Συγκεκριμένα, έχουμε ένα δίκτυο που η αξία κάθε βέλους (x,α) είναι $\ln p(x,\alpha)$, ή θέλουμε να βρούμε τη διαδρομή απο το 1 στο N που έχει τη μεγαλύτερη αξία. Η εξίσωση βελτιστοποίησης για το πρόβλημα αυτό είναι βέβαια :

$$v(x) = \max_{\alpha \in A(x)} \{\ln p(x,\alpha) + v(\overline{\alpha})\} \quad x = 1,2,\ldots,N-1$$

οριακή συνθήκη $v(N) = 0$, και η βέλτιστη πολιτική είναι επίσης βέλτιστη και για το αρχικό πρόβλημα.

ΠΡΟΒΛΗΜΑ 5 : *Η εξίσωση $x_{t+1} = gx_t + b\alpha_t + z_t$, $t=0,\ldots,N-1$ εκφράζει τη δυναμική του συστήματος, ενώ απο την έκφραση του κόστους βρίσκουμε ότι το τρέχον κόστος : $c_t(x_t,\alpha_t) = bx_t^2 + c\alpha_t^2$, ενώ το τερματικό κόστος : $\hat{c}(x_N) = lx_N^2$. Επειδή δεν υπάρχει περιορισμός ως προς τις τιμές της x_t θεωρούμε ότι $A_t = \mathbb{R} = (-\infty,\infty)$.*
Εχουμε πρόβλημα πεπερασμένου ορίζοντα με πιθανοθεωρητική δυναμική, γιατί η επόμενη κατάσταση εξαρτάται απο την τιμή της τ.μ. z_t.
Σύμφωνα με τα παραπάνω γράφουμε την εξίσωση βελτιστοποίησης

$$v(t,x_t) = \min_{\alpha_t \in \mathbb{R}} \{bx_t^2 + c\alpha_t^2 + Ev(t+1,gx_t + b\alpha_t + z_t)\}\ ,\ t = 0,1,\ldots,N-1$$

$$v(N,x_N) = lx_N^2\ ,\quad \text{όπου } E(Z_z) = 0\ ,\ \sigma^2(Z_z) = \sigma^2(\alpha_t^2)\ ,\ \text{ και ζητάμε την } v(0,x_0)\ .$$

Παρακάτω δίνουμε τη λύση του προβλήματος.

$$v_N(x) = lx^2,$$

$$v_{N-1}(x) = \min_{\alpha \in \mathbb{R}} \{bx^2 + c\alpha^2 + Ev(N,gx + h\alpha + z)\} =$$

$$= bx^2 + \min_{\alpha \in \mathbb{R}} \{c\alpha^2 + E[l(gx + b\alpha + z)^2\} = bx^2 + \min_{\alpha \in \mathbb{R}} \{f_{N-1}(\alpha)\}\ ,$$

$$f_{N-1}(\alpha) = c\alpha^2 + lE\{gx + h\alpha + z\}^2 = c\alpha^2 + lE\{(gx + h\alpha)^2 + z^2 + 2z(gx + h\alpha)\} =$$

$$= c\alpha^2 + l\{Ez^2 + (gx + b\alpha)^2 + 2(gx + h(\alpha)Ez\}\ .$$

Ομως $\quad Ez = 0\ ,\ Ez^2 = \sigma^2(z) + (Ez)^2 = \sigma^2(\mathrm{z}) = \sigma^2\alpha^2 \quad \Rightarrow$

$$\Rightarrow f_{N-1}(\alpha) \ = \ c\alpha^2 + l[\sigma^2\alpha^2 + (gx + b\alpha)^2]\ .$$

$$\frac{df_{N-1}(\alpha)}{d\alpha} = 2\alpha c + l[2\sigma^2\alpha + 2h(gx + h\alpha)] \Rightarrow$$

$$\frac{d^2f_{N-1}(\alpha)}{d\alpha^2} = 2c + 2\sigma^2 l + 2h^2.$$

Θεωρώντας $b_1c \geq 0$: $f''_{N-1}(\alpha) \geq 0 \Rightarrow$

$\Rightarrow f_{N-1}(\alpha)$: *κυρτή, άρα ελαχιστοποιείται για:*

$$f'_{N-1}(\alpha) = 0 \Rightarrow 2(c + l\sigma^2 + h^2)\alpha + 2hgx = 0 \Rightarrow \alpha^*_{N-1}(x) = -\frac{hg}{c+l\sigma^2+h^2} = -d_{N-1}x$$

και $\quad v_{N-1}(x) = bx^2 + cd_{N-1}^2x^2 + l\sigma^2d_{N-1}^2x^2 + (g - hd_{N-1})^2x^2 = l_{N-1}x^2.$

Θέτουμε (ή καλύτερα υποθέτουμε):

$$v_\kappa(x) = l_\kappa x\ ,\ \ \alpha^*_\kappa(x) = -d_\kappa x\ ,\ \ \kappa = n+1,\ldots,N\ \ ,\ \ l_\kappa \geq 0\ ,\ \text{για κάποιο}\ n : n \leq N-1\ .$$

Τότε θα έχουμε:

$$v_n(x) = v(n,x) = \min_{\alpha \in \mathbb{R}}\{bx^2 + c\alpha^2 + E_{n-1}v(gx + h\alpha + z)\} =$$

$$= \underset{\alpha \in \mathbb{R}}{n}\{bx^2 + c\alpha^2 + E[l_{n+1}(gx + h\alpha + z)^2]\} =$$

$$= bx^2 + \min_{\alpha \in \mathbb{R}}\{f_{n-1}(\alpha)\},\ \text{όπου:}$$

$$f_{n+1}(\alpha) = c\alpha^2 = c\alpha^2 + l_{n+1}E[(gx + h\alpha)^2 + z^2 + 2z(gx + h\alpha)] = (\text{όπως \& προηγουμένως})$$

$$= c\alpha^2 + l_{n-1}(\sigma^2\alpha^2 + (gx + h\alpha)^2).$$

Οπως και παραπάνω (με $l_{n+1} = l$) βρίσκουμε:

$$\alpha^\kappa_n(x) = -\tfrac{hg}{c + l_{n+1}\sigma^2 + h^2}x = -d_n x$$

$$v(n,x) = [b + cd_n^2 + l_{n+1}\sigma^2 + (g - hd_n)^2]x^2 = e_n x^2,\ \ \ e_n \geq 0$$

Δείξαμε δηλ. επαγωγικά ότι αν για κάποιο n ισχύει :

$$v_{n+1}(x) = l_{n+1}x^2,\ \ l_{n+1} \geq 0$$

$$\alpha_{n+1}(x) = -d_{n+1}x\ \ \text{τότε και για}\ \ n\ \ \text{θα έχουμε :}$$

$$v_n(x) = l_n x^2,\ \ l_n \geq 0\ \ \ \alpha_n(x) = -d_n x\ ,$$

ενώ αυτό ισχύει και για $n = N$, όπως είδαμε παραπάνω. Επομένως ισχ`υει $\forall n = 0,1,\ldots,N$ και έτσι μπορούμε να φτιάξουμε το παρακάτω αναδρομικό σχήμα για τη λύση του προβλήματος.

$$v(n,x) = l_n x^2$$

$$\alpha^\kappa(n,x) = -d_n x,\ \ n = 0,1,\ldots,N$$

όπου:

$$l_N = l,\ \ d_N = 0$$
$$l_n = b + cd_n^2 + l_{n+1}\sigma^2 + (g - hd_n)^2,$$

$$d_n = \tfrac{hg}{c + l_{n+1}\sigma^2 + h^2}\ ,\ \ \ n = 0,1,\ldots,N-1\ .$$

ΑΣΚΗΣΕΙΣ ΚΕΦΑΛΑΙΟΥ 1

1. **α)** *Να λυθεί το πρόβλημα διαδρομής ελάχιστου μήκους του σχήματος 1. με τις μεθόδους διαδοχικών προσεγγίσεων,*

i) στο χώρο τιμών
ii) στο χώρο πολιτικών

β) *Να δοθεί η διατύπωση του προβλήματος μέσω γραμμικού προγραμματισμού.*

Σχήμα 1

2. *Δίνεται το δίκτυο :*

α) *Να βρεθεί η διαδρομή ελάχιστου μήκους απο το (1) στο (6)*
i) Με τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο τιμών.
ii) Με τη μέθοδο των διαδοχικών προσεγγίσεων στο χώρο πολιτικών.

β) *Να δοθεί η διατύπωση του προβλήματος σαν πρόβλημα γραμμικού προγραμματισμού.*

3. *Εστω ένα δίκτυο (σχήμα 3) και έστω $\delta_1 = (x_o,\ldots,y,\ldots z)$ μια βέλτιστη διαδρομή απο το x_o στο z (τερματικό).*
Εστω $\delta_2 = (x_o,\ldots,y)$ και $\delta_3 = (y,\ldots,z)$ δυο κομμάτια της διαδρομής.
α) *Είναι η δ_3 βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το y και τελικότο z ; δώστε απόδειξη).*

β) *Είναι η δ_2 βέλτιστη διαδρομή για το πρόβλημα με αρχικό σημείο το x_0 και τελικότο y ; δώστε απόδειξη).*

γ) *Αν η απάντηση στο (β) είναι καταφατική, υπάρχει εξίσωση βελτιστοποίησης που βασίζεται στην ιδιότητα (β) ; (αν ναι, να βρεθεί).*

Σχήμα 3

4. *Ας θεωρήσουμε το γενικό πρόβλημα της διαδρομής ελάχιστου μήκους σε ένα δίκτυο. Να βρεθεί μια ικανή συνθήκη κάτω απο την οποία η βέλτιστη πολιτική είναι η μυωπική.*

5. *Ενας φοιτητής έχει αγοράσει ένα "όχι και πολύ καινούργιο" αυτοκίνητο, και σκοπεύει να κάνει τις διακοπές του πηγαίνοντας από την πόλη 1 στην πόλη 8 της εικόνας 1 , σελ. 1 . Σκέφτεται όμως ότι αν ακολουθήσει μια διαδρομή ανάμεσα σε δύο πόλεις όπου η θερμοκρασία είναι πολύ υψηλή, η μηχανή του αυτοκινήτου έχει κίνδυνο να σκάσει. Υπολογίζει επίσης ότι c_{ij} είναι η μέγιστη θερμοκρασία που θα συναντήσει κατά τη διαδρομή από την πόλη i στην πόλη j , όπου c_{ij} είναι οι αριθμοί που στο παράδειγμα της εικόνας 1 χρησιμοποιήθηκαν σαν κόστη των αντίστοιχων* *δρόμων. Θέλει επομένως να βρει εκείνη τη διαδρομή που ελαχιστοποιεί τη μεγαλύτερη θερμοκρασία που θα συναντήσει.*

α) *Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού, και να εφαρμοσθεί για τα δεδομένα της εικόνας 1.*

β) *Να διατυπωθεί η αντίστοιχη εξίσωση βελτιστοποίησης στη γενική περίπτωση δικτύων που δεν περιέχουν κύκλους.*

6. *Θεωρούμε το πρόβλημα εύρεσης της διαδρομής ελάχιστου μήκους από ένα αρχικό κόμβο (πηγή) σ' ένα τελικό (προορισμό), σ' ένα δίκτυο που δεν περιέχει κύκλους. Χρησιμοποιήστε τη μέθοδο του Δυναμικού Προγραμματισμού για να βρείτε τη δεύτερη–καλύτερη (αμέσως μεγαλύτερη από την ελάχιστη) διαδρομή από την πηγή στον προορισμό. (Υπόδειξη: 1. Η δεύτερη–καλύτερη διαδρομή πρέπει να διαφέρει από την ελάχιστη τουλάχιστον σε ένα βέλος. 2. Εναλλακτικά μπορείτε να αποδείξετε ότι η δεύτερη–καλύτερη διαδρομή από τον κόμβο j στον προορισμό επιτυγχάνεται αν πάμε από το j σε κάποιο k μέσω του βέλους (j,k), και μετά στον προορισμό ακολουθώντας είτε την ελάχιστη, είτε τη δεύτερη–καλύτερη διαδρομή από τον k στον προορισμό).*

7. *Θεωρούμε ένα $m \times n$ πίνακα $A = \{a_{ij}\}$. Θέλουμε να ξεκινήσουμε από τη θέση $(1,1)$ και να φτάσουμε στη (m,n) προχωρώντας κάθε φορά ένα βήμα είτε προς τα δεξιά είτε προς τα κάτω, αλλά κατά τέτοιο τρόπο ώστε το άθροισμα των a_{ij} των στοιχείων που θα διασχίσουμε να είναι όσο το δυνατό μικρότερο. Δείξτε πώς μπορούμε να βρούμε τα καλύτερα μονοπάτια.*

8. *Εστω το πρόβλημα :*

$$(\min) \quad w = \sum_{t=0}^{3} [2x^2(t) + \alpha^2(t)]$$

$$\begin{aligned} \pi. \quad & x(t+1) = x(t) - \alpha(t) \\ & x(t) \in X = [0,1] \quad \forall t, \\ & \alpha(t) \in A(x(t)) = [0,1], \\ & x(0) = 1. \end{aligned}$$

Να βρεθεί η βέλτιστη πολιτική $\pi = (\alpha_o, \alpha_1, \alpha_2)$.

9. *Εστω το πρόβλημα :*

$$(\max) \quad \sum_{j=1}^{n} c_j x_j$$

$$\begin{aligned} \pi. \quad & \alpha_1 \le x_1 \le b_1 \\ & \alpha_j x_j \le x_{j+1} \le b_j x_j, \qquad j = 1,2,\ldots,n-1, \end{aligned}$$

όπου c_j, α_j, b_j είναι γνωστοί αριθμοί.
Να δοθεί η διατύπωση του προβλήματος μέσω Δυναμικού Προγραμματισμού.

10. *Εστω το πρόβλημα :*

$$\begin{aligned} \max \quad & 3x + 5y \\ \pi. \quad & 2x + y \le 10 \\ & x + 2y \le 6 \\ & x \in \{0,1,4,6\} \\ & y \in [0,\infty]. \end{aligned}$$

Να δοθεί η διατύπωση του προβλήματος με τη μέθοδο του Δυναμικού Προγραμματισμού.

11. *Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης :*

$$\min \quad \sum_{j=1}^{N} \frac{x_j^2}{c_j} \quad ,$$

$$\pi. \quad \sum_{j=1}^{N} x_j \ = \ A \ ,$$

$$x_j \ \geq \ 0 \ , \quad j = 1,\ldots,N \ ,$$

όπου $c_j > 0$. Χρησιμοποιώντας μεθόδους Δ.Π., δείξτε ότι αν ξέρουμε τις τιμές $x_1, x_2, \ldots, x_{t-1}$, η βέλτιστη τιμή για το x_t είναι :

$$x_t \ = \ c_t \cdot \frac{A - x_1 - \cdots - x_{t-1}}{c_t + \cdots + c_N} \quad .$$

12 . Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης :

$$\max \quad x_1^{c_1} \ \cdot \ x_2^{c_2} \cdots x_N^{c_N} \ ,$$

$$\pi. \quad \sum_{j=1}^{N} x_j \ = \ A \ ,$$

$$x_j \ \geq \ 0 \ , \quad j = 1,\ldots,N \ ,$$

όπου $c_j \ > \ 0$. Χρησιμοποιώντας μεθόδους Δ.Π., βρείτε τη μορφή της βέλτιστης λύσης $x_1^*, x_2^*, \ldots, x_N^*$. Δείξτε πώς εξαρτάται η βέλτιστη τιμή της αντικειμενικής συνάρτησης από το A .

13 . Θεωρούμε τους αριθμούς $\alpha_1, \alpha_2, \ldots, \alpha_n$, όπου $\alpha_\kappa \geq 0 \quad k = 1,2,\ldots,n$, και $\sum_{\kappa=1}^{n} \alpha_\kappa = A$. Να αποδείξετε με επιχειρήματα Δυναμικού Προγραμματισμού ότι :

$$[(\tfrac{1}{n} \sum_{\kappa=1}^{n} \alpha_\kappa)]^n \geq \prod_{\kappa=1}^{n} \alpha_\kappa \ .$$

14 . Εστω το πρόβλημα :

$$\max \quad \sum_{t=1}^{n} (x_t)^2$$

$$\pi. \quad \sum_{t=1}^{n} (x_t) = 1$$

$$x_t \geq 0$$

Να λυθεί με τη μέθοδο του Δυναμικού Προγραμματισμού.

15 . Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες $E_1, E_2, \ldots E_n$. Ο χρόνος εκτέλεσης της E_i είναι γνωστή σταθερά T_i, $i = 1,2,\ldots,n$. Για την εκτέλεση των εργασιών υπάρχει ένας σταθμός εξυπηρέτησης. Υποθέτουμε ότι η εκτέλεση των εργασιών αρχίζει τη χρονική στιγμή $t = 0$. Κριτήριο βελτιστοποίησης είναι ο συνολικός χρόνος ροής (flow–time), που ορίζεται σαν το άθροισμα $\sum_{i=1}^{n}$, όπου t_i είναι η χρονική στιγμή που τελείωσε η εκτέλεση της εργασίας E_i .

Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που ελαχιστοποιεί το συνολικό χρόνο ροής.

Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

Παράδειγμα. Αν $n = 3$, $T_1 = 2$, $T_2 = 4$, $T_3 = 5$, και η εξυπηρέτηση γίνει με τη σειρά 2,3,1, ο χρόνος ροής θα είναι $4 + (4 + 5) + (4 + 5 + 2) = 4 + 9 + 11 = 24$.

16. Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες $E_1, E_2, \ldots E_n$. Ο χρόνος εκτέλεσης της E_i είναι γνωστή σταθερά T_i, $i = 1,2,\ldots,n$. Μπορούμε να εκτελούμε το πολύ μια εργασία (υπάρχει ένας μόνο σταθμός εξυπηρέτησης) και αν η εργασία E_i τελειώσει τη χρονική στιγμή t έχουμε ένα κέρδος $R_i(t)$.
Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που μεγιστοποιεί το συνολικό κέρδος.
Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

17. Ας θεωρήσουμε ότι υπάρχουν προς εκτέλεση n εργασίες $E_1, E_2, \ldots E_n$. Ο χρόνος εκτέλεσης της E_i είναι γνωστή σταθερά T_i, $i = 1,2,\ldots,n$. Για την εκτέλεση των εργασιών υπάρχουν k όμοιοι σταθμοί εξυπηρέτησης. Υποθέτουμε ότι η εκτέλεση των εργασιών αρχίζει τη χρονική στιγμή $t = 0$. Κριτήριο βελτιστοποίησης είναι ο συνολικός χρόνος που απαιτείται μέχρις ότου να τελειώσουν όλες οι εργασίες (makespan).

Το πρόβλημα είναι να βρεθεί η σειρά εκτέλεσης των εργασιών που ελαχιστοποιεί τον παραπάνω συνολικό χρόνο.

Να διατυπωθεί το πρόβλημα σαν πρόβλημα Δυναμικού Προγραμματισμού.

18. Ενα αεροπλάνο πρέπει να μεταφέρει υλικά σε μια επιστημονική μονάδα στο Νότιο Πόλο. Υπάρχουν τρεις τύποι υλικών, που για ευκολία τους ονομάζουμε 1, 2 και 3. Επειδή κάθε είδος υλικού έχει διαφορετική αξία ως προς το βαθμό αναγκαιότητάς του, μοντελοποιούμε τις διαφορές αυτές υποθέτοντας ότι βάρος α από το υλικό τύπου i έχει αξία ίση με $R_i(\alpha)$, όπου :

$$R_1(\alpha) = \alpha^2, \quad R_2(\alpha) = \alpha^2 + 2\alpha, \quad R_3(\alpha) = 2\alpha.$$

Το μεγαλύτερο βάρος που μπορεί να μεταφέρει το αεροπλάνο είναι 10 τόνοι.

Να βρεθούν οι ποσότητες από κάθε τύπο υλικού που πρέπει να φορτωθούν στο αεροπλάνο, έτσι ώστε να μεγιστοποιηθεί η συνολική αξία των μεταφερόμενων στη βάση υλικών, χρησιμοποιώντας τη μέθοδο του Δυναμικού Προγραμματισμού.

19. Θεωρούμε ένα δίκτυο επικοινωνίας N κόμβων, όπου οι κόμβοι συμβολίζουν πόλεις και τα βέλη τηλεφωνικές γραμμές μεταξύ των πόλεων. Κατά τα γνωστά συμβολίζουμε με $A(x)$ το σύνολο των βελών που έχουν σαν αρχικό κόμβο τον x, $x = 1,2,\ldots,N-1$. Υποθέτουμε ότι θέλουμε ν στείλουμε ένα μήνυμα από τον κόμβο 1 στον κόμβο N. Για κάθε βέλος $(x,\alpha) \in A(x)$, η πιθανότητα ότι το μήνυμα, περνώντας από τον x, θα φτάσει με επιτυχία στον $\overline{\alpha}(x)$, είναι γνωστή και ίση με $p(x,\alpha)$.

Θέλουμε να βρούμε το μονοπάτι από τον 1 στον N, που έχει τη μεγαλύτερη πιθανότητα επιτυχούς μετάδοσης του μηνύματος.

Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

20. Σε μια εθνική επιτροπή συμμετέχουν R αντιπρόσωποι. Ολη η χώρα διαιρείται σε s περιοχές, με $s < R$. Κάθε περιοχή j έχει πληθυσμό p_j. Για να είναι αυστηρά αναλογική η αντιπροσώπευση κάθε περιοχής στην επιτροπή, θα έπρεπε η περιοχή j να πάρει $r_j = R \cdot p_j / \sum_{i=1}^{s} p_i$ αντιπροσώπους. Αυτή όμως η κατανομή δεν είναι δυνατή, γιατί γενικά τα r_j δεν έχουν ακέραιες τιμές. Το πρόβλημα επομένως είναι να βρεθεί μια κατανομή αντιπροσώπων y_j, $j = 1,\ldots,s$ στις περιοχές, κατά τρόπο ώστε να ελαχιστοποιηθεί η μέγιστη απόλυτη διαφορά μεταξύ του y_j και της "ιδανικής" τιμής r_j, που παρατηρείται στις περιοχές, δηλαδή να ελαχιστοποιηθεί η παράσταση $\max(|y_1 - r_1|, |y_2 - r_2|, \ldots, |y_s - r_s|)$.

α) Να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

β) Να γίνει αριθμητική εφαρμογή του α), όταν $R = 4$, $s = 3$, $r_1 = 0.6$, $r_2 = 2.2$, $r_3 = 1.2$. Να συζητηθεί ποιοτικά η λύση που προκύπτει.

21 . Θεωρούμε το γνωστό πρόβλημα μεταφοράς. Υπάρχουν m σταθμοί αποστολής, καθένας από τους οποίους έχει διαθέσιμη ποσότητα S_i, και n σταθμοί ζήτησης, καθένας από τους οποίους ζητά ποσότητα D_j. Ισχύει η γνωστή παραδοχή ότι $\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} D_j$, όπως επίσης ότι τα S_i, D_j, και οι ποσότητες x_{ij} που αποστέλλονται από το σταθμό αποστολής i στο σταθμό ζήτησης j, πρέπει να είναι ακέραιοι αριθμοί. Το κόστος μεταφοράς x_{ij} μονάδων είναι γενικά μη γραμμική συνάρτηση $c_{ij}(x_{ij})$, $i = 1,\ldots,m$, $j = 1,\ldots,n$.

α) Μοντελοποιήστε το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού, για $m = 2$. (Υπόδειξη : Μπορείτε να χρησιμοποιήσετε την ισότητα συνολικής αποστολής και συνολικής ζήτησης, για να πάρετε μονοδιάστατη μεταβλητή κατάστασης).

β) Δείξτε πώς μεταβάλλεται η μοντελοποίηση στο α) όταν $m = 3$.

22 . Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης .

$$\min \quad f = \sum_{i=1}^{N} x_i \ ,$$

$$\begin{aligned} \pi. \quad & x_i + x_{i+1} \ \geq \ d_i \ , \ i = 1,2,\ldots,N-1 \ , \\ & x_i \ \geq \ 0 \ . \end{aligned}$$

Δείξτε ότι η ελάχιστη τιμή της f είναι ίση με $f^* = \max_{i=1,\ldots,N} d_i$, αρκεί ένα από τα d_i να είναι θετικό.

23 . Επαναλάβετε την άσκηση 7 όταν οι πρώτοι $N-1$ περιορισμοί αντικατασταθούν από τους $x_i + x_{i+1} + x_{i+2} \ \geq \ d_i$, $i = 1,2\ldots,N-2$.

24 . Χρησιμοποιώντας τη μέθοδο του Δυναμικού Προγραμματισμού και με την εισαγωγή κατάλληλων βοηθητικών μεταβλητών, μοντελοποιήστε το παρακάτω πρόβλημα ελαχιστοποίησης :

$$\min_{x_i \geq 0} \left[\frac{x_1}{x_2+x_3} + \frac{x_2}{x_3+x_4} + \cdots + \frac{x_{n-1}}{x_n+x_1} + \frac{x_n}{x_1+x_2} \right] .$$

25 . Θεωρούμε το πρόβλημα ελαχιστοποίησης :

$$\min \quad \sum_{k=1}^{N} g_k(x_k,x_{k+1}) + \sum_{k=1}^{N} h_k(x_k) \ ,$$

$$\begin{aligned} \pi. \quad & x_{n+1} \ = \ x_1 \ , \\ & 0 \leq x_k \leq b_k \ , \ k = 1,\ldots,N \ , \\ & \sum_{k=1}^{N} \phi_k(x_k) \ \geq \ c \ , \end{aligned}$$

όπου κάθε $\phi_k(x)$ είναι μια γνωστή γνήσια αύξουσα συνάρτηση του x, με $\phi_k(0) = 0$.
Εισάγουμε το παρακάτω βοηθητικό πρόβλημα :

$$\min \quad g(u,x_2) + g(x_2,x_3) + \cdots + g(x_{N-1},x_N) + g(x_N,v) + \sum_{k=2}^{N} h_k(x_k) \ ,$$

$$\begin{aligned} \text{υπό περιορ.} \quad & 0 \leq x_k \leq b_k \ , \ k = 2,\ldots,N \ , \\ & \sum_{k=2}^{N} \phi_k(x_k) \ \geq \ c \ . \end{aligned}$$

Δείξτε ότι αν το ελάχιστο του παραπάνω προβλήματος ως προς $x_2, \ldots, x_N$ είναι ίσο με $F(u,v,c)$, τότε το ελάχιστο του αρχικού προβλήματος είναι ίσο με

$$\min_{0 \le x_1 \le b_1} F(x_1, x_1, c - \phi_1(x_1)) .$$

26. Για την εύρεση του ελάχιστου στην άσκηση 10, θεωρείστε την παρακάτω ακολουθία συναρτήσεων για $t = 2,3,\ldots,N-1$:

$$F_t(u,v,c) = \min_{x_k,\, k=t,\ldots N} \left[g(u,x_t) + g(x_t,x_{t+1}) + \cdots + g(x_{N-1},x_N) + g(x_N,v) + \sum_{k=t}^{N} h_k(x_k) \right]$$

$$F_N(u,v,c) = \min_{x_N} \left[g(u,x_N) + g(x_N,v) + h_N(x_N) \right] ,$$

όπου τα $x_2,\ldots,x_N$ υπόκεινται στους περιορισμούς που αναφέρθηκαν στην άσκηση 25, ενώ για τις επιτρεπτές τιμές του c ισχύει η υπόθεση ότι $\sum_{k=t}^{N} \phi_k(b_k) \ge c$, όπου b_k είναι γνωστοί θετικοί αριθμοί.

Δείξτε ότι η ισοδύναμη με το παραπάνω πρόβλημα εξίσωση βελτιστοποίησης είναι

$$F_t(u,v,c) = \min_{x_t \in D_t} \left[g(u,x_t) + h_t(x_t) + F_{t+1}(x_t,v,c - \phi_t(x_t)) \right] ,$$

όπου ο χώρος αποφάσεων D_t για το x_t καθορίζεται από τις παρακάτω σχέσεις :

$$0 \le x_t \le b_t$$
$$\sum_{k=t+1}^{N} \phi_k(b_k) \ge c - \phi_t(x_t) .$$

27. Μοντελοποιήστε τα παρακάτω προβλήματα με τη μέθοδο του Δυναμικού Προγραμματισμού, όταν το κριτήριο βελτιστοποίησης είναι i. ελαχιστοποίηση και ii. μεγιστοποίηση. Με z συμβολίζουμε την αντικειμενική συνάρτηση.

α) $z = cx_1^2 + (x_1 + cx_2)^2 + (x_1 + x_2 + cx_3)^2 + \cdots + (x_1 + x_2 + \cdots + x_{N-1} + cx_N)^2$,
υ.π. $x_1^2 + x_2^2 + \cdots + x_N^2 = 1$.

β) $z = cx_1^2 + (x_1 + cx_2)^2 + (x_1 + cx_2 + c^2x_3)^2 + \cdots + (x_1 + cx_2 + \cdots + c^{N-2}x_{N-1} + c^{N-1}x_N)^2$,
υ.π. $x_1^2 + x_2^2 + \cdots + x_N^2 = 1$.

γ) $z = cx_1^2 + (x_1 + cx_2)^2 + [x_1 + cx_2 + (c+d)x_3]^2 + \cdots$
$+ [x_1 + cx_2 + (c+d)x_3 + (c+2d)x_4 + \cdots + (c+(N-2)d)x_N]^2$,
υ.π. $x_1^2 + x_2^2 + \cdots + x_N^2 = 1$.

28. Να λυθεί με τη μέθοδο του Δυναμικού Προγραμματισμού το πρόβλημα :

$$\max \quad \sum_{i=1}^{N} c_i x_i ,$$

$$\pi. \quad \sum_{i=1}^{N} x_i^2 = 1 ,$$
$$0 \le x_1 \le x_2 \le \ldots \le x_N ,$$

όπου c_i , $i = 1,\ldots,N$ είναι γνωστοί μη αρνητικοί συντελεστές.

29. Σε προβλήματα εύρεσης ελάχιστων και μέγιστων ιδιοτιμών πινάκων (οπως π.χ. στον πίνακα Jacobi) προκύπτει συχνά η παρακάτω τετραγωνική μορφή, που πρέπει να ελαχιστοποιηθεί ή να μεγιστοποιηθεί ως προς x_i , $i = 1,\ldots,N$.

$$f_N(\underline{x}) = \sum_{i=1}^{N} c_i x_i^2 + 2\sum_{i=1}^{N-1} b_i x_i x_{i+1} ,$$

όπου τα x_i είναι περιορισμένα στη N–διάστατη μοναδιαία σφαίρα S, δηλαδή $\underline{x} \in S \iff \sum_{i=1}^{N} x_i^2 = 1$.

Θεωρούμε τις παρακάτω ακολουθίες :

$$v_N(c) = \max_{\underline{x} \in S} [f_N(\underline{x}) + 2cx_N],$$

$$w_N(c) = \min_{\underline{x} \in S} [f_N(\underline{x}) + 2cx_N].$$

Χρησιμοποιώντας μεθόδους του Δυναμικού Προγραμματισμού βρείτε αναδρομικές σχέσεις που συνδέουν τα $v_N(c)$, $w_N(c)$ με τα $v_{N-1}(c)$, $w_{N-1}(c)$ αντίστοιχα.

30. Λύστε την άκηση 29 όταν

$$f_N(\underline{x}) = \sum_{i=1}^{N} c_i x_i^2 + 2\sum_{i=1}^{N-1} b_i x_i x_{i+1} + 2\sum_{i=1}^{N-2} d_i x_i x_{i+2}.$$

31. Θεωρούμε το παρακάτω πρόβλημα βελτιστοποίησης :

$$\min \quad F_N(x_1,x_2,\ldots,x_N) = \phi_1(x_1) + \phi_2(x_2) + \cdots + \phi_N(x_N)$$

$$\begin{aligned} \pi. \quad & x_i \geq 0 \quad i = 1,\ldots,N, \\ & x_1 \geq r_1, \\ & x_1 + x_2 \geq r_2, \\ & \vdots \\ & x_1 + x_2 + \cdots x_N \geq r_N. \end{aligned}$$

Ορίζουμε την ακολουθία

$$f_k(z) = \min_{\underline{x}} \sum_{i=k}^{N} \phi_i(x_i), \text{ με νέους περιορισμούς :}$$

$$\begin{aligned} & x_i \geq 0 \quad i = 1,\ldots,N, \\ & x_k \geq r_k - z, \\ & x_k + x_{k+1} \geq r_{k+1} - z, \\ & \vdots \\ & x_k + \cdots + x_N \geq r_N - z, \end{aligned}$$

όπου $z \geq 0$, $k = 1,2,\ldots,N$. Δείξτε ότι :

$$f_k(z) = \min_{x_k \geq 0,\, x_k \geq r_k - z} [\phi_k(x_k) + f_{k+1}(z + r_k)], \; k = 1,2,\ldots,N,$$

και επομένως ότι $\min_{\underline{x}} F_N(x_1,\ldots,x_N) = f_1(0)$.

32. Χρησιμοποιώντας επιχειρήματα Δυναμικού Προγραμματισμού βρείτε μια μέθοδο για να προσεγγίσετε τη συνάρτηση $f(x)$ στο διάστημα $[a,b]$ μέσω της γραμμικής συνάρτησης $rx + s$, σύμφωνα με τα εξής μέτρα απόκλισης :

α) $\int_a^b (f(x) - rx - s)^2\, dx$

β) $\max_{a \leq x \leq b} |f(x)\text{-}rx\text{-}s|$.

<u>***Υπόδειξη***</u>. Βρείτε μια αναδρομική διαδικασία για να προσδιορίσετε τα r,s που ελαχιστοποιούν τα παραπάνω μέτρα απόκλισης.

33. *Γενικεύστε το αποτέλεσμα της προηγούμενης άσκησης, για συνάρτηση προσέγγισης όχι γραμμική, αλλά πολυωνυμική βαθμού* n.

ΑΣΚΗΣΕΙΣ ΚΕΦΑΛΑΙΟΥ 2

1. *Δίνεται το δίκτυο του σχήματος 1 , όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι το συνολικό εκπτωτικό κόστος*

$\sum_{t=0} \beta^t c(x_t,\alpha_t)$, *με* $\beta = 0.9$, *όταν* $x_o = 1$.

Σχήμα 1

2. *Δίνεται το δίκτυο του σχήματος 2 , όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι*

α) *Το συνολικό εκπτωτικό κόστος*
$\sum_{t=0} \beta^t c(x_t,\alpha_t)$, *με* $\beta = 0.5$,

β) *Το μέσο κόστος*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{t=0}^{n} c(x_t,\alpha_t) ,$$

όταν $x_o = 1$.

Σχήμα 2

3. *Δίνεται το δίκτυο του σχήματος 3 , όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι το μέσο κόστος*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{t=0}^{n} c(x_t,\alpha_t) ,$$

όταν $x_o = 1$.

Σχήμα 3

4. *Δίνεται το δίκτυο του σχήματος 4 , όπου οι αριθμοί δίπλα στα βέλη συμβολίζουν μήκη διαδρομών. Να βρεθεί η βέλτιστη πολιτική για το πρόβλημα άπειρου ορίζοντα, όπου το κριτήριο βελτιστοποίησης είναι το συνολικό εκπτωτικό κόστος*
$\sum_{t=0} \beta^t c(x_t,\alpha_t)$, *με* $\beta = 0.9$, *όταν* $x_o = 1$.

5. *Θεωρείστε τον παρακάτω σύνθετο αλγόριθμο διαδοχικών προσεγγίσεων για τη λύση των εξισώσεων βελτιστοποίησης στο πρόβλημα άπειρου ορίζοντα με πεπερασμένους χώρους καταστάσεων και αποφάσεων.*

Εφαρμόζουμε τη διαδικασία διαδοχικών προσεγγίσεων στο χώρο τιμών για N *συνεχόμενα βήματα. Μετά χρησιμοποιούμε τις τιμές των* $\alpha^{(N)}(x)$ *που έχουμε βρει και μ' αυτές εφαρμόζουμε μια φορά το τμήμα υπολογισμού των* w *της διαδικασίας διαδοχικών προσεγγίσεων στο χώρο των πολιτικών. Σταματάμε αν ικανοποιούνται οι συνθήκες αυτής της μεθόδου για τα* Δw. *Διαφορετικά επαναλαμβάνουμε όλη την παραπάνω μέθοδο, με αρχική τιμή για τα* $v^{(N+1)}$ *τις τιμές των* w *που έχουμε υπολογίσει από τις διαδοχικές προσεγγίσειsν στο χώρο των πολιτικών. Εξηγήστε γιατί η μέθοδος που περιγράψαμε συγκλίνει στη λύση των εξισώσεων βελτιστοποίησης σε πεπερασμένο αριθμό βημάτων.*

6. *Θεωρούμε την παρακάτω εξίσωση βελτιστοποίησης*

$$v(x) = \max_{0 \le \alpha \le x} \left[g(\alpha) + h(x-\alpha) + v(p\alpha + q(x-\alpha)) \right],$$

όπου κάνουμε τις εξής παραδοχές για την ασυμπτωτική συμπεριφορά των g *και* h

α) $g(\alpha) \sim c_1\alpha^d$, $h(\alpha) \sim c_2\alpha^d$, $c_1, c_2, d > 0$, *όταν* $\alpha \to \infty$.

β) $g(\alpha) \sim c_1\alpha^{d_1}$, $h(\alpha) \sim c_2\alpha^{d_2}$, $c_1, c_2, d_1, d_2 > 0$, *όταν* $\alpha \to \infty$.

Και στις δύο περιπτώσεις βρείτε την ασυμπτωτική συμπεριφορά της $v(x)$, *όταν* $x \to \infty$.

ΑΣΚΗΣΕΙΣ ΚΕΦΑΛΑΙΟΥ 3

1. *Ενα πολιτικό κόμμα σχεδιάζει οικονομική εξόρμηση και σκοπεύει να διαθέσει χρηματικό ποσό S (το πολύ) για διαφημιστική εκστρατεία διάρκειας N εβδομάδων. Από προηγούμενες στατιστικές εκτιμήσεις είναι γνωστό ότι αν στην αρχή της k εβδομάδας έχει ήδη συγκεντρωθεί ποσό x, τότε η πιθανότητα μέσα στη διάρκεια της εβδομάδας να συγκεντρωθεί επιπλέον ποσό x_k, αν διατεθεί για διαφήμιση ποσό s_k, είναι ίση με $p_k(x_k/x,s_k)$, 'οπου p_k, $k = 1,\ldots,N$, γνωστές συναρτήσεις. Το πρόβλημα είναι να βρεθεί η διαφημιστική στρατηγική (όσον αφορά τη διάθεση χρημάτωνσε διαφημίσεις) που μεγιστοποιεί το αναμενόμενο συνολικό ποσό που θα συγκεντρωθεί (συμπεριλαμβανομένων και των τυχόν αδιάθετων ποσών από το διαφημιστικό κεφάλαιο S).*

α) *Δείξτε ότι η βέλτιστη στρατηγική μπορεί να βρεθεί χρησιμοποιώντας μεθόδους Δυναμικού Προγραμματισμού.*

β) *Δείξτε πώς πρέπει να αλλάξει η μοντελοποίηση στο α), όταν υπάρχει κεφάλαιο S στην αρχή της πρώτης εβδομάδας διαθέσιμο για διαφημίσεις, αλλά ακόμη το κόμμα είναι διατεθειμένο να διαθέσει όσα επιπλέον χρήματα μαζέψει κατα τη διάρκεια της διαφημιστικής εκστρατείας για περαιτέρω διαφήμιση μέχρι το τέλος των N εβδομάδων.*

2. *Ο διευθυντής ενός σούπερ μάρκετ αντιμετωπίζει το εξής πρόβλημα σε σχέση με την εξυπηρέτηση των πελατών. Υπάρχουν στατιστικά στοιχεία σχετικά με τον αριθμό των πελατών που μπαίνουν στο κατάστημα κάθε χρονική περίοδο (ας πούμε 15 λεπτών), όπως επίσης και για τον αριθμό των πελατών που εξυπηρετούνται μέσα σε μια περίοδο. Ο τελευταίος αριθμός εξαρτάται τόσο από τον αριθμό πελατών που βρίσκονται στο κατάστημα, όσο και από τον αριθμό των ανοιχτών ταμείων. Ετσι έχει εκτιμήσει ότι $p_t(e)$ είναι η πιθανότητα e πελάτες να μπουν στο κατάστημα την περίοδο t (όπου για λόγους απλότητας υποθέτουμε ότι μπαίνουν όλοι μαζί στην αρχή της περιόδου). Επίσης $q_t(m/x,s)$ είναι η πιθανότητα να εξυπηρετηθούν m πελάτες στη διάρκεια της περιόδου t, δεδομένου ότι στην αρχή της περιόδου υπάρχουν x πελάτες στο κατάστημα και s ανοιχτά ταμεία. Εστω ότι w είναι το κόστος ανά περίοδο για κάθε ανοιχτό ταμείο. Επίσης ο διευθυντής έχει μια σχετική εκτίμηση για το κόστος που προκύπτει από την αναμονή των πελατών και αυτό είναι h_t για κάθε πελάτη που μένει στο κατάστημα στο τέλος της περιόδου t.*

α) *Να βρεθεί ένα μοντέλο Δυναμικού Προγραμματισμού για τον υπολογισμό της βέλτιστης πολιτικής κατά τις ώρες αιχμής, που υπολογίζονται σε N περιόδους, ούτως ώστε να ελαχιστοποιείται το αναμενόμενο συνολικό κόστος.*

β) *Δείξτε πώς αλλάζει η μοντελοποίηση στο α) όταν υπάρχει πρόσθετο κόστος K για κάθε ταμείο που ανοίγει στην αρχή μιας περιόδου (δηλαδή έχουμε επιπλέον κόστος rK στη διάρκεια μιας περιόδου, αν r ταμεία ανοίξουν στην αρχή αυτής της περιόδου.*

3. *Μια μεγάλη εταιρεία συνεργάζεται με μια τράπεζα για ευκολότερη διεκπεραίωση των οικονομικών της συναλλαγών. Οι οικονομικοί όροι της συνεργασίας αυτής είναι οι εξής.*

Οι συναλλαγές της εταιρείας με την τράπεζα (καταθέσεις και αναλήψεις) μπορούν να γίνονται στην αρχή κάθε μέρας, και για κάθε κατάθεση ή ανάληψη η τράπεζα χρεώνει ποσό ίσο με D ή W αντίστοιχα. Κατά τη διάρκεια της μέρας το υπόλοιπο του λογαριασμού της εταιρείας έχει διακυμάνσεις που οφείλονται είτε σε καταθέσεις τρίτων στο όνομα της εταιρείας, είτε σε αναλήψεις από τρίτους επιταγών που η εταιρεία έχει εκδώσει. Αν το υπόλοιπο πέσει κάτω από μηδέν, η τράπεζα δανείζει αυτόματα το λογαριασμό με τα απαιτούμενα ποσά, με επιτόκιο r ανά ημέρα.

Η εκτίμηση της εταιρείας για το "κόστος ευκαιρίας" είναι s ανά ημέρα (δηλαδή θα μπορούσε να κερδίσει ποσό s κάθε μέρα από κάθε χρηματική μονάδα που θα επένδυε αλλού αντί να την καταθέσει στην τράπεζα). Υποθέτουμε ότι τα r και s υπολογίζονται πάνω στο υπόλοιπο του λογαριασμού, όπως αυτό έχει διαμορφωθεί στο τέλος της μέρας.

Εστω q το ποσό κατά το οποίο μεταβάλλεται το υπόλοιπο του λογαριασμού στη διάρκεια μιας μέρας (το q μπορεί να είναι είτε θετικό είτε αρνητικό) εξ αιτίας των συναλλαγών της εταιρείας με τρίτους. Το q είναι τυχαία μετβλητή που ακολουθεί γνωστή συνάρτηση κατανομής $p_t(q)$, κατά την t ημέρα.

α) *Ζητείται να βρεθεί η πολιτική καταθέσεων/αναλήψεων που ελαχιστοποιεί το αναμενόμενο συνολικό κόστος της εταιρείας, σε μια περίοδο N ημερών.*

β) *Δείξτε πώς μεταβάλλεται το μοντέλο στο α) όταν τα κόστη συναλλαγών W και D, όπως επίσης και οι τόκοι από τυχόν δανεισμό αφαιρούνται από το λογαριασμό αντί να πληρώνονται απευθείας.*

4. *Ο διευθυντής προσωπικού μιας εταιρείας θέλει να προσλάβει ένα υπάλληλο καί παίρνει συνεντεύξεις από τους υποψήφιους. Απο πρίν έχει διακρίνει τρείς κατηγορίες υποψηφίων, μέτριους (1), καλούς (2) και πολυ καλούς (3). Επίσης έχει προηγούμενες εκτιμήσεις για τα ποσοστά των υποψηφίων που ανήκουν σε κάθε κατηγορία, και επομένως άν πάρει μια συνέντευξη η πιθανότητα ο υποψήφιος να ανήκει στην κατηγορία i είναι ίση με p_i, $i = 1,2,3$. Ο διευθηντής δεν μπορεί να δεί πάνω απ᾽ο N υποψήφιους. Αν κρίνει κάποιον πολύ καλό φυσικά θα τον προσλάβει, ενώ αν τον κρίνει μέτριο δεν χάνει τίποτα να εξετάσει και τον επόμενο. Το πρόβλημα βρίσκεται στο τι πρέπει να κάνει αν κρίνει κάποιον καλό. Αν τον διώξει, μπορεί όλοι οι επόμενοι να είναι μέτριοι, ενώ αν τον προσλάβει χάνει την ευκαιρία να προσλάβει ένα πολύ καλό μετά. Εδώ υποθέτουμε οτι αν δέν προσλάβει ενα υποψήφιο μόλις πάρει την συνέντευξη, δεν εχει ευκαιρία να τον προσλάβει αργότερα. Αν q_i είναι η "άξία" που εχει για το διευθυντή η πρόσληψη υπαλλήλου της κατηγορίας i, ᾽οπου προφανώς $q_1 < q_2 < q_3$, να βρεθεί η πολιτική πρόσληψης που μεγιστοποιεί το αναμενόμενο κέρδος απ᾽ ᾽ολη τη διαδικασία. Ζητείται να μοντελοποιηθεί το πρόβλημα με τη μέθοδο του Δ.Π. για διάφορες δυνατές περιπτώσεις :*

α) *Υπάρχει κόστος c για κάθε συνέντευξη.*

β) *Ο αριθμός των υποψηφίων που μπορεί να δεί ο διευθυντής είναι απεριόριστος $(N = \infty)$, και $c > 0$. Ποιά είναι η βέλτιστη πολιτική σ'αυτή την περίπτωση ᾽οταν $c = 0$;*

γ) *$N = \infty$, $c > 0$, και υπάρχει και εκπτωτικός παράγοντας β, $0 \le \beta \le 1$.*

δ) *Γράψτε το πρωτεύον και το δυαδικό γραμμικό πρόγραμμα που αντιστοιχεί στο γ).*

ε) *Υπολογίστε τη βέλτιστη πολιτική στο γ) όταν $q_1 = 1$, $q_2 = 2$, $q_3 = 3$, $p_1 = 0.3$, $p_2 = 0.5$, $p_3 = 0.2$, $c = 0.15$, $\beta = 1$.*

στ) *Αν ο διευθυντής έχει την ευκαιρία να αναβάλει την απόφαση σχετικά μ' ᾽εναν υποψήφιο, μέχρι να δεί και άλλους, σε ποι᾽α περίπτωση θα ήθελε να εκμεταλλευτεί αυτή τη δυνατότητα; Βρείτε π᾽ως αλλάζει η μοντελοποίηση στο γ) με την επιπλέον υπόθεση. Επίσης απαντήστε ξανά στα δ) και ε).*

(Γενική Υπόδειξη. Σαν κατάσταση x_t στο μοντέλο δυναμικού προγραμματισμού μπορείτε να χρησιμοποι᾽ησετε την καλύτερη κατηγορία υποψήφιου που ᾽εχει δεί ο διευθυντής στις $t-1$ προηγούμενες συνεντεύξεις).

5. *Κάθε μήνα η τράπεζα Ελλάδος ορίζει το ποσό πού πρέπει να δανεισθεί απ᾽ο τις διεθνείς χρηματαγορές, για να ικανοποιήσει τις ζητήσεις για δάνεια των πελατών της. Οι συνολικές απαιτήσεις τον μήνα t ε᾽ιναι d_t. Η πραγματική τιμή της d_t είναι τυχαία μεταβλητή που ακολουθεί κατανομή $p(d_t)$, υποθέτουμε όμως οτι η τιμή d_t είναι γνωστή απ᾽ο την αρχή της περιόδου t, πρίν παρθούν οι αντίστοιχες αποφάσεις. Κάθε μήνα η τράπεζα μπορεί να δανεισθεί κεφάλαια για 1, 2 ᾽η 3 μήνες, σε ποσότητες α_1, α_2, α_3. Απ' ᾽οσα είπαμε παραπάνω, η ποσότητα $\alpha_1 + \alpha_2 + \alpha_3$ μπορε᾽ι να χρησιμοποιειθεί για να καλύψει την d_t, η $\alpha_2 + \alpha_3$ για την d_{t+1}, και η α_3 για την d_{t+2}. Υποθέτουμε ακόμη οτι το κόστος δανεισμού ποσού α_i, $i = 1,2,3$ την περίοδο t, εξαρτάται απ᾽ο μια παράμετρο r, και δίνεται απο τη συνάρτηση $c_{ti}(\alpha_i/r)$, ᾽οπου η r ακολουθεί την πιθανοθεωρητική κατανομή $q_t(r)$ και η τιμή της είναι γνωστή στήν αρχή της περιόδου t πρίν ληφθούν οι αποφάσεις. Υποθέτουμε τέλος οτι όλες οι τυχα᾽ιες μεταβλητ᾽ες ε᾽ιναι ανεξάρτητες μεταξύ τους.*

α) *Να διατυπωθεί ᾽ενα μοντέλο Δυναμικο᾽υ Προγραμματισμο᾽υ για την ελαχιστοποίηση του αναμενόμενου συνολικού κόστους, υπ᾽ο τον περιορισμό οτι όλες οι* απαιτήσεις για δάνεια πρέπει να ικανοποιηθούν. Ο εκπτωτικ`os παράγοντας είναι ίσος με β , $0 < \beta < 1$.

β) Βρείτε πως απλοποιείται (αν απλοποιείται) το α) όταν η τιμή της d_t γ`ινεται γνωστή μετά την απόφαση της περιόδου $t$ . Υποθέστε `οτι όταν η d_t ξεπερνάει τα διαθέσιμα κεφάλαια προς δανεισμό, υπάρχει μοναδιαίο κόστος απώλειας ίσο με c .

6. Θεωρείστε τη γενική εξ`ισωση βελτιστοποίησης άπειρου ορίζοντα με πιθανοθεωρητική δυναμική για το αναμεν`ομενο εκπτωτικ`ο κ`οστος

$$v(x) = \min_{x \in D(x)} \{ c(x,\alpha) + \beta \sum_{x' \in X} p[X_{t+1} = x'/X_t = x, A_t = \alpha] \, \upsilon(x') \} .$$

Υποθέστε οτι έχετε εκτιμήσεις για τα άνω και κάτω φράγματα U_x , L_x των `αγνωστων τιμών $v(x)$: $L_x < v(x) < U_x$. Δείξτε ότι δεν είναι ποτέ βέλτιστο να χρησιμοποιήσουμε την απόφαση $\alpha' \in D(x)$ αν

$$c(x,\alpha') + \beta \sum_{x' \in X} p(x'/x, \alpha') \cdot L_{x'} > \min_{x \in (D(x)} + \{ c(x,\alpha) + \sum_{x' \in X} p(x'/x,\alpha) \cdot U_{x'} \} .$$

Προτείνετε πώς μπορεί να χρησιμοποιηθεί η παραπάνω ιδιότητα για να επιταχυνθεί η σύγκλιση στους αλγόριθμους των διαδοχικών προσεγγίσεων.

7. Σε κάθε στάδιο μιάς εξακολουθητικής διαδικασίας, μπορούμε να πάρουμε μία από δύο διαθέσιμες αποφάσεις. Αν πάρουμε την πρώτη απόφαση μπορούμε να κερδίσουμε μία μονάδα με πιθανότητα p_1 , δύο με πιθανότητα p_2 , ενώ με πιθανότητα p_3 δεν κερδίζουμε τίποτα και όλη η διαδικασία σταματά. Η δεύτερη απόφαση έχει αντίστοιχο σύνολο πιθανοτήτων p_1' , p_2' , p_3' .

Θέλουμε να βρούμε την ακολουθία αποφάσεων που μεγιστοποιεί την πιθανότητα να κερδίσουμε n μονάδες πρίν σταματήσει η διαδικασία. Να μοντελοποιηθεί το πρόβλημα χρησιμοποι`ωντας τη μέθοδο του Δυναμικο`υ Προγραμματισμο`υ.
(Υπόδειξη. Θεωρείστε το τρέχον κέρδος ίσο με 0 και το τερματικό κέρδος ίσο με 1 ή 0 , ανάλογα με το αν έχουν συγκεντρωθεί ή οχι n μονάδες στο τέλος του παιχνιδιού).

8. Ενας άνθρωπος στέκεται στην ουρά περιμένοντας να εξυπηρετηθεί, ενώ μπροστά του βρίσκονται N άτομα. Αν περιμένει μέχρι να τελειώσει θά έχει κέρδος r . Ομως έχει κόστος c για κάθε χρονική μονάδα που στέκεται περιμένοντας. Η πιθανότητα να τελειώσει `ενα άτομο απ' αυτά που βρίσκονται μπροστά του κάθε χρονική μονάδα ε`ιναι ίση με p . Το πρόβλημα είναι να βρεί την πολιτική αναμονής που μεγιστοποιεί το αναμενομένο κε`ρδος του.
Μοντελοποιήστε τοπρόβλημα με μεθόδους του Δυναμικο`υ Προγραμματισμο`υ και προσπαθείστε να λύσετε την εξίσωση βελτιστοποίησης.

9. Ενας υπάλληλος πηγαίνει με το αυτοκίνητο του στη δουλειά και θέλει να βρεί μια βέλτιστη πολιτική όσον αφορά το παρκάρισμα. Μπορεί να βάλει το αυτοκίνητο σε γκαράζ, απόφαση που έχει κόστος B , είτε μπορεί να προσπαθήσει να παρκάρει στο δρόμο. Σ'αυτη την περίπτωση όμως, όσο πιο μακριά απ`ο το χώρο της δουλειάς του αφήσει το αυτοκίνητο , τόσο περισσότερο περπάτημα θα πρέπει να κάνει. Εστω οτι στο δρόμο υπάρχουν N τετράγωνα πριν και N μετά το κτίριο που δουλεύει, στα οποία μπορεί να βρεί θεση να παρκάρει, και τα οποία συμβολίζουμε με $-N, -N+1, \dots, -1, 0, 1, \dots, N-1, N$. Το τετράγωνο με αριθμ`ο 0 είναι μπροστά στο κτίριο, κι εκε`ι απαγορεύεται η στάθμευση. Το κόστος που έχει αν αφήσει το αυτοκίνητο στο τετράγωνο με αριθμό n , εκτιμά `οτι είναι ίσο με $|n|$. Εστω p_n , $n = -N, \dots, -1, 1, \dots, N$ η πιθανότητα να βρεί κενή θέση στο n τετράγωνο.

α) Βρείτε ένα μοντέλο Δυναμικο`υ Προγραμματισμο`υ που δίνει τη βέλτιστη πολιτική ως πρός το αναμένομενο κόστος. Υποθέστε οτι κάθε φορά ο υπάλληλος βλέπει μόνο ένα τετράγωνο .
β) Γράψτε το αντίστοιχο πρωτεύον και το δυικό γραμμικό πρόγραμμα αν $N = 3$.

(Υπόδειξη . Η Κατάσταση του συστήματος μπορεί να είναι ο αριθμός του τετραγώνου και το αν είναι ή οχι γεμάτο).

10 . *Ας περιγράψουμε ένα τυχερό παιγνίδι στο οποίο θα αναφερθούμε συχνά και απ`ο τις επόμενες ασκήσεις (11 εως 16) .*
Ο παίκτης παίρνει πληροφορίες για τα αποτελέσματα μιάς σειράς ανεξάρτητων αγώνων μπάσκετ. Το πρόβλημα είναι οτι τις πληροφορίες τις παίρνει μέσα απο ενα τηλεπικοινωνιακό κανάλι που εχει θόρυβο . Ετσι για καθε αγώνα το κανάλι εχει πιθανότητα ίση με p να μεταδώσει το σωστό αποτέλεσμα και $q = 1\text{-}p$ το αντίθετο. Αφού πάρει μιά πληροφορία, έχει τη δυνατότητα να στοιχηματίσει υπέρ της μιάς ή της άλλης ομάδας όσο ποσό θέλει από όσο έχει στη διάθεση του εκείνη τη στιγμή. Υποθέτοντας οτι ο παίκτης ξεκινάει με αρχικό ποσό x, δείξτε χρησιμοποι`οντας επιχειρήματα Δ.Π., οτι για να μεγιστοποιήσει
το αναμενόμενο κεφά
λαιο του στο τέλος N σταδίων τού παιγνιδιού, πρέπει να στοιχιματίσει ολόκληρο το κεφάλαιο του σε κάθε στάδιο αν $p > 1/2$ και καθόλου αν $p < 1/2$.

3.11)Στο παιγνίδι της άσκησης 3.10, υποθέτουμε οτι ο παίκτης θέλει να μεγιστοποιήσει την αναμενόμενη τιμή του λογαρίθμου του κεφαλαίου του μετά N στάδια. Αν χρησιμοποίηση την ίδια πολιτική σε κάθε στάδιο, βρείτε το λόγο τού ποσού που στοιχηματίζει προς το συνολικό κεφάλαιο.

3.12) Ξανά στην άσκηση 3.10. Ας υποθέσουμε οτι ο παίκτης θέλει να μεγιστοποιήσει την αναμενόμενη τιμή του κεφαλαίου του μετά N στάδια παιγνιδιού.
Εστω οτι η $f(x)$ δηλώνει την αναμενόμενη τιμή που πετυχαίνει όταν χρησιμοποιεί τη βέλτιστη πολιτική, ξεκινά με αρχικό ποσό x και παίζει N φορές. Δείξτε οτι
$f(x) = \max[pf_N(x + y) + qf_N(x\text{-}y)]$, $N = 1,2,...$
$f_1(x) = \max [p \log(x + y) + q \log(x\text{-}y)]$.

3.12) Στην άσκηση 3.12 δείξτε επαγωγικά οτι:
$f_n(x) = \log x + N k$,
όπου
$k = \max[p \log (1 + r) + q \log(1\text{-}r)]$, *και*
επομένως υπάρχει ενας αριθμός r_o έτσι ώστε η βέλτιστη πολιτική σε κάθε στάδιο ορίζεται από τη σχέση $y = r_o\, x$.

3.14) Θεωρείστε την περίπτωση, όπου οι πιθανότητες σωστής μετάδοσης του παιγνιδιού εξαρτώνται απο το στάδιο του παιγνιδιού. Βρείτε τις αντίστοιχες εξισώσεις βελτιστοποίησης και προσδιορίστε τη δομή της βέλτιστης πολιτικής.

3.15) Αυξάνοντας κατά ένα ακόμη βήμα την πολυπλοκότητα του μοντέλου, θεωρούμε οτι τα διαδοχικά σήματα του καναλιού δεν είναι ανεξάρτητα. Ετσι πού η πθανότητα σωστής μετάδοσης στο στάδιο k εξαρτάται απο τη μετάδοση στο στάδιο $k\text{-}1$, με τρόπο που θα ορισθεί παρακάτω. Για $x > 0$ ορίζουμε τις ποσότητες :

$f(x)$: αναμενόμενη τιμή του τελικού κεφαλαίου πού θα αποκτηθεί στα υπόλοιπα k στάδια από τα N της όλης διαδικασίας, αν αρχίσουμε με κεφάλαιο x την πληροφορία οτι η μετάδοση του $k\text{-}1$ στάδιου έγινε σωστά, και χρησιμοποιόντας βέλτιστη πολιτική.

$g_k(x)$: η αντίστοιχη με την $f_k(x)$ τιμή όταν στο $k\text{-}1$ στάδιο η μετάδοση ήταν λαθεμένη.

Τότε δείξτε οτι η εξίσωση βελτιστοποίησης χωρίζεται στις παρακάτω δύο εξισώσεις:

$$f_k(x) = \max_{0 \le y \le x} [p_{N-k+1} f_{k-1}(x+y) + (1-p_{N-k+1}) g_k(x-y)] ,$$

$$g_k(x) = \max_{0 \le y \le x} [r_{N-k+1} f_{k-1}(x+y) + (1-r_{N-k+1}) g_{k-1}(x-y)] ,$$

οπου p_k *: η πιθανότητα σωστής μετάδοσης του k-οστού σήματος, δεδομένου ότι το k-1 μεταδόθηκε σωστά.*

r_k *: η πιθανότητα σωστής μετάδοσης k-οστού σήματος λάθος.*

Δείξτε οτι $f_k(x) = logx + \alpha_k$, $g_k(x) = logx + b_k$. *Προσδιορίστε τα* α_k, b_k *καθώς επίσης και τη δομή της βέλτιστης πολιτικής.*