ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΕΦΑΡΜΟΣΜΕΝΕΣ ΜΑΘΗΜΑΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ

ΜΑΡΚΟΒΙΑΝΕΣ ΔΙΑΔΙΚΑΣΙΕΣ ΑΠΟΦΑΣΕΩΝ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΜΠΑΛΗΣ ΑΘΑΝΑΣΙΟΣ

Τριμελής επιτροπή εξέτασης:

Κοκκίνης Βασίλειος, Επίκουρος Καθηγητής Ε.Μ.Π

Κολέτσος Ιωάννης, Επίκουρος Καθηγητής Ε.Μ.Π (Επιβλέπων)

Στεφανέας Πέτρος, Επίκουρος Καθηγητής Ε.Μ.Π

Αθήνα, Σεπτέμβριος 2017

Copyright © 2017 Μπαλής Αθανάσιος

Εγκρίθηκε από :

Επιβλέπων: κ. Κολέτσος Ιωάννης

Επίκουρος Καθηγητής Ε.Μ.Π.

Ευχαριστίες

Πρώτα απ' όλα , θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα της μεταπτυχιακής διπλωματικής μου εργασίας, Επίκουρο Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου κ. Ιωάννη Κολέτσο, για την πολύτιμη και καθοριστική βοήθειά του, τόσο στην διεκπαιρέωση του παρόντος έργου , όσο και για τις γνώσεις και τις εμπειρίες που αποκόμισα από την συμμετοχή μου στις παραδόσεις του πάνω στο μάθημα της Επιχειρησιακής Έρευνας στην σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Ε.Μ.Π. Ομολογώ ότι θα συμμετείχα ξανά και με μεγάλη ευχαρίστηση σε μαθήματα που διδάσκει ο κ. Κολέτσος – τον ευχαριστώ ιδιαιτέρως για αυτό. Κατά δεύτερον, θα ήθελα να ευχαριστήσω όλους τους καθηγητές του Τμήματος για όσα με δίδαξαν κατά την διάρκεια της φοίτησής μου . Δεν θα μπορούσα να παραλείψω την ευγνωμοσύνη μου προς την οικογένειά μου για την ηθική υποστήριξή τους όλα αυτά τα χρόνια-ιδαιτέρως δε προς τους γονείς μου Κωνσταντίνο και Χρυσούλα Μπαλή, αλλά και προς τα αδέλφια μου για όλες τις γόνιμες συζητήσεις που έχουμε τακτικά. Επιπλέον, θα ήθελα να ευχαριστήσω από καρδιάς και τους πνευματικούς μου πατέρες Κωνσταντίνο και Σωτήριο για όλα όσα απλόχερα μου προσφέρουν. Τέλος και τω εν Τριάδι υμνουμένω Θεώ,δόξα,ευχαριστία και προσκύνησις.

Θανάσης Μπαλής

Σεπτέμβριος 2017

ΠΕΡΙΕΧΟΜΕΝΑ

Ενότητα *σελίδα*

Περίληψη…….. 1

Εισαγωγή……..... 2

1. Βασικές Έννοιες…….. 4

1.1 Βασικά στοιχεία από Θεωρία Μαρκοβιανών Αλυσίδων…….. 4

1.2 Εισαγωγή στις Μαρκοβιανές Αλυσίδες Αποφάσεων……... 10

1.2.1 Πιθανότητες μετάβασης και κόστη αποφάσεων…… 10

1.2.2 Επιλογές πολιτικής……..... 14

1.2.3 Κριτήριο βελτιστοποίησης……..... 15

1.2.4 Παραδείγματα Εφαρμογών…… 17

2. Οι Μέθοδοι Επίλυσης …… 27

2.1 Η Μέθοδος Βελτίωσης της πολιτικής……... 27

2.2 Η Μέθοδος Επαναληπτικής τιμής……..... 37

2.3 Η Επίλυση με Γραμμικό Προγραμματισμό…….. 44

2.3.1 Το απλό μοντέλο…… 44

2.3.2 Η περίπτωση των στοχαστικών περιορισμών…… 50

2.4 Σύγκριση των μεθόδων επίλυσης……..... 53

3. Λυμένες εφαρμογές …… 55

4.Συγκλίσεις των μεθόδων…….. 75

4.1 Σύγκλιση της Μεθόδου Βελτίωσης της πολιτικής…… 75

4.2 Σύγκλιση της Μεθόδου Επαναληπτικής τιμής……..... 78

Βιβλιογραφία……... 80

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία αποτελεί μία ανασκόπηση των πιο βασικών σημείων των Μαρκοβιανών Διαδικασιών Αποφάσεων (Μ.Δ.Α) .

Αρχικά γίνεται εισαγωγή στις Στοχαστικές Διαδικασίες Αποφάσεων. Πρόκειται για στοχαστικές διαδικασίες με κόστη και αποφάσεις . Η διαφορά από τις απλές στοχαστικές διαδικασίες έγκειται στο γεγονός ότι στις πρώτες η μετάβαση από μία κατάσταση i σε μία κατάσταση j εξαρτάται και από την απόφαση α που θα λάβουμε. Κάθε μία από αυτές τις μεταβάσεις συνεπάγεται κάποιο κόστος $C(i,\alpha)$ με αποτέλεσμα μετά από n βήματα να προκύπτει ένα συνολικό κόστος $V_n(i,R)$ με R τον τρόπο επιλογής των αποφάσεων. Ο στόχος μας είναι να ελαχιστοποιήσουμε το συνολικό αναμενόμενο (μέσο) κόστος ανά χρονική μονάδα (συμβολικά g*) επιλέγοντας την κατάλληλη πολιτική R . Άρα το ζητούμενο είναι η εύρεση του ελάχιστου μέσου ρυθμού κόστους g* και της βέλτιστης πολιτικής R με την οποία αυτό επιτυγχάνεται. Περιορίζουμε την μελέτη μας σε στάσιμες πολιτικές οι οποίες δεν εξαρτώνται ούτε από το χρονικό βήμα επιλογής της απόφασης, ούτε από την παρελθούσα εξέλιξη της διαδικασίας.

Για την επίλυση του παραπάνω προβλήματος χρησιμοποιούνται τρεις διαφορετικοί τρόποι :

- η μέθοδος βελτίωσης της πολιτικής
- η μέθοδος επαναληπτικής τιμής
- η θεώρηση του γραμμικού προγραμματισμού.

Κάθε μία από τις τρεις παραπάνω θεωρήσεις πλαισιώνεται από τις απαραίτητες αποδείξεις θεωρημάτων, οι οποίες γίνονται ιδιαίτερα αναλυτικά. Από τα κεφάλαια αυτά δεν θα μπορούσαν να λείπουν οι αντίστοιχοι αλγόριθμοι, οι οποίοι αποτελούν τον πυρήνα της παρούσας εργασίας.

Τέλος,το θεωρητικό κομμάτι της εργασίας συνοδεύεται από υποδειγματικά λυμένα παραδείγματα των παραπάνω μεθόδων σε διάφορες εφαρμογές όπως σε συστήματα συντήρησης μηχανημάτων, στις τηλεπικοινωνίες ή την οικονομία.

Αναλυτικότερα , στο 1ο κεφάλαιο γίνεται η εισαγωγή στις Μαρκοβιανές Διαδικασίες Αποφάσεων η οποία πλαισιώνεται από αρκετά παραδείγματα. Στο 2ο κεφάλαιο,γίνεται η ανάλυση των τριών μεθόδων επίλυσης,των αντίστοιχων αλγορίθμων και η απόδειξη των σχετικών θεωρημάτων,ενώ στο 3ο κεφάλαιο παρουσιάζονται λυμένες εφαρμογές των αλγορίθμων σε επιλεγμένα παραδείγματα του 1ου κεφαλαίου. Τέλος,στο 4ο κεφάλαιο έχουμε τις αποδείξεις των συγκλίσεων των αλγορίθμων επίλυσης .

Η παρούσα διπλωματική εργασία είναι έργο ανθρώπινο και όχι θέσφατο . Ευελπιστώ ότι με το έργο αυτό θα βοηθήσει όποιον φοιτητή ή συνάδελφο ενδιαφέρεται να γνωρίσει κάποιες εισαγωγικές πτυχές των Μ.Δ.Α και θα αποτελέσει σημαντικό ερέθισμα για περαιτέρω έρευνα και εμβάθυνση τόσο στην θεωρία όσο και στις εφαρμογές αυτού του ενδιαφέροντος τομέα της Μαθηματικής Επιστήμης.

ΕΙΣΑΓΩΓΗ

Οι Μαρκοβιανές αλυσίδες αποφάσεων αποτελούν σημείο τομής δύο ιδιαίτερα εφαρμοσμένων κλάδων της Μαθηματικής Επιστήμης: των Στοχαστικών Ανελίξεων και της Επιχειρησιακής Έρευνας.

Η Επιχειρησιακή Έρευνα πραγματεύεται την μαθηματική μοντελοποίηση και επίλυση προβλημάτων βελτιστοποίησης .Αυτό σημαίνει ότι στόχος της Επιχειρησιακής Έρευνας είναι η ανάπτυξη επιστημονικών μεθόδων για τον καλύτερο σχεδιασμό και έλεγχο οργανωμένων συστημάτων, ώστε η λειτουργία τους να επιτυγχάνει το βέλτιστο αποτέλεσμα. Με τον όρο σύστημα εννοούμε ένα σύνολο αποτελούμενο από μηχανές ή από ανθρώπους ή και από τα δύο,ενώ ο όρος βέλτιστο αποτέλεσμα παραπέμπει συνήθως σε προβλήματα μεγιστοποίησης κέρδους ή ελαχιστοποίησης κόστους (αρνητικό κέρδος).

Η επιστημονική μελέτη και η μαθηματική θεμελίωση της Επιχειρησιακής Έρευνας μέσω της ανάπτυξης κατάλληλης θεωρίας και των σχετικών αποδείξεων έγινε μετά τον Β΄ Παγκόσμιο Πόλεμο. Είχε προηγηθεί η χρησιμοποίησή της κατά την διάρκεια του πολέμου σε μεγάλης κλίμακας επιχειρήσεις όπως η Μάχη της Αγγλίας και η Μάχη του Ατλαντικού[1]. Εξάλλου εκεί οφείλει και την ονομασία της : Έρευνα Επιχειρήσεων ή Επιχειρησιακή Έρευνα (Ε.Ε)-δηλώνοντας τις στρατιωτικές επιχειρήσεις.

Μετά τον πόλεμο οι επιστήμονες δεν άργησαν να συνειδητοποιήσουν την μεγάλη αξία της Ε.Ε σε ένα ευρύ φάσμα εφαρμογών , όπως η παραγωγή , η οικονομία, οι μεταφορές,ο έλεγχος αποθεμάτων, τα δίκτυα, η εξυπηρέτηση πελατών, η πρόσληψη και κατανομή του προσωπικού και άλλα. Λόγω ακριβώς αυτών των εφαρμογών της, η Ε.Ε έχει διατηρήσει και το όνομά της ως Έρευνα Επιχειρήσεων όχι μόνο στρατιωτικού αλλά πλέον και οικονομικού χαρακτήρα.Ουσιαστικά,τα αποτελέσματα της Ε.Ε είναι ποσοτικές αναλύσεις που χρησιμοποιούνται από την διοίκηση για την λήψη κατάλληλων αποφάσεων με στόχο την βελτιστοποίηση του μελετώμενου συστήματος.

Όσον αφορά τις Στοχαστικές Διαδικασίες , αυτές πραγματεύονται την μελέτη στοχαστικών συστημάτων (συστήματα που εξαρτώνται από την τυχαιότητα[2]) τα οποία εξελίσσονται στον χρόνο (ή πιο σπάνια στον χώρο)[3]. Μερικά χαρακτηριστικά προβλήματα που επιλύονται με χρήση της Θεωρίας Στοχαστικών Διαδικασιών είναι τα εξής[4]:

---

[1] Κολέτσος Ι., Στογιάννης Δ. (2012), Εισαγωγή στην Επιχειρησιακή Έρευνα.

[2] Για τους οπαδούς της αιτιοκρατίας (ντετερμινισμός), ο όρος τύχη μπορεί να ερμηνευθεί στα μαθηματικά ως η δυσκολία πρόβλεψης αποτελέσματος. Δηλαδή αν για παράδειγμα επιθυμούμε να προβλέψουμε το αποτέλεσμα που θα φέρει η ρίψη ενός ζαριού θα μπορούσαμε να λύσουμε ένα σύστημα πολλών εξισώσεων και αγνώστων με δεδομένα την πίεση, την θερμοκρασία του τόπου, τις διαστάσεις του ζαριού, το βάρος, την αρχική του ταχύτητα και στροφορμή, την ώθηση, την δύναμη, κλπ. Κάτι τέτοιο όμως έχει αυξημένη πολυπλοκότητα και μάλιστα για ένα απλό πείραμα όπως η ρίψη ενός ζαριού. Η δυσκολία επίλυσης τέτοιων πειραμάτων με ντετερμινιστικά μέσα, οδηγεί στη υιοθέτηση του πειράματος τύχης και του όρου τυχαιότητα!

[3] Φακίνου Δ. (2012), Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες, σελ. 7.

[4] Susan H. Xu (2008), Operations Research and Management Science Handbook, Κεφ. 8: σελ. 1-4.

- Ποιο είναι το ποσοστό αγοράς που αναμένεται να κατέχει μία εταιρεία για περίοδο τριών μηνών σε συνθήκες μεγάλου ανταγωνισμού των προϊόντων της;
- Ποιο είναι το μέσο κόστος που αναμένεται να επιβαρυνθεί μια ασφαλιστική εταιρεία λόγω ενός συγκεκριμένου πελάτη σε βάθος χρόνου; Πως πρέπει να διαμορφωθούν τα ασφάλιστρα του ώστε να μην ζημιωθεί η εταιρεία;
- Ποιος είναι ο αναμενόμενος χρόνος εμφάνισης βλάβης ενός μεγάλου μηχανικού/ηλεκτρονικού συστήματος σε σειρά ή σε παράλληλη σύνδεση δεδομένης της κατανομής βλαβών στο παρελθόν;Τι πρέπει να προτιμηθεί : η συντήρησή του ή η αντικατάστασή του ;
- Ποια η πιθανότητα να αλλάξει η οικονομική/κοινωνική θέση ένας εργαζομένου στο άμεσο μέλλον ή σε βάθος χρόνου, δεδομένων των οικονομικών συγκυριών και της κοινωνικής κινητικότητας σε μία χώρα ή περιοχή; Ποιος είναι ο αναμενόμενος χρόνος για αυτήν την αλλαγή;

Το εύρος των εφαρμογών των Στοχαστικών Διαδικασιών (ή Ανελίξεων) είναι πραγματικά εντυπωσιακό, περιλαμβάνοντας μεταξύ άλλων εφαρμογές στην Μηχανική, την Φυσική, την Ιατρική, την Βιολογία,την Γενετική, τις ασφαλίσεις, τις επενδύσεις και τις χρηματοοικονομικές αναλύσεις,τις ουρές αναμονής-εξυπηρέτησης,την κοινωνιολογία και την ψυχολογία,τις τηλεπικοινωνίες,τα δίκτυα,τις στρατιωτικές επιχειρήσεις κ.ά. Ταυτόχρονα,ο κλάδος αυτός προσφέρεται και για καθαρά μαθηματικούς ερευνητικούς σκοπούς παρουσιάζοντας εξαιρετικό ενδιαφέρον από μαθηματικό-θεωρητική σκοπιά.

Όσον αφορά τις Μαρκοβιανές Διαδικασίες , με αυτές μελετούμε στοχαστικά συστήματα που έχουν την Μαρκοβιανή ιδιότητα - δηλαδή συστήματα των οποίων η μελλοντική εξέλιξη εξαρτάται μόνο από το παρούσα κατάστασή τους και όχι από την παρελθούσα ιστορία τους. Ο Ρώσος μαθηματικός A.Markov (1906) ήταν ο πρώτος που εισήγαγε και μελέτησε την εν λόγω ιδιότητα, η οποία χαρακτηρίζει πολλά πραγματικά συστήματα και ως εκ τούτου οι εφαρμογές της παρουσιάζουν μεγάλη ποικιλομορφία.

Οι Μαρκοβιανές διαδικασίες αποφάσεων (Μ.Δ.Α) αποτελούν υποπερίπτωση των Μαρκοβιανών Ανελίξεων με μεγάλη χρησιμότητα στην Επιχειρησιακή Έρευνα, από την οποία αντλούν διάφορα εργαλεία. Η αρχή τίθεται από την δεκαετία του 1950,όταν ο R. Bellman ανέπτυξε την θεωρία του δυναμικού προγραμματισμού και εν συνεχεία στην δεκαετία του 1960, ο R.Howard στηριζόμενος στην θεωρία των Μαρκοβιανών αλυσίδων και τον δυναμικό προγραμματισμό ανακαλύπτει τον αλγόριθμο βελτίωσης της πολιτικής. Από το σημείο αυτό και έπειτα η ανάπτυξη των Μ.Δ.Α είναι ραγδαία . Ταυτόχρονα, η χρησιμότητά τους απλώνεται σε ένα μεγάλο εύρος τομέων όπως η ρομποτική, ο αυτοματισμός, η οικονομία, η παραγωγή, η αξιοπιστία, οι ουρές αναμονής και ο έλεγχος αποθεμάτων. Εν τέλει, η εισαγωγή στις εν λόγω στοχαστικές διαδικασίες και η ανασκόπηση κάποιων από τις εφαρμογές τους αποτελεί το αντικείμενο μελέτης της παρούσας Μεταπτυχιακής Διπλωματικής Εργασίας .

ΚΕΦΑΛΑΙΟ 1

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Περίληψη: Στο συγκεκριμένο κεφάλαιο παρατίθενται κάποια βασικά στοιχεία από την θεωρία των Μαρκοβιανών αλυσίδων και στην συνέχεια γίνεται η εισαγωγή στην θεωρία των Μαρκοβιανών Διαδικασιών Αποφάσεων . Τέλος,παρουσιάζονται κάποια ενδεικτικά παραδείγματα αντιπροσωπευτικών εφαρμογών των Μ.Δ.Α.

1.1 ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΑΠΟ ΘΕΩΡΙΑ ΜΑΡΚΟΒΙΑΝΩΝ ΑΛΥΣΙΔΩΝ

Μία στοχαστική διαδικασία (ή ανέλιξη) από το T στο S είναι μία συλλογή από τυχαίες μεταβλητές (τ.μ.) $\{X(t,\omega)\}_{t\in T,\omega\in\Omega}$,ορισμένες σε ένα χώρο πιθανότητας

$(\Omega, \mathcal{F}, P)$, όπου

T= $\mathbb{N}_0$ ={0,1,2,…} ή T=[0, +∞) ο χρονικός ορίζοντας (συνήθως το T είναι το σύνολο χρόνων).

S: ο χώρος καταστάσεων της διαδικασίας,δηλαδή το πεδίο τιμών της $X(t,\omega)$ για κάθε t∈T και ω∈Ω. Αν το S είναι αριθμήσιμο ή πεπερασμένο, κάνουμε λόγο για στοχαστική αλυσίδα.

Ω: είναι ο δειγματικός χώρος (δ.χ.),δηλαδή το σύνολο όλων των δυνατών αποτελεσμάτων του υπό μελέτη πειράματος τύχης[5].

$\mathcal{F}$: σ-άλγεβρα των ενδεχομένων[6],θα λέγαμε ότι είναι το πεδίο ορισμού των ενδεχομένων του Ω (υπάρχουν περιπτώσεις δ.χ. Ω όπου δεν μπορούμε να θεωρήσουμε κάθε υποσύνολό του ως ενδεχόμενο[7]) .

P: μέτρο πιθανότητας, δηλαδή μία συνάρτηση P:$\mathcal{F}\rightarrow\mathbb{R}$ τέτοια ώστε

1. P(Ω)=1,
2. P(A)≥ 0, ∀ A∈$\mathcal{F}$,
3. $P(\bigcup_{\nu=1}^{\infty} A_\nu) = \sum_{\nu=1}^{\infty} P(A_\nu)$, ∀ ακολουθία $\{A_\nu\}_{\nu\in\mathbb{N}}$ της $\mathcal{F}$ ξένων ανά δύο ενδεχομένων $(A_i \cap A_j = \emptyset\ \forall\ i \neq j)$.

Μία στοχαστική διαδικασία μπορεί να παρατηρηθεί κάτω από δύο οπτικές γωνίες .Για κάθε t∈T, η τ.μ. X_t : Ω → S με ω↦ $X_t(\omega)$ περιγράφει την κατάσταση της

[5] Ο αναγνώστης ας μην ξεχνά ότι στα στοχαστικά συστήματα (συστήματα επηρεαζόμενα από τον παράγοντα τύχη) η αρχή γίνεται εξ' ορισμού με την μελέτη ενός πειράματος τύχης. Βλ. Φακίνου, Δ. (2012). «Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες», σελ. 14.

[6] Ενδεχόμενο ορίζεται το σύνολο που αποτελείται από ένα ή περισσότερα στοιχεία του Ω, δηλαδή από ένα ή περισσότερα αποτελέσματα του πειράματος τύχης. Βλ. Κούτρα, Μ. (2004). «Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές», μερος Ι, σελ. 17.

[7] Βλ. Hoel, P. ,Port, S. , Stone, C. (2002). «Εισαγωγή στη Θεωρία Πιθανοτήτων», σελ. 6-7.

διαδικασίας την χρονική στιγμή t, ενώ η συλλογή όλων των X_t δείχνει την κατάσταση του συστήματος για κάθε χρονική στιγμή t∈T.

Εναλλακτικά,μπορούμε για κάθε ω∈Ω να μελετήσουμε την τροχιά της X_ω(t) καθώς το t μεταβάλλεται στον χρόνο T. Δηλαδή, X_ω: T→ S με t ↦$X_\omega(t)$,ενώ η συλλογή αυτών (για όλα τα ω∈Ω) μας δίνει μία τυχαία συνάρτηση,δηλαδή μία τ.μ. με πεδίο ορισμού το Ω και πεδίο τιμών το S^T (συναρτήσεις από το T στο S).

Η μελέτη μας επικεντρώνεται σε Μαρκοβιανές αλυσίδες,δηλαδή διαδικασίες από το T=$\mathbb{N}_0$ σε αριθμήσιμο (ή απλά πεπερασμένο) χώρο S,οι οποίες έχουν την μαρκοβιανή ιδιότητα :

$\mathrm{P}(X_{n+1} = j|\, X_0, X_1, \dots, X_n) = \mathrm{P}(X_{n+1} = j|\, X_n)$, n∈ $\mathbb{N}_0$, j∈ S (1.1.1).

Η σχέση (1.1.1) δηλώνει ότι δοθείσης της τιμής της τ.μ. X_n(παρόν), η τ.μ. X_{n+1} (μέλλον) είναι στοχαστικά ανεξάρτητη των τ.μ. $X_0, X_1, \dots, X_{n-1}$ (παρελθούσα κατάσταση της διαδικασίας). Μία «λογοτεχνική» επεξήγηση της μαρκοβιανής ιδιότητας είναι ότι το μέλλον της Μαρκοβιανής αλυσίδας εξαρτάται από το παρελθόν μόνο μέσω του παρόντος!

Η πιθανότητα $p_{ij}(n, n+1) := \mathrm{P}(X_{n+1} = j|\, X_n = i)$ καλείται πιθανότητα μετάβασης 1ης τάξης από την κατάσταση i στην κατάσταση j κατά το (n+1)-οστό βήμα. Αν οι εν λόγω πιθανότητες δεν εξαρτώνται από το χρονικά βήμα n (στάσιμες) , δηλαδή

$$p_{ij} = \mathrm{P}(X_{n+1} = j|\, X_n = i) = \mathrm{P}(X_1 = j|\, X_0 = i)$$

τότε κάνουμε λόγο για χρονικά ομογενείς αλυσίδες $\{X_n\}$. Επιπλέον,μπορούμε να θεωρήσουμε τον πίνακα $\mathbb{P}$ =$\{p_{ij}\}_{i,j\in S}$,ο οποίος καλείται πίνακας πιθανοτήτων μετάβασης 1ης τάξης της αλυσίδας και είναι ένας στοχαστικός πίνακας[8]. Αξίζει να σημειωθεί ότι μία μαρκοβιανή αλυσίδα είναι πλήρως ορισμένη όταν δίνεται ο πίνακας $\mathbb{P}$ και η αρχική κατανομή $\pi_0(j) = P(X_0 = j)$ (συνάρτηση πιθανότητας της αρχικής τ.μ. X_0). Σε αυτήν την περίπτωση, είναι εύκολο να διαπιστώσουμε ότι για να βρεθεί η αλυσίδα στις καταστάσεις $i_0, i_1, \dots, i_{n-1}, i_n$ ∈ S διαδοχικά, η πιθανότητα[9] είναι

$\mathrm{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n) = \pi_0(j) \cdot p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$ (1.1.2).

Ορίζουμε επίσης την κατανομή π_n της κατάστασης της αλυσίδας ως την κατανομή για την οποία ισχύει $\pi_n(j) = P(X_n = j)$ ∀ j ∈S.Τότε από το Θεώρημα Ολικής Πιθανότητας , έχουμε

$\pi_{n+1}(j) = \sum_{i\in S} P(X_n = i) P(X_{n+1} = j|\, X_n = i) = \sum_{i\in S} \pi_n(i) \cdot p_{ij}$ (1.1.3),

ή ισοδύναμα σε γλώσσα πινάκων : $\pi_{n+1} = \pi_n \mathbb{P}$, ∀ n∈ $\mathbb{N}_0$ (1.1.4),

---

[8] Στοχαστικός πίνακας: έχει όλα τα στοιχεία του μη αρνητικά και επιπρόσθετα τα στοιχεία κάθε γραμμής του αθροίζουν στην μονάδα.
Βλ. Λουλάκη, Μ.(2015) «Στοχαστικές Διαδικασίες», σελ. 7.

[9] Ο τύπος προκύπτει μέσα από διαδοχικές εφαρμογές του ορισμού της δεσμευμένης πιθανότητας.
Βλ. Φακίνου, Δ. (2012). «Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες», σελ. 215.

και επομένως $\pi_n = \pi_0 \mathbb{P}^n$, όπου $\mathbb{P}^n = \{p_{ij}^{(n)}\}_{i,j\in S}$ ο πίνακας πιθανοτήτων μετάβασης ν-οστής τάξης (ή n βημάτων) ,δηλαδή[10]

$$p_{ij}^{(n)} = \mathrm{P}(X_n = j|\, X_0 = i).$$

Μάλιστα, αν επιλέξουμε $\pi_0 = \pi$ τέτοια ώστε $\pi = \pi \cdot \mathbb{P}$, προκύπτει $\pi_n = \pi$ ∀ n∈ ℕ, δηλαδή η κατανομή κατάστασης της αλυσίδας είναι στάσιμη (ανεξάρτητη του n). Τότε λέμε ότι η αλυσίδα έχει αναλλοίωτη ή στάσιμη κατανομή και ότι βρίσκεται σε κατάσταση στατιστικής ισορροπίας . Συνεπώς, για να βρούμε την στάσιμη κατανομή π αρκεί να λύσουμε το σύστημα :

$$\begin{cases} \pi = \pi \cdot \mathbb{P} \\ \sum_{i\in S} \pi(i) = 1 \\ \pi(i) \geq 0\,, \forall\, i \in \mathrm{S} \end{cases} \quad (1.1.5)\,.$$

Ένα ιδιαίτερα ενδιαφέρον θέμα είναι κατά πόσον η αλυσίδα μπορεί να μετακινείται από την μία κατάσταση σε άλλη, με ένα ή περισσότερα βήματα ή και όχι. Θα λέμε ότι δύο καταστάσεις i, j∈ S επικοινωνούν μεταξύ τους αν υπάρχουν m, n ≥0:

$p_{ij}^{(n)} > 0$ και $p_{ji}^{(m)} > 0$ (αν ισχύει μόνο ένα από τα δύο, τότε λέμε ότι μόνο μία έχει πρόσβαση στην άλλη). Η επικοινωνία των καταστάσεων είναι σχέση ισοδυναμίας που διαμερίζει τον S σε κλάσεις , ανοιχτές ή κλειστές. Σε μία ανοιχτή κλάση C ,υπάρχουν i∈C , $j \notin$C και n≥0 με $p_{ij}^{(n)} > 0$, δηλαδή οι καταστάσεις της C έχουν πρόσβαση σε καταστάσεις εκτός της C . Αν μία κλάση δεν είναι ανοιχτή ,θα λέγεται κλειστή. Εύκολα γίνεται αντιληπτό ότι αν η αλυσίδα βρεθεί σε μία κλειστή κλάση C ($X_n = i, i \in$ C) μένει για πάντα σε αυτήν!

Ωστόσο, ένα επιπλέον ζήτημα που μπορεί να ανακύψει είναι το πόσο χρόνο δαπανά η αλυσίδα σε κάθε κατάσταση – δηλαδή πόσες φορές επισκέπτεται κάθε κατάσταση. Για κάθε κατάσταση i∈C ορίζουμε το συνολικό πλήθος των επισκέψεων της αλυσίδας στην i ως V(i)=$\sum_{k=0}^{\infty} \mathbb{1}\{X_n = i\}$ (τ.μ. με τιμές στον n∈ $\mathbb{N}_0 \cup \{\infty\}$) . Για τα ω∈Ω για τα οποία V(i)<∞ , η αλυσίδα εγκαταλείπει τελικά την i και δεν επιστρέφει ποτέ σ' αυτήν. Φυσικά, αν V(i)=∞ , η αλυσίδα δεν σταματά ποτέ να επισκέπτεται την i . Συνεπώς, αν $\mathrm{P}(\mathrm{V(i)} < \infty|\, X_0 = i) = 0$, η κατάσταση i καλείται επαναληπτική, ενώ αν $\mathrm{P}(\mathrm{V(i)} < \infty|\, X_0 = i) = 1$, η κατάσταση i καλείται παροδική. Αποδεικνύεται για μια κατάσταση i ότι είτε θα είναι παροδική είτε θα είναι επαναληπτική! Μία αλυσίδα θα καλείται μη-υποβιβάσιμη αν ο χώρος καταστάσεων S είναι ολόκληρος μία κλειστή κλάση.

Ορίζουμε ως χρόνο πρώτης επιστροφής στην κατάσταση x∈S την τ.μ.

$T_i = inf\{k > 0: X_k = i\}$. Μάλιστα , για μία επαναληπτική κατάσταση x αν ισχύει επιπλέον[11] $\mathrm{E}_i[T_i] < +\infty$, τότε κάνουμε λόγο για μία γνήσια ή θετικά επαναληπτική κατάσταση. Οι ιδιότητες της παροδικότητας και επαναληπτικότητας (γνήσιας και μη)

---

[10] Αποδεικνύεται ότι $p_{ij}^{(n)} = \sum_{i_1,\dots,i_{n-1}}^{\cdots} \sum (p_{i\,i_1} \cdots p_{i_{n-1}j})$. Βλ. Λουλάκη, Μ.(2015) «Στοχαστικές Διαδικασίες», σελ. 15.

[11] $\mathrm{E}_i[T_i] = \mathrm{E}[T_i \;|X_0 = i]$: δεσμευμένη μέση τιμή.

είναι ιδιότητες κλάσης, που σημαίνει ότι αν μία κατάσταση x έχει μία από τις ιδιότητες αυτές,τότε όλες οι καταστάσεις της κλάσης στην οποία η x ανήκει,έχουν την ίδια ιδιότητα. Είναι εύλογο ότι οι παροδικές κλάσεις είναι ανοιχτές (και αντίστροφα κάθε ανοιχτή κλάση είναι παροδική),ενώ οι επαναληπτικές κλάσεις είναι κλειστές(το αντίστροφο δεν ισχύει−γενικά οι πεπερασμένες κλειστές κλάσεις είναι επαναληπτικές), που σημαίνει ότι η αλυσίδα θα αφήσει κάποια στιγμή τις παροδικές καταστάσεις με πιθανότητα 1 και μετά από πεπερασμένο αριθμό βημάτων θα βρεθεί σε κάποια επαναληπτική κατάσταση ,οπότε και θα παραμείνει για πάντα στην κλάση των επαναληπτικών καταστάσεων. Επιπλέον,αξίζει να σημειωθεί ότι π(i) =0 για κάθε κατάσταση i που δεν είναι γνήσια επαναληπτική.

Σε αρκετές εφαρμογές,αν η αλυσίδα ξεκινήσει σε μία κατάσταση x θα την επισκέπτεται ανά τακτά χρονικά διαστήματα – π.χ. σε άρτιους χρόνους . Τότε λέμε ότι η κατάσταση x είναι περιοδική και ότι η αλυσίδα παρουσιάζει την ιδιότητα της περιοδικότητας. Ορίζουμε ως περίοδο μίας κατάστασης i∈S,τον μέγιστο κοινό διαιρέτη d(i) του συνόλου {n∈ℕ: $p_{ii}^{(n)}$>0} (σύνολο δυνατών χρόνων επιστροφής στην i). Αν d(i)=1,η κατάσταση i καλείται απεριοδική. Έπεται ότι αν p_{ii} >0 τότε η κατάσταση x είναι απεριοδική και αν η αλυσίδα είναι μη-υποβιβάσιμη,όλες οι καταστάσεις είναι απεριοδικές ,οπότε γίνεται λόγος για απεριοδική αλυσίδα.

Υποθέτουμε ότι η μαρκοβιανή αλυσίδα $\{X_n\}_{n\in\mathbb{N}_0}$ είναι μη-υποβιβάσιμη, γνησίως επαναληπτική και απεριοδική. Τότε ,ανεξαρτήτως της αρχικής κατανομής της ισχύει $\lim_{n\to\infty} \pi_n = \pi$ και $\lim_{n\to\infty} p_{ij}^{(n)} = \pi(j)$, $\forall\ i \in S$, όπου η π(·) είναι η λύση του (1.1.5)[12]. Επίσης , το ποσοστό σε βάθος χρόνου που ξοδεύει η αλυσίδα σε κάθε κατάσταση i∈S, είναι π(i) (εργοδικό θεώρημα[13]). Δηλαδή

$$P\left[\frac{V_n(i)}{n} \to \pi(i), \forall\, i \in S\right] = 1 \quad (1.1.6).$$

Τέλος, αν η μετάβαση από μία κατάσταση σε άλλη κατάσταση συνεπάγεται μία αμοιβή ή κάποιο κόστος (αρνητική αμοιβή) της μορφής $R_n = R_n(X_{n-1}, X_n), n \in \mathbb{N}$ για το n-οστό βήμα , τότε :

- Η συνολική αμοιβή στα πρώτα n βήματα είναι C(n) = $\sum_{s=1}^{n} R_s(X_{s-1}, X_s)$,

$n \in \mathbb{N}$.

- Η μέση αμοιβή για μετάβαση i→j είναι r_{ij} =E[$R_n(i,j)$]<∞, i, j ∈S.
- Η μέση αμοιβή για ένα βήμα που αντιστοιχεί σε αναχώρηση της $\{X_n\}$ από την κατάσταση i είναι $r_i = \sum_j r_{ij} \cdot p_{ij}$, i∈S.
- Αποδεικνύεται[14] ότι για τον ρυθμό αμοιβής (μέση αμοιβή σ' ένα βήμα) ισχύει
  - $\lim_{n\to\infty} \frac{C(n)}{n} = \sum_j r_j \cdot \pi_j$ με πιθανότητα 1,

[12] Βλ. Λουλάκη, Μ.(2015) «Στοχαστικές Διαδικασίες», σελ. 94.
[13] Βλ. Λουλάκη, Μ.(2015) «Στοχαστικές Διαδικασίες», σελ. 99.
[14] Βλ. Βλ. Φακίνου, Δ. (2012). «Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες», σελ. 256.

- $\lim_{n\to\infty} \frac{\mathrm{E[C(n)]}}{n} = \sum_j r_j \cdot \pi_j$. (1.1.7)

Οι παραπάνω τύποι ισχύουν εφόσον η αλυσίδα έχει αναλλοίωτη κατανομή π. Επιπρόσθετα έχει γίνει η υπόθεση ότι οι R_nείναι ανεξάρτητες τ.μ. η κατανομή κάθε μίας από τις οποίες εξαρτάται από την εκάστοτε μετάβαση από κατάσταση σε κατάσταση και όχι από το χρονικό βήμα n .

Παράδειγμα 1.1(πρόβλημα εμπορικού αντιπροσώπου):

Ένας εμπορικός αντιπρόσωπος (Α) μετακινείται μεταξύ τριών πόλεων 1,2,3 στις οποίες έχει πελάτες . Ο Α επιλέγει την πόλη στην οποία θα μεταβεί με τέτοιο τρόπο ώστε είναι πολύ δύσκολο να την προβλέψει κανείς. Ωστόσο, κατόπιν συλλογής στατιστικών στοιχείων από τις επισκέψεις του,έχουμε τα εξής αποτελέσματα (ποσοστά):

- Αν μία ημέρα ο Α βρίσκεται στην πόλη 1, τότε την επόμενη ημέρα θα παραμείνει στην 1 με πιθανότητα 50% (1/2) ή θα μεταφερθεί τυχαία σε μία από τις άλλες δύο πόλεις .
- Αν κάποια ημέρα ο Α βρίσκεται στην πόλη 2 ή 3 , τότε την επόμενη ημέρα μεταβαίνει στην 1 με πιθανότητα 66.66% (2/3) , αλλιώς πηγαίνει στην άλλη πόλη.

Επίσης το κόστος μεταφοράς (σε χρηματικές μονάδες) του Α από την i-οστή πόλη στην j-οστή δίνεται από τον πίνακα

$$\mathrm{R}=(r_{ij}) = \begin{pmatrix} 0 & 2 & 4 \\ 2 & 0 & 3 \\ 4 & 3 & 0 \end{pmatrix}.$$

i) Να μοντελοποιηθεί το πρόβλημα στην μορφή στοχαστικής διαδικασίας.Είναι αυτή μαρκοβιανή;

ii) Να βρεθεί η πιθανότητα στις πρώτες επτά ημέρες ο Α να διαγράψει την πορεία

(1, 2, 1, 1, 3, 2, 3).

iii) Να βρεθεί η στάσιμη κατανομή π της μαρκοβιανής διαδικασίας καθώς και το ποσοστό σε βάθος χρόνου που ο Α βρίσκεται στην πόλη 1.

iv) Ποιο είναι το μέσο κόστος μεταφοράς του Α ανά ημέρα;

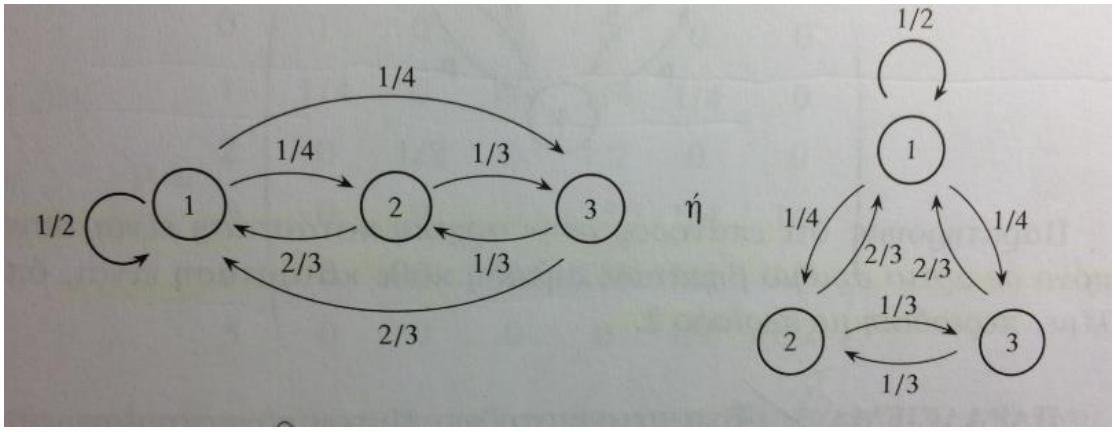
Λύση:

i) Έστω X_n: η πόλη στην οποία βρίσκεται ο Α την n-οστή ημέρα, n∈$\mathbb{N}_0$. Τότε ο χώρος καταστάσεων είναι S={1, 2, 3} και ο πίνακας πιθανοτήτων μετάβασης 1^{ης} τάξης από την κατάσταση i στην j είναι

$$\mathbb{P}=(p_{ij}) = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}.$$

Η αλυσίδα $\{X_n\}_{n\in\mathbb{N}_0}$είναι μαρκοβιανή αφού η πόλη στην οποία θα βρεθεί ο Α την επόμενη ημέρα $(X_{n+1} = j)$ εξαρτάται μόνο από την πόλη στην οποία βρίσκεται "σήμερα" $(X_{n+1} = i)$ και όχι από την παρελθούσα πορεία του $(X_0, X_1, \dots, X_{n-1})$.

Επιπλέον, όλες οι καταστάσεις επικοινωνούν μεταξύ τους , οπότε η αλυσίδα είναι μη-υποβιβάσιμη με πεπερασμένο χώρο καταστάσεων άρα είναι γνησίως επαναληπτική και μάλιστα απεριοδική (αφού p_{11} >0). Στο σχήμα 1.1 δίνεται ο γράφος[15] της αλυσίδας, δηλαδή το διάγραμμα πιθανοτήτων μετάβασης 1ης τάξης (δίκτυο με κορυφές τα στοιχεία του S και βέλη στις δυνατές μεταβάσεις με τις αντίστοιχες πιθανότητες μετάβασης).



Σχήμα 1.1: Ο γράφος των δυνατών μεταβάσεων για την αλυσίδα του παραδείγματος 1.1

ii) Ουσιαστικά ζητείται η πιθανότητα :

p = P$(X_0 = 1, X_1 = 2, X_2 = 1, X_3 = 1, X_4 = 3, X_5 = 2, X_6 = 3)$

που από τον τύπο (1.1.2) έπεται ότι p $= \pi_0(1) \cdot p_{12}p_{21}p_{11}p_{13}p_{32}p_{23} =$

$= \pi_0(1) \cdot \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{\pi_0(1)}{432}$, όπου $\pi_0(1) = P(X_0 = 1)$.

Αν έχουμε $X_0 = 1$, τότε $\pi_0(1) = 1$ και p= $1/432$, ενώ αν $X_0 = 2$ ή 3, τότε $\pi_0(1) = 0$ και p=0 . Φυσικά μπορούμε να υποθέσουμε ότι η τ.μ. X_0 ακολουθεί μία συγκεκριμένη κατανομή και να υπολογίσουμε την $P(X_0 = 1)$ από αυτήν. Για παράδειγμα,αν ο Α είναι εξ' ίσου πιθανό να βρίσκεται αρχικά σε οποιαδήποτε πόλη,που σημαίνει ότι $P(X_0 = i) = 1/3$ και άρα έχουμε $\pi_0(1) = 1/3$ και p=$1/1296$.

iii) Για να βρούμε την στάσιμη κατανομή π ≥0 αρκεί να λύσουμε το σύστημα (1.1.5):

$$\begin{cases} \pi = \pi P \\ \sum_{i\in S} \pi(i) = 1 \end{cases} \Leftrightarrow \begin{cases} \pi_1 = \pi_1 \cdot \frac{1}{2} + \pi_2 \cdot \frac{2}{3} + \pi_3 \cdot \frac{2}{3} \\ \pi_2 = \pi_1 \cdot \frac{1}{4} + \pi_2 \cdot 0 + \pi_3 \cdot \frac{1}{3} \\ \pi_3 = \pi_1 \cdot \frac{1}{4} + \pi_2 \cdot \frac{1}{3} + \pi_3 \cdot 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

η λύση του οποίου προκύπτει εύκολα :

[15] βλ. Φακίνου Δ. (2012). «Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες», σελ.217.

$$\pi = (\pi_1, \pi_2, \pi_3) = \left(\frac{8}{14}, \frac{3}{14}, \frac{3}{14}\right).$$

Τέλος, από τον τύπο (1.1.6) έχουμε ότι το ποσοστό ημερών σε βάθος χρόνου που ο Α βρίσκεται στην πόλη 1 είναι :

$\frac{V_n(1)}{n} \xrightarrow{n\to\infty} \pi(1) = \frac{8}{14} = 57.14\%$.

iv) Αρχικά υπολογίζουμε τις ποσότητες $r_i = \sum_{j=1}^{3} r_{ij} \cdot p_{ij}$, $i = 1,2,3$:

$r_1 = r_{11} \cdot p_{11} + r_{12} \cdot p_{12} + r_{13} \cdot p_{13} = 3/2$ και ομοίως $r_2 = 7/3$, $r_3 = 11/3$. Τελικά, το μέσο κόστος μεταφοράς ανά ημέρα είναι (1.1.7) :

$r = \sum_{j=1}^{3} r_j \cdot \pi_j = r_1 \cdot \pi_1 + r_2 \cdot \pi_2 + r_3 \cdot \pi_3 = \frac{15}{7} = 2{,}14$ χρηματικές μονάδες ανά ημέρα.

1.2 ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΜΑΡΚΟΒΙΑΝΕΣ ΑΛΥΣΙΔΕΣ ΑΠΟΦΑΣΕΩΝ

1.2.1 Πιθανότητες μετάβασης και κόστη αποφάσεων

Προτού διατυπώσουμε την γενική θεωρία των Μ.Δ.Α,ας εξετάσουμε το παράδειγμα 1.1 από μία διαφορετική σκοπιά. Υποθέτουμε ότι οι πόλεις 1,2,3 αντιπροσωπεύουν την πρωτεύουσα, ένα νησί και μία άλλη πόλη βαθιά στην ενδοχώρα αντίστοιχα. Επιπλέον κάνουμε την εξής σημαντική διαφοροποίηση: ο Α έχει πλέον την δυνατότητα *απόφασης*,την οποία θα συμβολίζουμε με α. Έστω α={0,1} όπου α=0 σημαίνει ότι ο Α δεν μεταβαίνει στην πρωτεύουσα (πόλη 1), ενώ α=1 σημαίνει ότι ο Α μεταβαίνει (ή παραμένει) σ' αυτήν. Συγκεκριμένα:

- αν ο Α βρίσκεται στην 1 και επιλέξει να φύγει (α=0),τότε όπως έχει παρατηρηθεί μεταβαίνει με πιθανότητα 2/5 σε άλλη πόλη, ενώ αν επιλέξει να παραμείνει στην 1 (α=1), παραμένει σε αυτήν.
- Ωστόσο, αν ο Α βρίσκεται στην πόλη 2 (νησί) και επιλέξει να αναχωρήσει για την πρωτεύουσα(α=1),τότε αυτό συμβαίνει με πιθανότητα 2/3 (π.χ. λόγω δυσμενών καιρικών συνθηκών), αλλιώς παραμένει στην 2.
- Επιπλέον, αν ο Α βρίσκεται στην 3, και αποφασίσει να μεταβεί στην 1 (α=1), τότε το επιτυγχάνει με πιθανότητα 3/4 (π.χ. λόγω απεργιών ή δυσμενών καιρικών συνθηκών), αλλιώς μεταβαίνει στην πόλη 2.
- Τέλος, αν ο Α ευρισκόμενος στην 2 ή την 3, επιλέξει να μην μεταβεί στην 1 (α=0), τότε μεταβαίνει στην 3 με πιθανότητα 2/3 ή μεταβαίνει στην 2.

Παρατηρούμε ότι οι αποφάσεις του Α μετέβαλλαν κατά πολύ τις πιθανότητες μετάβασης. Πλέον δεν μπορούμε να κάνουμε λόγο για έναν πίνακα πιθανοτήτων p_{ij}, αλλά για διάφορους πίνακες πιθανοτήτων $p_{ij}(\alpha_i)$, i, j=1,2,3 ανάλογα με την απόφαση α_i =0 ή 1 που πήρε ο Α , όταν βρισκόταν στην κατάσταση i (παρακάτω θα δούμε ότι ο συμβολισμός α_i δεν είναι πρακτικός). Για παράδειγμα,αν $\alpha_1 = 1$, $\alpha_2 = 1$, $\alpha_3 = 0$, έχουμε τον πίνακα πιθανοτήτων:

$$\mathbb{P}=[p_{ij}(\alpha_i)] = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array}.$$

Εδώ αξίζει να σημειωθεί ότι αν η αλυσίδα βρεθεί στην κατάσταση 1 θα παραμείνει για πάντα σ' αυτήν .Αυτό σημαίνει ότι ο χώρος καταστάσεων S={1,2,3} διαμερίζεται στην κλειστή κλάση C_1={1} και τις ανοιχτές κλάσεις C_2={2}, C_3={3} . Ωστόσο, κάτι τέτοιο δεν θα συνέβαινε αν είχαμε θέσει $\alpha_1 = 0$, αφού τότε $p_{11}(0) = 1/5$ και $p_{12}(0) = p_{13}(0) = 2/5 > 0$, που δηλώνει ότι όλες οι καταστάσεις του S επικοινωνούν μεταξύ τους και άρα η αλυσίδα είναι μη υποβιβάσιμη. Επομένως, η λήψη αποφάσεων ενδέχεται να προκαλεί δραστικές αλλαγές στην δομή του χώρου καταστάσεων.

Επιπρόσθετα είναι λογικό να υποθέσουμε ότι και τα κόστη μεταφοράς r_{ij} αλλάζουν με παρόμοιο τρόπο όπως οι πιθανότητες p_{ij} - δηλαδή εξαρτώνται και αυτά από τις αποφάσεις a_i που λαμβάνουμε. Μάλιστα, μπορούμε να υποθέσουμε ότι τα κόστη $r_{ij}(\alpha)$ είναι ουσιαστικά της μορφής $r_i(\alpha)$, όπως φαίνεται παρακάτω.

Σύμφωνα με τα όσα προηγήθηκαν, οι Μ.Δ.Α διαφέρουν ιδιαίτερα από τις απλές μαρκοβιανές ανελίξεις , οπότε γίνεται αναγκαίο να μελετηθούν ξεχωριστά. Ας αρχίσουμε με την παράθεση των απαραίτητων ορισμών .

Ορισμός Μ.Δ.Α:

Έστω στοχαστική διαδικασία $\{X_n\}_{n\in\mathbb{N}_0}$ με πεπερασμένο χώρο καταστάσεων S και έστω A ένα πεπερασμένο σύνολο αποφάσεων. Υποθέτουμε ότι για κάθε i∈ S , υπάρχει ένα σύνολο A(i) δυνατών αποφάσεων (επιλογών), όπου προφανώς A=$\bigcup_{i\in S}$ A(i). Όταν $X_n = i$, i∈ S και λαμβάνεται η απόφαση $\alpha_n = \alpha \in$ A(i) , τότε δύο πράγματα συμβαίνουν :

1. Έχουμε ένα άμεσο ή αναμενόμενο κόστος $c_i(a)$ ή $C(i,\alpha)$.
2. Η κατάσταση την επόμενη χρονική στιγμή n+1 είναι η $X_{n+1} = j$, j ∈S με πιθανότητα $p_{ij}(\alpha)$, ανεξάρτητα από την παρελθούσα εξέλιξη της διαδικασίας (επίσης $\sum_{j\in S} p_{ij}\,(a) = 1$).

Σχολιάζουμε τα βασικά σημεία του ορισμού. Όσον αφορά το 1ο σημείο, τονίζουμε ότι αντί για κόστος, ενδεχομένως να έχουμε αμοιβή ή κέρδος (αρνητικό κόστος). Σε κάθε περίπτωση, τα $C(i,\alpha)$ είναι φραγμένα[16] , δηλαδή $|C(i,\alpha)| \leq C$, ∀ i∈ S , α ∈ A . Σε αρκετές εφαρμογές, τα κόστη είναι συναρτήσεις της μορφής $C(i,\alpha,j)$, που σημαίνει ότι το κόστος ή η αμοιβή που θα προκύψει αν κατά την χρονική στιγμή n όντες στην κατάσταση i,λάβουμε την απόφαση $\alpha_n = \alpha$,εξαρτάται από την κατάσταση j στην οποία θα βρεθεί η αλυσίδα μας την επόμενη χρονική στιγμή n+1. Παρ' όλα αυτά, κάτι τέτοιο δεν είναι πρόβλημα αφού μπορούμε να χρησιμοποιήσουμε

[16] Βλ.Ross, S. M. (1983). Introduction to Stochastic Dynamic Programming, σελ.29.

τις ιδιότητες της μέσης τιμής, οπότε παίρνουμε ισοδύναμα[17] το αναμενόμενο κόστος να εξαρτάται από τα i, α_n :

$$C(i,\alpha) = \sum_{j\in S} C(i,\alpha,j)\cdot p_{ij}\,(a)\ ,\ \forall\ i\in \mathrm{S}.$$

Το 2ο σημείο γράφεται ισοδύναμα :

$$\mathrm{P}(X_{n+1} = j|\ X_0,\alpha_0,X_1,\alpha_1,\dots,X_n = i,\alpha_n = a) = p_{ij}(\alpha)\ .$$

Επομένως, παρατηρούμε ότι τόσο οι πιθανότητες μετάβασης, όσο και τα κόστη εξαρτώνται μόνο από την παρούσα κατάσταση της αλυσίδας και από την απόφαση που έχει επιλεγεί κατά την παρούσα χρονική στιγμή και όχι από το παρελθούσες καταστάσεις ή τις αποφάσεις. Αυτή είναι ακριβώς η μαρκοβιανή ιδιότητα, εξ' ου και η ονομασία Μαρκοβιανές Διαδικασίες Αποφάσεων.

Επιπλέον,έχουμε υποθέσει ότι $\sum_{j\in S} p_{ij}\,(a) = 1-$ κάτι τέτοιο δεν είναι γενικά απαραίτητο[18]. Όμως,η μελέτη μίας άλλης περίπτωσης πέραν της ισότητας ξεφεύγει από τους στόχους της συγκεκριμένης εργασίας .

Τέλος, να σημειωθεί ότι σε όλα τα παραπάνω $\alpha_n = \alpha$,που σημαίνει ότι η απόφαση α εξαρτάται[19] από την χρονική στιγμή n∈ $\mathbb{N}_0$ και άρα οφείλουμε να γράψουμε $C(i,\alpha_n)$ και $p_{ij}(\alpha_n)$. Στην περίπτωση που δεν υπάρχει εξάρτηση από το χρονικό βήμα n κάνουμε λόγο για χρονικά ομογενείς Μ.Δ.Α. Στις χρονικά ομογενείς Μ.Δ.Α οι πιθανότητες και τα κόστη καθορίζονται μόνο από την απόφαση που επιλέχθηκε ανεξαρτήτως της χρονικής στιγμής που έγινε η επιλογή της.

Συνοπτικά, Μ.Δ.Α καλείται η συλλογή < T, S, A, $p_{ij}(\alpha_n)$, $C(i,\alpha_n)$> ,όπου T το σύνολο χρόνου, S ο χώρος καταστάσεων και $p_{ij}(\alpha_n)$, $C(i,\alpha_n)$ οι πιθανότητες μετάβασης και τα κόστη αποφάσεων αντίστοιχα. Στην παρούσα εργασία θα επικεντρωθούμε σε T=$\mathbb{N}_0$:απεριόριστο (μη φραγμένο) χρονικό ορίζοντα και S: πεπερασμένο.

Παράδειγμα 1.2 (απλή Μ.Δ.Α) :

Υποθέτουμε το εξής απλό μοντέλο: μία Μ.Δ.Α έχει χώρο καταστάσεων που αποτελείται μόνο από δύο καταστάσεις s_1, s_2 με αντίστοιχα σύνολα αποφάσεων $\alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}$ και $\alpha_{2,1}$. Όταν $X_n = s_1$ και επιλέγεται η απόφαση $\alpha_{1.1}$ τότε προκύπτει μία αμοιβή 5 μονάδων και η αλυσίδα μεταβαίνει την επόμενη χρονική στιγμή στην s_2 με πιθανότητα 0.5, ενώ παραμένει στην s_1 με πιθανότητα επίσης 0.5. Αν $X_n = s_1$ και επιλεγεί η $\alpha_{1,2}$, τότε η διαδικασία καταλήγει με βεβαιότητα στην s_2 και η αμοιβή είναι 10 μονάδες. Αν επιλεγεί η $\alpha_{1,3}$ δεν υπάρχει αμοιβή ή κόστος αλλά η αλυσίδα παραμένει στην s_1. Τέλος,αν $X_n = s_2$,τότε θα επιλεγεί αναγκαστικά η $\alpha_{2,1}$ και θα προκύψει κόστος μίας μονάδας, ενώ η διαδικασία θα παραμείνει σίγουρα στην s_2 .

---

[17]Βλ. Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 20.

[18] Βλ.Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 20.

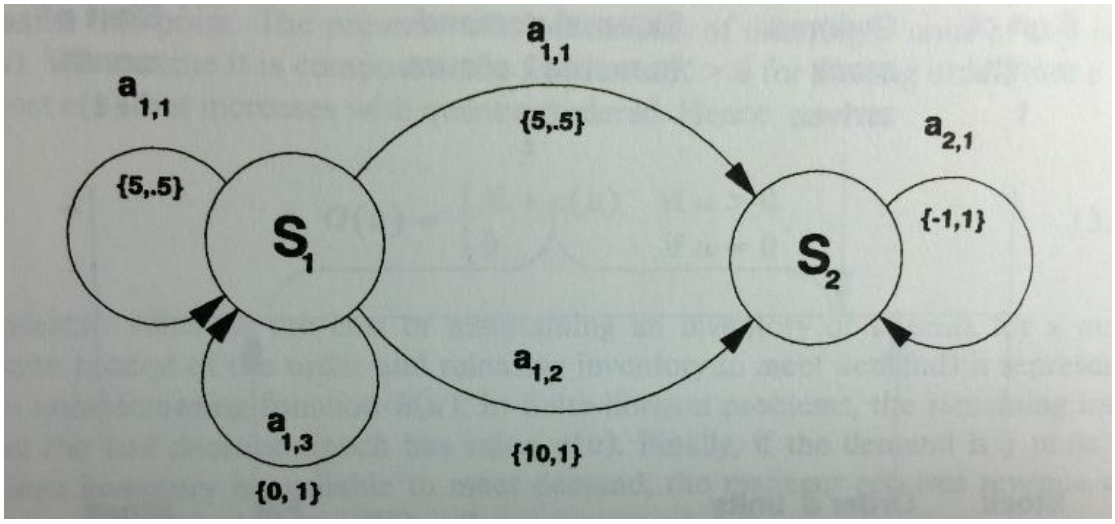
[19] Πιο αναλυτικά βλ. Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 17-18.

Εν συντομία η Μ.Δ.Α μοντελοποιείται ως εξής :

- Χώρος καταστάσεων S = $\{s_1, s_2\}$,
- Δυνατές αποφάσεις: $\alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}$ και $\alpha_{2,1}$ άρα χώρος αποφάσεων A=$\{\alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}, \alpha_{2,1}\}$ και μάλιστα:
  - A(s_1) = $\{\alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}\}$.
  - A(s_2) = $\{\alpha_{2,1}\}$.
- Πιθανότητες μετάβασης 1ης τάξης $\mathbb{P}(\alpha)$:
  - $p_{s_1 s_1}(\alpha_{1,1}) = 0.5, p_{s_1 s_2}(\alpha_{1,1}) = 0.5$, $p_{s_1 s_2}(\alpha_{1,2}) = 1, p_{s_1 s_1}(\alpha_{1,3}) = 1$
  - $p_{s_2 s_2}(\alpha_{2,1}) = 1.$
- Κόστη αποφάσεων:
- $C(s_1, \alpha_{1,1}) = +5, C(s_1, \alpha_{1,2}) = +10, C(s_1, \alpha_{1,3}) = 0.$
- $C(s_2, \alpha_{2,1}) = -1.$

Παρατηρούμε ότι εδώ τόσο οι πιθανότητες μετάβασης, όσο και τα κόστη δεν εξαρτώνται από το χρονικό βήμα,οπότε η εν λόγω αλυσίδα είναι χρονικά ομογενής.

Στο σχήμα 1.2 φαίνεται ο γράφος[20] της συγκεκριμένης Μ.Δ.Α . Να σημειώσουμε ότι οι καταστάσεις βρίσκονται σε κύκλο,ενώ μετάβαση από την κατάσταση i στην j συμβολίζεται με καμπυλοειδές βέλος. Σε κάθε βέλος αντιστοιχεί ένα άγκυστρο του οποίου το πρώτο στοιχείο αντιπροσωπεύει την αμοιβή ή το κόστος της σχετικής μετάβασης,ενώ το δεύτερο στοιχείο είναι η πιθανότητα της μετάβασης αυτής.



Σχήμα 1.2: Ο γράφος της μαρκοβιανής αλυσίδας αποφάσεων του παραδείγματος 1.2

[20] Βλ.Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 37.

1.2.2 Επιλογές Πολιτικής

Κατ' αρχάς, πριν δώσουμε τον ορισμό της "πολιτικής" στην θεωρία των Μ.Δ.Α, οφείλουμε να διακρίνουμε τα διάφορα είδη κανόνων απόφασης ή τρόπων λήψης απόφασης. Όταν $X_n = i$,οι τρόποι να διαλέξουμε μία απόφαση είναι οι εξής[21]:

1. Μαρκοβιανός αιτιοκρατικός (ντετερμινιστικός): στην περίπτωση αυτή ο τρόπος λήψης απόφασης είναι μία συνάρτηση $d_n(i)$ που εξαρτάται μόνο από το παρόν χρονικό βήμα n και την παρούσα κατάσταση i της αλυσίδας (μαρκοβιανός) και επιλέγει μία απόφαση με πιθανότητα 1(αιτοκρατικός),δηλαδή $\alpha_n = d_n(i)$, όπου $d_n: \mathrm{S} \to A\ ,\ i \mapsto d_n(i) \in A(i)$.
2. Παρελθοντικά εξαρτώμενος αιτιοκρατικός τρόπος: εδώ ο τρόπος λήψης απόφασης εξαρτάται από την παρελθούσα ιστορία,οπότε αποτελεί μία συνάρτηση της $h_n = (s_1, a_1, \dots, s_{n-1}, a_{n-1}, s_n)$, όπου s_k=X_k, $a_k \in A(s_k)$, $k = 1,2, \dots n$, η κατάσταση και η ληφθείσα απόφαση την χρονική στιγμή k. Όταν $X_n = s_n = i$, η απόφαση $\alpha_n = d_n(\cdot)$ λαμβάνεται από το σύνολο δυνατών αποφάσεων $A(s_n)$.Ως εκ τούτου έχουμε
   $d_n: \mathrm{S} \times \mathrm{A} \times \mathrm{S} \times \cdots \times \mathrm{A} \times \mathrm{S} \to A\ ,\quad h_n \mapsto d_n(h_n) \in A(s_n)$.
3. Τυχαιοποιημένος κανόνας λήψης αποφάσεων: διακρίνεται σε δύο υποκατηγορίες:
   i. στον μαρκοβιανό τυχαιοποιημένο κανόνα, κατά τον οποίο όταν $X_n = i$ η λήψη της απόφασης α∈ $A(i)$ γίνεται σύμφωνα με μία κατανομή πιθανοτήτων $q_n(i, a)$, α∈ $A(i)$.
   ii. και στον παρελθοντικά εξαρτώμενο τυχαιοποιημένο κανόνα, κατά τον οποίο όταν $X_n = i$ η λήψη της απόφασης α∈ $A(i)$ εξαρτάται και από την παρελθούσα εξέλιξη h_n της διαδικασίας και των αντίστοιχων αποφάσεων ακολουθώντας την κατανομή $q_n(h_n, a)$, α∈ $A(i)$, $i = s_n$.

Ορισμός πολιτικής:

Μία ακολουθία τρόπων λήψης αποφάσεων για κάθε χρονική στιγμή n (όχι αναγκαστικά ίδιου είδους από τα προαναφερόμενα τέσσερα για κάθε n) ονομάζεται πολιτική.

Ουσιαστικά, πρόκειται για μία στρατηγική που μας εφοδιάζει με τον κανόνα επιλογής απόφασης για κάθε χρονική στιγμή και κάτω από οποιαδήποτε προηγούμενη εξέλιξη ή απόφαση της αλυσίδας. Σε περίπτωση που η πολιτική δεν εξαρτάται από το χρονικό βήμα,καλείται *στάσιμη πολιτική*[22]. Το ενδιαφέρον μας επικεντρώνεται στην κλάση των στάσιμων μαρκοβιανών αιτιοκρατικών πολιτικών[23],δηλαδή στις ακολουθίες[24] :

$$R=(R_i)_{i\in S} \text{ με } R_i: \mathrm{S} \to \mathrm{A}\ , i \mapsto R_i \in A(i), i \in \mathrm{S}.$$

[21] Βλ.Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 21-22.
[22] Βλ. Φακίνου Δ.(2007). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 362.
[23] Θα δούμε παρακάτω γιατί συμβαίνει κάτι τέτοιο.
[24] Εδώ χρησιμοποιούμε ακολουθία ως προς τις καταστάσεις i∈S και όχι ως προς τον χρόνο n αφού τέτοια εξάρτηση δεν υφίσταται (εξ' ορισμού των στάσιμων πολιτικών).

Αν συμβολίσουμε με X_n(R) την κατάσταση της υπό μελέτη αλυσίδας τη χρονική στιγμή n όταν εφαρμόζεται η στάσιμη μαρκοβιανή αιτιοκρατική πολιτική R, τότε εύκολα γίνεται αντιληπτό ότι η στοχαστική διαδικασία $\{X_n(\mathrm{R})\}_{\mathrm{n}\in\mathbb{N}_0}$είναι μία μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης $p_{ij}(R_i)$,οι οποίες διαβάζονται ως πιθανότητες μετάβασης από την i-οστή στην j-οστή κατάσταση κάτω από την πολιτική R=$(R_i)_{i\in S}$. Υποθέτοντας πεπερασμένο χώρο καταστάσεων και θέτοντας χάρην απλοποίησης συμβολισμού S={1,2,3,…,s}, έχουμε τον πίνακα πιθανοτήτων:

$$\mathbb{P}(\mathrm{R})=\begin{pmatrix} p_{11}(R_1) & p_{12}(R_1) & \cdots & p_{1s}(R_1) \\ p_{21}(R_2) & p_{22}(R_2) & \cdots & p_{2s}(R_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_{s1}(R_s) & p_{s2}(R_s) & \cdots & p_{ss}(R_s) \end{pmatrix}.$$

Στα παρακάτω αν δεν σημειώνεται διαφορετικά, θεωρούμε δεδομένη την εξής υπόθεση:

Για κάθε στάσιμη μαρκοβιανή πολιτική, η μαρκοβιανή αλυσίδα $\{X_n(\mathrm{R})\}_{\mathrm{n}\in\mathbb{N}_0}$ είναι μη-υποβιβάσιμη ή πιο γενικά έχει μοναδική κλειστή κλάση επικοινωνίας (που σημαίνει ότι επιτρέπονται και παροδικές καταστάσεις εκτός αυτής),η οποία ως πεπερασμένη[25] θα είναι (γνησίως) επαναληπτική .

Επίσης όταν θα γίνεται λόγος για στάσιμη πολιτική,θα εννοείται μία στάσιμη μαρκοβιανή αιτιοκρατική.

1.2.3 Κριτήριο βελτιστοποίησης

Ένα ιδιαίτερα ενδιαφέρον ερώτημα είναι το πώς θα επιλέξουμε την στάσιμη πολιτική ή ισοδύναμα με ποιο κριτήριο θα γίνει αυτή η επιλογή. Εφόσον τα $C(i,\alpha)$ εκφράζουν κόστος θα ασχοληθούμε με το κριτήριο της ελαχιστοποίησης του ρυθμού κόστους που αποδεικνύεται ως το καταλληλότερο σε πλήθος εφαρμογών. Βέβαια,αν τα $C(i,\alpha)$ αντιπροσωπεύουν κέρδος (αρνητικό κόστος) το κριτήριο θα είναι η μεγιστοποίηση του ρυθμού κέρδους,που είναι ισοδύναμο με το προηγούμενο.

Θεωρούμε την Μ.Δ.Α $\{X_n(\mathrm{R})\}_{\mathrm{n}\in\mathbb{N}_0}$ όπου R: στάσιμη πολιτική και υποθέτουμε ότι $X_0 = i$ για κάποιο $i\in$ S και $p_{ij}^{(k)}(R)$ οι πιθανότητες μετάβασης k-οστής τάξης. Τότε το μέσο (ή αναμενόμενο) συνολικό κόστος στις n πρώτες χρονικές στιγμές (ή βήματα) όταν $X_0 = i$ και εφαρμόζεται η R είναι:

$$\mathrm{V_n}(i,R) = \mathrm{E}_R[\sum_{k=0}^{n-1} C(X_k, a_k)|X_0 = i\ ]$$

ή από τον ορισμό και την γραμμικότητα της (δεσμευμένης) μέσης τιμής :

$$\mathrm{V_n}(i,R) = \sum_{k=0}^{n-1}\sum_{j\in S} p_{ij}^{(k)}(R)\cdot C\big(j,R_j\big).$$

[25] Ο αναγνώστης ας μην λησμονεί την υπόθεση του πεπερασμένου χ.κ. S.

Στην διεθνή βιβλιογραφία συχνά απαντάται ο ορισμός

$$\mathrm{V_n}(i,R) = \mathrm{E}_R[\sum_{k=0}^{n-1} \lambda^k C(X_k, a_k)|X_0 = i\ ]$$

με 0<λ<1:παράγοντας προεξόφλησης με τον οποίο εκφράζεται το γεγονός ότι κέρδος $C(X_k, a_k)$ που θα πραγματοποιηθεί στο μέλλον αξίζει λιγότερο από ότι αν πραγματοποιούνταν σήμερα[26] (π.χ. ενδεχομένως λόγω ύπαρξης ρίσκου). Στο παρόν έργο ο χρονικός ορίζοντας δεν προκαλεί τέτοιες μεταβολές οπότε θεωρούμε λ=1 .

Ο αντικειμενικός σκοπός μας είναι να μελετήσουμε την συμπεριφορά του ρυθμού του μέσου κόστους σε βάθος χρόνου,δηλαδή το όριο $\lim_{n\to\infty} \frac{\mathrm{V}_n(i,R)}{n}$ για κάθε $i\in$ S. Όμως αφού η $\{X_n(\mathrm{R})\}_{\mathrm{n}\in\mathbb{N}_0}$ έχει μοναδική κλειστή κλάση επικοινωνίας (θετικά επαναληπτική), το όριο είναι μοναδικό και ανεξάρτητο[27] της αρχικής κατάστασης i :

$$\mathrm{g(R)} = \lim_{n\to\infty} \frac{\mathrm{V}_n(i,R)}{n} = \sum_{j\in S} \pi_j(R) \cdot C\big(j, R_j\big), \qquad (1.1.8)$$

με $\pi_j(R)$, $j\in$ S την στάσιμη κατανομή της αλυσίδας κάτω από την πολιτική R και καλείται ρυθμός μέσου κόστους ή σε βάθος χρόνου μέσο κόστος ανά μονάδα χρόνου[28].

Συνεπώς , μία στάσιμη πολιτική R^* θα *καλείται βέλτιστη ως προς τον μέσο ρυθμό κόστους* (ή πιο απλά ως προς το μέσο κόστος) αν για κάθε στάσιμη πολιτική R ισχύει $\mathrm{g(R^*)} \leq \mathrm{g(R)}$.

Επιπλέον, δεδομένου ότι ο χρονικός ορίζοντας είναι απεριόριστος και ο χώρος καταστάσεων είναι πεπερασμένος,*αποδεικνύεται ότι πάντα θα υπάρχει μία στάσιμη πολιτική η οποία θα είναι βέλτιστη* ως προς το παραπάνω κριτήριο.

Το προαναφερθέν αποτέλεσμα έχει αυξημένη σημασία, διότι κατοχυρώνει το δικαίωμα της αναζήτησης μίας βέλτιστης πολιτικής ανάμεσα στις στάσιμες. Λαμβάνοντας αυτό υπ' όψιν και χρησιμοποιώντας την (1.1.8) για *κάθε* δυνατή στάσιμη πολιτική μπορούμε να υπολογίσουμε την στάσιμη κατανομή και τον αντίστοιχο ρυθμό κόστους σε κάθε περίπτωση· εν τέλει επιλέγουμε την πολιτική που τον ελαχιστοποιεί. Δυστυχώς η τακτική αυτή έχει αυξημένη πολυπλοκότητα ακόμη και για μικρό χ.κ. S. Για παράδειγμα,για χ.κ. πλήθους N με 2 δυνατές αποφάσεις (τις ίδιες για κάθε κατάσταση), από την απλή πολλαπλασιαστική αρχή έχουμε 2^N διαφορετικές στάσιμες πολιτικές και για N=10 έχουμε συνολικά 1024 στάσιμες πολιτικές. Το εμπόδιο αυτό υπερπηδάται με μεθόδους που θα μελετήσουμε στο κεφάλαιο 2. Τέλος,υπάρχουν αρκετές περιπτώσεις στις οποίες μία τυχαιοποιημένη πολιτική είναι προτιμότερη (βέλτιστη) από μία στάσιμη,κάτι που θα αντιμετωπίσουμε στο επόμενο κεφάλαιο (υποπαράγραφος 2.3.2).

---

[26] Βλ. Ross S.M.(1983).Introduction to Stochastic Dynamic Programming, σελ. 31.

[27] Βλ. Φακίνου Δ.(2007). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 216 και Λουλάκη Μ.(2015). «Στοχαστικές Διαδικασίες», σελ. 100-101.

[28] βλ.Henk C. Tijms, Stochastic Modelling and Analysis: A Computational Approach, 1986, σελ. 165.

1.2.4 Παραδείγματα Εφαρμογών

Οι εφαρμογές των Μ.Δ.Α παρουσιάζουν εντυπωσιακό εύρος και μεγάλη ποικιλία. Ο αυτοματισμός , η οικονομία, η παραγωγή, η αξιοπιστία, οι ουρές αναμονής και ο έλεγχος αποθεμάτων είναι μερικοί από τους τομείς των εφαρμογών αυτών. Σε αυτήν την υποπαράγραφο θα δούμε κάποια ενδεικτικά παραδείγματα[29] από τις εφαρμογές των Μ.Δ.Α. Ωστόσο, η εύρεση της βέλτιστης πολιτικής και του ελάχιστου ρυθμού κόστους θα γίνουν στο Κεφάλαιο 3, αφού δεν έχουν παρουσιαστεί ακόμα οι μέθοδοι για κάτι τέτοιο.

Παράδειγμα 1.3 (συντήρηση μηχανήματος):

Ένα μηχάνημα ελέγχεται στην αρχή κάθε ημέρας και κατηγοριοποιείται σε μία κατάσταση i με i= 1, 2, …, N. Λόγω της χρήσης και της φθοράς που αντιμετωπίζει το μηχάνημα προϊόντος του χρόνου κάθε κατάσταση i είναι καλύτερη από την αμέσως επόμενη i+1 (i = 1, 2,…, N – 1). Ωστόσο,υπάρχει και η δυνατότητα επισκευής του μηχανήματος. Αν η παρούσα κατάσταση του μηχανήματος είναι i και το μηχάνημα δεν επισκευαστεί,τότε στην αρχή της επόμενης ημέρας θα βρεθεί στην κατάσταση j (που προφανώς δεν είναι καλύτερη της i άρα j ≥ i) με πιθανότητα q_{ij} , j ≥ i. Φυσικά, $q_{ij} = 0$ για j < i (το μηχάνημα δεν μπορεί να βρεθεί σε καλύτερη κατάσταση την επόμενη ημέρα εφόσον δεν επισκευάζεται) και $\sum_j q_{ij} = 1$.

Η κατάσταση i=N συμβολίζει βλάβη του μηχανήματος (ας πούμε ολική),οπότε το μηχάνημα δεν μπορεί να λειτουργήσει αν δεν επισκευαστεί. Η επισκευή αυτή (ολική) απαιτεί δύο ημέρες,στο τέλος των οποίων το μηχάνημα επιστρέφει στην κατάσταση 1. Για τις υπόλοιπες καταστάσεις i = 2, 3,…,N-1,μπορούμε είτε να το επισκευάσουμε (μερική επισκευή) είτε να μην το επισκευάσουμε καθόλου και η λειτουργία του να συνεχίσει ως έχει. Η μερική επισκευή του μηχανήματος διαρκεί μία ημέρα και φέρνει το μηχάνημα πάλι στην κατάσταση 1. Προφανώς κατά την διάρκεια των ημερών επισκευής (μερικής ή ολικής) το μηχάνημα δεν λειτουργεί.

Το ζητούμενο είναι να βρεθεί μία πολιτική συντήρησης,η οποία να ελαχιστοποιεί το ποσοστό ημερών που το μηχάνημα δεν λειτουργεί.

Απάντηση:

Αρχικά, μοντελοποιούμε το πρόβλημα. Το σύστημά μας περιγράφεται από μία Μαρκοβιανή διαδικασία (αλυσίδα) αποφάσεων όπως φαίνεται παρακάτω:

- Ορίζουμε X_n : την κατάσταση του μηχανήματος στην αρχή της n-οστής ημέρας .
- Χώρος καταστάσεων S = {1, 2, 3,…, N, N+1},

όπου η κατάσταση N+1 συμβολίζει ότι το μηχάνημα βρίσκεται στην 2η ημέρα (ολικής) επισκευής του (η συγκεκριμένη επισκευή άρχισε στην κατάσταση N).

[29] Βλ. Φακίνου Δ.(2007). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 371-373, 379,384-386,462-465.

- Δυνατές αποφάσεις $\alpha = \begin{cases} 0, \text{δεν επισκευάζω} \\ 1, \quad \text{επισκευάζω} \end{cases}$, άρα χώρος αποφάσεων A={0,1} και μάλιστα:
  - A(1) = {0} (δηλαδή στην κατάσταση 1 δεν επισκευάζω- το μηχάνημα βρίσκεται στην ιδανική κατάσταση λειτουργίας του)
  - A(i) = {0, 1} , i= 2,3,…, N-1 (δυνατότητα επιλογής επισκευής ή όχι στις ενδιάμεσες καταστάσεις)
  - A(N) = A(N+1) = {1} (το μηχάνημα είναι σε ολική επισκευή).
- Πιθανότητες μετάβασης 1ης τάξης $\mathbb{P}(R)$:
  - $p_{i1}(1) = 1, \quad 1 < i < N$ (σε περίπτωση επισκευής το μηχάνημα μεταφέρεται στην κατάσταση 1 υποχρεωτικά)
  - $p_{NN+1}(1) = p_{N+1,1}(1) = 1$
  - $p_{ij}(0) = q_{ij} \quad , j \geq i \, , \quad 1 \leq i < N$
  - 0 , αλλιώς.
- Το κόστος επισκευής είναι μία ημέρα , άρα υποθέτουμε κόστος 1. Οπότε

$$C(i,a) = \begin{cases} 0, \alpha\nu\ \alpha = 0 \\ 1, \alpha\nu\ \alpha = 1 \end{cases}.$$

Τελικά, το συνολικό ποσοστό χρόνου μη λειτουργίας είναι $g(R) = \sum_{i \in S} C(i, R_i) \cdot \pi_i(R)$. Συνεπώς,ζητείται να βρεθεί κατάλληλη πολιτική R ώστε να ελαχιστοποιηθεί η ποσότητα g(R).

Για να αντιληφθούμε καλύτερα την έννοια της στάσιμης πολιτικής ας επιλέξουμε την πολιτική R=(0,1,1,1,…, 1,1). Εδώ $R_1 = 0, R_i = 1, i = 2,3, \dots, N, N+1$ που σημαίνει ότι όταν βρισκόμαστε στην κατάσταση 1 δεν επισκευάζουμε – σε οποιαδήποτε άλλη, επισκευάζουμε.

Για την αλυσίδα X_n κάτω από την πολιτική R ισχύει ότι κάθε κατάσταση i,

i= 2, 3,…, N-1 θα βρεθεί στην κατάσταση 1 με πιθανότητα 1 (λόγω επισκευής). Για τις υπόλοιπες πιθανότητες μετάβασης 1ης τάξης ισχύουν τα γνωστά. Επομένως,ο αντίστοιχος πίνακας της αλυσίδας $\{X_n(R)\}_{n \in \mathbb{N}_0}$ είναι ο παρακάτω:

$$\mathbb{P}(R) = \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & \cdots & N & N+1 \end{array} \\ \begin{pmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1N} & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \\ N-1 \\ N \\ N+1 \end{array}.$$

Όπως φαίνεται, όλες οι καταστάσεις επικοινωνούν μεταξύ τους , οπότε ο χώρος καταστάσεων είναι μία κλειστή κλάση. Επίσης είναι εύκολο να υποτεθεί απεριοδικότητα (π.χ. αν $q_{11} > 0$[30] ή και λόγω της δομής του πίνακα πιθανοτήτων

[30] Βλ. Μ. Λουλάκη(2015) «Στοχαστικές Διαδικασίες», σελ. 89.

γενικότερα). Σε κάθε περίπτωση,ας βρούμε την στάσιμη κατανομή π. Για να γίνει κάτι τέτοιο, αρκεί να λύσουμε το σύστημα:

$$\begin{cases} \pi = \pi P \\ \sum_{i\in S} \pi(i) = 1 \end{cases} \Leftrightarrow \begin{cases} \pi_1 = \pi_1 q_{11} + \pi_2 + \pi_3 + \ldots + \pi_{N-1} + \pi_{N+1} \\ \pi_2 = \pi_1 q_{12} \\ \pi_3 = \pi_1 q_{13} \\ \vdots \\ \pi_N = \pi_1 q_{1N} \\ \pi_{N+1} = \pi_N \\ \pi_1 + \pi_2 + \pi_3 + \ldots + \pi_N + \pi_{N+1} = 1 \end{cases} .$$

Οπότε αντικαθιστώντας τα $\pi_2, \pi_3, \ldots, \pi_N, \pi_{N+1}$ στην τελευταία εξίσωση έχουμε

$$\pi_1 \cdot (1 + q_{12} + q_{13} + \ldots + 2q_{1N}) = 1$$

$$\pi_1 = (1 + q_{12} + q_{13} + \ldots + 2q_{1N})^{-1} := \beta.$$

Συνεπώς,η στάσιμη κατανομή της αλυσίδας $\{X_n(R)\}_{n\in\mathbb{N}_0}$ είναι η

$$\pi_j(R) = \begin{cases} \beta \quad , \quad j = 1 \\ \beta \cdot q_{1j} \, , \quad j = 2,3,\ldots,N \\ \beta \cdot q_{1N} \, , \quad j = N+1 \end{cases}$$

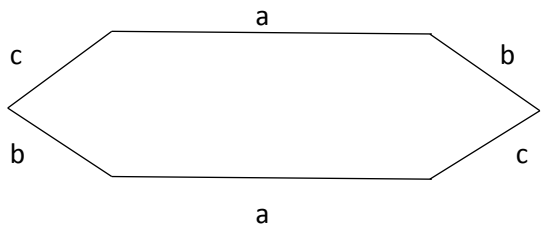
και άρα $g(R) = \sum_{i\in S} C(i, R_i) \cdot \pi_i(R) = 0 \cdot \beta + \sum_{j=2}^{N} 1 \cdot \beta \cdot q_{1j} + 1 \cdot \beta \cdot q_{1N}$

$$g(R) = \beta \cdot \ (q_{12} + q_{13} + \cdots + 2q_{1N}) = \frac{q_{12} + q_{13} + \cdots + 2q_{1N}}{1 + q_{12} + q_{13} + \cdots + 2q_{1N}} \ .$$

Ωστόσο,το ερώτημα παραμένει: είναι αυτή η βέλτιστη πολιτική συντήρησης για το συγκεκριμένο μηχάνημα;

Παράδειγμα 1.4 (ρύθμιση μηχανήματος κοπής):

Σε ένα εργοστάσιο κοπής μετάλλων,μία από τις μηχανές παράγει εξάγωνα πλαίσια της παρακάτω μορφής:



Η εν λόγω μηχανή αποτελείται από τρία ζεύγη κοπτήρων,καταχρηστικά ονομαζόμενα ως a, b, c και τα οποία χρησιμοποιούνται για την κοπή των πλευρών a, b, c του παραπάνω σχήματος αντίστοιχα .Δυστυχώς κάθε ζεύγος μπορεί να απορυθμίζεται από την κανονική του θέση *κατά την διάρκεια* κοπής ενός πλαισίου,αλλά κατόπιν υπάρχει η δυνατότητα επαναρύθμισής του. Συγκεκριμένα,μετά από κάθε κοπή μπορούμε να ρυθμίζουμε την μηχανή ώστε όλα τα ζεύγη κοπτήρων να έχουν την

κανονική τους θέση−διαδικασία που συνεπάγεται ένα κόστος 10 χρηματικών μονάδων. Αντιθέτως,αν κοπεί ένα πλαίσιο όταν k ζεύγη κοπτήρων δεν έχουν την κανονική τους θέση, προκύπτει ένα κόστος 4k χρηματικών μονάδων.

Έστω ότι οι δυνατές καταστάσεις στις οποίες μπορεί να βρεθεί η μηχανή είναι οι εξής πέντε (5) :

1. και τα τρία ζεύγη βρίσκονται στην σωστή θέση,

2. μόνο τα ζεύγη b, c βρίσκονται στην σωστή θέση,

3. μόνο το ζεύγος b βρίσκεται στην σωστή θέση,

4. μόνο το ζεύγος c βρίσκεται στην σωστή θέση,

5. κανένα ζεύγος δεν βρίσκεται στην σωστή θέση,

ενώ ο πίνακας πιθανοτήτων μετάβασης ($1^{ης}$ τάξης) της μηχανής από μία κατάσταση σε μία άλλη είναι ο παρακάτω:

$$\mathbb{Q} = (q_{ij}) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 3/4 & 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 1/4 & 1/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Ο στόχος είναι να βρεθεί μία πολιτική ρύθμισης των κοπτήρων με την οποία ελαχιστοποιείται ο ρυθμός κόστους.

Απάντηση:

Το σύστημά μας περιγράφεται από μία Μαρκοβιανή αλυσίδα αποφάσεων όπως φαίνεται παρακάτω:

- X_n: η κατάσταση της μηχανής (ως προς την θέση των κοπτήρων) μόλις πριν την κοπή του ($n+1$) – οστού πλαισίου.
- Χώρος καταστάσεων S= {1,2,3,4,5}: με στοιχεία τις πέντε καταστάσεις i όπως αυτές ορίσθηκαν στην εκφώνηση της άσκησης.
- Δυνατές αποφάσεις $\alpha = \begin{cases} 0, & \text{καμμία επέμβαση στους κοπτήρες} \\ 1, & \text{επαναρύθμιση όλων των κοπτήρων} \end{cases}$, άρα χώρος αποφάσεων A={0,1} και μάλιστα A(1) = {0},A(i) = {0, 1}, ∀ i∈S-{1}.
- Πιθανότητες μετάβασης $1^{ης}$ τάξης :
  - $p_{ij}(0) = q_{ij}$, ∀ i, j∈S ,
  - $p_{ij}(1) = q_{1j}$, ∀ i∈S γιατί αν η διαδικασία βρίσκεται σε οποιαδήποτε κατάσταση i , τότε κάτω από την απόφαση α=1, επιστρέφει στην κατάσταση 1, οπότε για να μεταβεί στην j έχει ξεκινήσει από την 1.
- Κόστη απόφασης (αναμενόμενα):
  - C(1,0) =$p_{11}(0) \cdot (4 \cdot 0)$ + $p_{12}(0) \cdot (4 \cdot 1)$=1 , αν X_n=1 ,τότε α=0 και X_{n+1}=1 (k=0 ζεύγη απορυθμισμένα) με πιθανότητα $p_{11}(0)$ ή X_{n+1}=2 (k=1 ζεύγος απορυθμισμένο) με πιθανότητα $p_{12}(0)$.
  - C(2,0) =$p_{22}(0) \cdot (4 \cdot 1)$ + $p_{23}(0) \cdot (4 \cdot 2) + p_{24}(0) \cdot (4 \cdot 2)$= 6 .

- C(3,0) =$q_{33} \cdot (4 \cdot 2) + q_{35} \cdot (4 \cdot 3)$= 9 .
- C(4,0) =$q_{44} \cdot (4 \cdot 2) + q_{45} \cdot (4 \cdot 3)$= 10.
- C(5,0) =$q_{55} \cdot (4 \cdot 3)$= 12.
- C(i,1) =$10 + p_{i1}(1) \cdot (4 \cdot 0) + p_{i2}(1) \cdot (4 \cdot 1) = 10 + q_{12} \cdot 4 = 11, \forall$ i=2,3,4,5.

Το ζητούμενο είναι η ελαχιστοποίηση του $g(R) = \sum_{j \in S} C(j, R_j) \cdot \pi_j(R)$ και η εύρεση της αντίστοιχης πολιτικής R .

Παράδειγμα 1.5 (διαχείριση αποθήκης):

Μία εταιρεία κατασκευής λυόμενων κατοικιών διαθέτει αποθήκη χωρητικότητας τεσσάρων (4) τέτοιων μονάδων. Κάθε εβδομάδα κατασκευάζονται k μονάδες με πιθανότητα $p_k = \begin{cases} 1/8, k = 0 \\ 1/2, \ k = 1 \\ 1/4, k = 2 \\ 1/8, k = 3 \end{cases}$. Στην περίπτωση που οι παραγόμενες μονάδες υπερβαίνουν την παρούσα υπολοιπόμενη χωρητικότητα της αποθήκης,τότε το περίσσευμα μεταφέρεται σε άλλο αποθηκευτικό χώρο με κόστος 30 χρηματικών μονάδων ανά μονάδα προϊόντος. Βέβαια,στο τέλος κάθε εβδομάδας υπάρχει η δυνατότητα μεταφοράς μονάδων,το πάγιο κόστος της οποίας είναι 25 χρηματικές μονάδες,ενώ υπάρχει και ένα μεταβλητό κόστος που ανέρχεται σε 5 χρηματικές μονάδες ανά μονάδα προϊόντος.

Ένα από τα στελέχη της εταιρείας πρότεινε την εξής πολιτική: αν στο τέλος της εβδομάδας οι αποθηκευμένες μονάδες είναι περισσότερες από δύο,τότε όλες οι μονάδες μεταφέρονται. Σε αντίθετη περίπτωση, δεν μεταφέρεται καμμία μονάδα. Είναι η πολιτική αυτή βέλτιστη με κριτήριο την ελαχιστοποίηση του ρυθμού κόστους;

Απάντηση:

Το σύστημά μας περιγράφεται από μία Μαρκοβιανή αλυσίδα αποφάσεων. Συγκεκριμένα:

- X_n: πλήθος μονάδων που βρίσκονται στην αποθήκη στο τέλος της n – οστής εβδομάδας.
- Χώρος καταστάσεων S= {0,1,2,3,4}.
- Δυνατές αποφάσεις $\alpha = \begin{cases} 0, & \text{δεν μεταφέρω καμμία μονάδα} \\ 1, & \text{μεταφέρω όλες τις μονάδες} \end{cases}$, άρα χώρος αποφάσεων A={0,1} και A(i) = {0, 1}, ∀ i=1,2,3,4, ενώ A(0) = {0}.
- Πιθανότητες μετάβασης 1ης τάξης :
  - $p_{00}(0) = p_0$, αφού αν X_n=0, τότε α=0 και αν X_{n+1}=0, τότε δεν παράχθηκε καμμία μονάδα (πιθανότητα p_0). Ομοίως,
  - $p_{01}(0) = p_1$, $p_{02}(0) = p_2$, $p_{03}(0) = p_3$, ενώ $p_{04}(0) = 0$,αφού αν X_n=0, τότε α=0 και για να ισχύει X_{n+1}=4,πρέπει να παραχθούν 4 μονάδες (πιθανότητα 0 στο παράδειγμά μας).

- Ομοίως, $p_{10}(0) = 0$, $p_{11}(0) = p_0$, $p_{12}(0) = p_1$, $p_{13}(0) = p_2$. Για παράδειγμα, με X_n=1 και α=0 αν X_{n+1}=2, τότε παράχθηκε μία μονάδα (πιθανότητα p_1).
- $p_{20}(0) = 0, p_{21}(0) = 0, p_{22}(0) = p_0, p_{23}(0) = p_1$.
- $p_{30}(0) = 0, p_{31}(0) = 0, p_{32}(0) = 0, p_{33}(0) = p_0$.
- $p_{40}(0) = p_{41}(0) = p_{42}(0) = p_{43}(0) = 0$.

Αφήσαμε τις πιθανότητες $p_{i4}(0)$, i=1,2,3,4 για να τις υπολογίσουμε ξεχωριστά:

- $p_{14}(0) = p_3 = 1/8$, γιατί αν X_n=1 και α=0,τότε για να προκύψει X_{n+1}=4, πρέπει να παραχθούν k=3 μονάδες.
- $p_{24}(0) = p_2 + p_3 = 3/8$, γιατί αν X_n=2 και α=0,τότε για να προκύψει X_{n+1}=4, πρέπει να παραχθούν k=2 ή k=3 μονάδες.
- $p_{34}(0) = p_1 + p_2 + p_3 = 7/8$, γιατί αν X_n=3 και α=0,τότε για να προκύψει X_{n+1}=4, πρέπει να παραχθούν k=1 ή k=2 ή k=3 μονάδες.
- $p_{44}(0) = 1$, γιατί αν X_n=4 και α=0, τότε όσες μονάδες και να παραχθούν προκύπτει X_{n+1}= 4.

Εν συντομία,

1. $p_{ij}(0) = p_{j-i}, i \leq j < 4$,
2. $p_{i4}(0) = \sum_{\substack{k:\, i+k>4 \\ k=0,1,2,3}} p_k,\ i = 1,2,3,4$,
3. αλλιώς 0.
4. Τέλος , $p_{ij}(1) = p_j$, $\forall\, i \in$S-{0}, j=0,1,2,3 , ενώ $p_{i4}(1) = 0$,

αφού για α=1 η αλυσίδα ουσιαστικά θα ξεκινήσει από την κατάσταση i=0, οπότε αποθηκεύονται οι k μονάδες που παράγονται με πιθανότητες p_k.

- Κόστη απόφασης (αναμενόμενα):
  - C(0,0) =0
  - C(1,0) =$p_{11}(0) \cdot 0 + p_{12}(0) \cdot 0 + p_{13}(0) \cdot 0 + p_{14}(0) \cdot 0$=
    $= p_0 \cdot 0 + p_1 \cdot 0 + p_2 \cdot 0 + p_3 \cdot 0 = 0$,
    διότι αν X_n=1, α=0 και π.χ. X_{n+1}=3,τότε παράχθηκαν k=2 μονάδες με πιθανότητα p_2, ενώ αν X_{n+1}=4 τότε παράχθηκαν k=3 μονάδες με πιθανότητα p_3. Ομοίως,
  - C(2,0) = $p_{22}(0) \cdot 0 + p_{23}(0) \cdot 0 + p_{24}(0) \cdot 0 + p_3 \cdot 30 \cdot (2 + 3 - 4)$=
    $= p_0 \cdot 0 + p_1 \cdot 0 + p_2 \cdot 0 + p_3 \cdot 30 \cdot 1$ =3.75.
  - C(3,0) = $p_{33}(0) \cdot 0 + p_{34}(0) \cdot 0 + p_2 \cdot 30 \cdot (3 + 2 - 4) + p_3 \cdot 30 \cdot (3 + 3 - 4)$= 15.
  - C(4,0) =$p_{44}(0) \cdot 0 + p_1(0) \cdot 30 \cdot (4 + 1 - 4) + p_2 \cdot 30 \cdot (4 + 2 - 4) + p_3 \cdot 30 \cdot (4 + 3 - 4)$= 41.25.

Ισοδύναμα: C(i,0) = $\sum_{\substack{k:\, i+k>4 \\ k=0,1,2,3}} 30(i + k - 4)p_k$,όπου k είναι οι μονάδες που παράχθηκαν κατά την n-οστή εβδομάδα .

- C(i,1) =$25 + 5i$, $\forall$ i=1,2,3,4 και εδώ η κατάσταση 0 εξαιρείται αφού A(0)={0} .

- Οπότε C(1,1)=30, C(2,1)=35, C(3,1)= 40, C(4,1)=45.

*Σημείωση:*Γενικά, ισχύει ότι C(i,1) $=25+5i+\sum_{\substack{k>4\\k\in\mathbb{N}}}30(k-4)p_k$,αν υπήρχε η επιπλέον υπόθεση ότι κατά την διάρκεια της εβδομάδας μπορούσαν να παραχθούν περισσότερες των τεσσάρων μονάδων (k>4),διότι τότε θα είχαμε ένα αναμενόμενο κόστος αναγκαστικής μεταφοράς του πλεονάσματος $\sum_{\substack{k>4\\k\in\mathbb{N}}}30(k-4)p_k$.

Το ζητούμενο παραμένει η ελαχιστοποίηση του $g(R)=\sum_{j\in S}C(j,R_j)\cdot\pi_j(R)$ και η εύρεση της αντίστοιχης πολιτικής R για την επίτευξή του.

Παράδειγμα 1.6 (πρόβλημα επιλογής διαφήμισης):

Μία εταιρεία παρακολουθεί τις πωλήσεις ενός νέου προϊόντος της, οι οποίες μπορεί να είναι υψηλές ή χαμηλές. Τα δεδομένα που έχουν συλλεγεί από την εταιρεία δείχνουν ότι αν οι πωλήσεις είναι υψηλές σε έναν μήνα,τότε θα συνεχίσουν να είναι υψηλές και τον επόμενο μήνα με πιθανότητα 50%, ενώ αν είναι χαμηλές, η πιθανότητα να γίνουν υψηλές είναι 20%. Ωστόσο,υπάρχει η δυνατότητα βελτίωσης των μηνιαίων πωλήσεων μέσω κατάλληλης διαφημιστικής εκστρατείας. Σε αυτήν την περίπτωση, οι προαναφερόμενες πιθανότητες γίνονται 80% και 40% αντίστοιχα.

Τα μηνιαία κέρδη της εταιρείας σε εκατομμύρια ευρώ έχουν ως εξής. Αν δεν χρησιμοποιηθεί διαφήμιση και οι πωλήσεις είναι υψηλές τότε τα μηνιαία κέρδη θα είναι 10 και 4,αναλόγως αν τον επόμενο μήνα οι πωλήσεις είναι υψηλές ή όχι. Στην περίπτωση κατά την οποία οι πωλήσεις είναι χαμηλές, τότε τα μηνιαία κέρδη θα είναι 7 και -2 (ζημία), αναλόγως αν τον επόμενο μήνα οι πωλήσεις είναι υψηλές ή όχι . Βέβαια, αν χρησιμοποιηθεί διαφήμιση οι παραπάνω ποσότητες γίνονται 7, 6 και 3, -5 αντίστοιχα. Το ζητούμενο είναι να βρεθεί πολιτική που μεγιστοποιεί το ρυθμό κέρδους.

Απάντηση:

Το πρώτο βήμα είναι η μοντελοποίηση του προβλήματος. Το σύστημά μας περιγράφεται από μία Μαρκοβιανή αλυσίδα αποφάσεων όπως φαίνεται παρακάτω:

- Ορίζουμε X_n : οι πωλήσεις της εταιρείας κατά τον n-οστό μήνα .
- Χώρος καταστάσεων S = {"υψηλές", "χαμηλές"} := {1, 2} .
- Δυνατές αποφάσεις $\alpha=\begin{cases}0, & \text{δεν χρησιμοποιείται διαφήμιση}\\ 1, & \text{χρησιμοποιείται διαφήμιση}\end{cases}$, άρα χώρος αποφάσεων A={0,1} και μάλιστα A(1) = A(2) = {0, 1}.
- Συμβολίζουμε με C(i,j,α) το κέρδος της εταιρείας (σε εκατομμύρια ευρώ) όταν οι μηνιαίες πωλήσεις X_n από i γίνουν j έχοντας επιλεγεί η απόφαση α∈A , i,j∈S. Τότε έχουμε :
  - C(1,1,0) = 10 αφού αν οι πωλήσεις είναι υψηλές και παραμείνουν υψηλές χωρίς να χρησιμοποιηθεί διαφήμιση,το κέρδος της εταιρείας είναι 10 .
  - C(1,2,0) = 4 αφού αν οι πωλήσεις από υψηλές γίνουν χαμηλές και δεν έχει χρησιμοποιηθεί διαφήμιση, το κέρδος της εταιρείας είναι 4 .
  - Ομοίως C(2,1,0) = 7 και C(2,2,0)= -2 .

- C(1,1,1)= 7 αφού αν οι πωλήσεις είναι υψηλές και παραμείνουν υψηλές έχοντας χρησιμοποιηθεί διαφήμιση, το κέρδος της εταιρείας είναι 7 .
- Ομοίως C(1,2,1) = 6 , C(2,1,1) = 3 , C(2,2,1) = − 5 .

Το ζήτημα που ανακύπτει είναι ότι με βάση την θεωρία που έχει διατυπωθεί, το κόστος πρέπει να είναι της μορφής C(i, α). Προτού προχωρήσουμε στην συγκεκριμένη μετατροπή ας δούμε τους διάφορους πίνακες πιθανοτήτων μετάβασης πρώτης τάξης κάτω από διάφορες πολιτικές που μπορούν να προκύψουν.

Επιλέγοντας την R_1= (0, 0) δηλαδή την πολιτική μη διαφήμισης σε κάθε περίπτωση, έχουμε

$$\mathbb{P}(R_1) = \begin{array}{cc} 1 & 2 \\ \end{array} \begin{pmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix} \begin{array}{c} 1 \\ 2 \end{array}$$

Για R_2= (1, 1) (επιλογή διαφήμισης και στις δύο καταστάσεις του S) έχουμε

$$\mathbb{P}(R_2) = \begin{array}{cc} 1 & 2 \\ \end{array} \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix} \begin{array}{c} 1 \\ 2 \end{array}$$

Για R_3= (0, 1) (όταν οι πωλήσεις είναι υψηλές δεν προχωρούμε σε διαφήμιση, ενώ όταν είναι χαμηλές προχωρούμε) έχουμε

$$\mathbb{P}(R_3) = \begin{array}{cc} 1 & 2 \\ \end{array} \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{pmatrix} \begin{array}{c} 1 \\ 2 \end{array}$$ και τα λοιπά.

Αξίζει να σημειωθεί ότι τα στοιχεία των θέσεων 11 και 21 είναι γνωστά από την εκφώνηση , ενώ τα υπόλοιπα συμπληρώνονται λόγω των ιδιοτήτων του στοχαστικού πίνακα $\mathbb{P}(R_i)$.

Ας υπολογίσουμε τα κέρδη C(i,α) :

- Το αναμενόμενο κέρδος αν η αλυσίδα βρίσκεται στην κατάσταση i=1 και επιλεγεί η απόφαση της μη χρησιμοποίησης διαφήμισης (α=0) είναι C(1,0) = $p_{11}(0)\cdot 10 + p_{12}(0)\cdot 4$ = 0.5·10 + 0.5·4 = 7.
- Ομοίως C(1,1) = $p_{11}(1)\cdot 7 + p_{12}(1)\cdot 6$ = 0.8·7 + 0.2·6 = 6.8
- C(2,0) = $p_{21}(0)\cdot 7 + p_{22}(0)\cdot(-2)$ = 0.2·7 + 0.8·(-2)= − 0.2
- C(2,1) = $p_{21}(1)\cdot 3 + p_{22}(1)\cdot(-5)$ = 0.4·7 + 0.6·(-2)= 1.6 .

Δηλαδή εν συντομία έχουμε C(1,0) = 7, C(1,1) = 6.8, C(2,0) = − 0.2 , C(2,1) = 1.6 .

Συνεπώς, το ζητούμενο τυποποιείται ως max $g(R) = \sum_{j\in S} C(j, R_j)\cdot \pi_j(R)$ (1.5.1) ή ως

$$-\min[-g(R)] = -\min\{\sum_{j\in S}[-C(j,R_j)]\cdot \pi_j(R)\} \quad (1.5.2)\ .$$

Ωστόσο, για να είμαστε συνεπείς με την θεωρία που διατυπώσαμε, το πρόβλημα πρέπει να αναχθεί στην μορφή ελαχιστοποίησης του ρυθμού κόστους. Αυτό επιτυγχάνεται αν θεωρήσουμε την ίδια αλυσίδα X_n , με κόστη C ΄(i,α) = − C(i,α) και χρησιμοποιήσουμε τον τύπο (1.5.2). Επομένως, αρκεί να βρούμε την τιμή min [$\sum_{j\in S} C'(j, R_j)\cdot \pi_j(R)$]

(ελαχιστοποίηση ρυθμού κόστους) και κατόπιν θα πάρουμε την αντίθετη ποσότητα [μεγιστοποίηση ρυθμού κέρδους – σχέση (1.5.1)[31]].

Παράδειγμα 1.7 (σύστημα τηλεπικοινωνιών):

Δίνεται σύστημα με δύο κανάλια αναμετάδοσης μηνυμάτων 1 και 2, τα οποία διαθέτουν κοινό αποθηκευτικό χώρο μνήμης M μηνυμάτων. Τα κανάλια αυτά δέχονται μηνύματα τύπου 1 και 2,με χρόνους αφίξεων που ακολουθούν ανεξάρτητες διαδικασίες Poisson με ρυθμούς λ_1 και λ_2 αντίστοιχα. Ωστόσο,το i-οστό κανάλι μπορεί να μεταδώσει μόνο μηνύματα τύπου i , για αυτό τον λόγο όταν ένα μήνυμα φτάνει σε ένα κανάλι αποφασίζεται η αποδοχή ή η απόρριψή του. Κάθε κανάλι μεταδίδει ένα μήνυμα κάθε φορά με χρόνο αποστολής που ακολουθεί την εκθετική κατανομή με παράμετρο μ_i , για το κανάλι i=1,2 .

Ως μέτρο λειτουργικότητας του συστήματος επιλέχθηκε ο σταθμισμένος μέσος $\overline{w}_*$ των απορριφθέντων μηνυμάτων τύπου 1 και 2 με συντελεστές βαρύτητας γ_1 και γ_2 αντιστοίχως. Ποια είναι η βέλτιστη πολιτική αποδοχής/απόρριψης μηνυμάτων με κριτήριο την ελαχιστοποίηση του $\overline{w}_*$;

Απάντηση:

Φαίνεται λογικό να θεωρήσουμε τις στιγμές αφίξεων ως στιγμές αποφάσεων. Όμως στην περίπτωση αυτή η εύρεση των πιθανοτήτων μετάβασης είναι αρκετά δύσκολη. Για να παρακάμψουμε την δυσκολία αυτή εισάγουμε ως τεχνητές στιγμές αποφάσεων τις στιγμές ολοκλήρωσης της μετάδοσης μηνυμάτων με επιπλέον τεχνητή απόφαση την μη-επέμβαση στο σύστημα (k=0). Αυτό σημαίνει ότι ο χώρος καταστάσεων αποτελείται από τριάδες της μορφής (x, y, k) , όπου (x , y, 0) εκφράζει το γεγονός ότι ένα μήνυμα αναμεταδίδεται αφήνοντας στον κοινό χώρο αναμονής x το πλήθος μηνύματα τύπου 1 και y μηνύματα τύπου 2. Αντίστοιχα,(x, y, k) με k=1,2 συμβολίζει την άφιξη ενός μηνύματος τύπου k στον κοινό χώρο αναμονής x το πλήθος μηνύματα τύπου 1 και y μηνύματα τύπου 2.

Οι τεχνητές στιγμές απόφασης μας διευκολύνουν στο γεγονός ότι άμεση μετάβαση από κάθε κατάσταση είναι δυνατή μόνο προς τις γειτονικές καταστάσεις και ως εκ τούτου οι πιθανότητες μετάβασης είναι πιο εύκολο να προσδιοριστούν- αρκεί να διακρίνουμε περιπτώσεις.

- Ορίζουμε X_n : κατάσταση του συστήματος στην n-οστή αναμετάδοση .
- Χώρος καταστάσεων S = {s = (i_1, i_2, k): $i_1, i_2 = 0,1,2, \dots, M$ $(i_1 + i_2 \leq M)$, $k = 0,1,2$}
    1. όπου i_1, i_2 : το πλήθος μηνυμάτων τύπου 1,2 αντίστοιχα που βρίσκονται στον χώρο αναμονής και
    2. k =1,2 το είδος του αφιχθέντος μηνύματος, ενώ
    3. k=0 : μη-επέμβαση στο σύστημα την στιγμή που ολοκληρώνεται η αναμετάδοση.

[31] Δεν ξεχνούμε ότι το κέρδος μπορεί να οριστεί και ως αρνητικό κόστος και αντίστροφα. Βλ. Φακίνου, Δ. (2007). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 361.

- Δυνατές αποφάσεις $\alpha = \begin{cases} 0, & \text{απόρριψη μηνύματος} \\ 1, & \text{αποδοχή μηνύματος} \end{cases}$, άρα χώρος αποφάσεων A={0,1} (α=0 όταν $i_1 + i_2 = M$) .
- C(s, α)=$\begin{cases} \gamma_1, \alpha\nu\ s = (i_1, i_2, 1)\ \kappa\alpha\iota\ \alpha = 0 \\ \gamma_2, \alpha\nu\ s = (i_1, i_2, 2)\ \kappa\alpha\iota\ \alpha = 0. \\ 0, \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$
- Αν $\lambda(i_1, i_2) = \lambda_1 + \lambda_2 + \mu_1\delta(i_1) + \mu_2\delta(i_2)$, όπου δ(t)=$\begin{cases} 0, t = 0 \\ 1, t > 0 \end{cases}$, τότε από στοιχειώδη θεωρία ουρών αναμονής[32] προκύπτει:
  - Για $s = (i_1, i_2, k)$ και $\alpha = 0$:

$$p_{su}(0) = \begin{cases} \dfrac{\lambda_1}{\lambda(i_1, i_2)}, u = (i_1, i_2, 1) \\ \dfrac{\lambda_2}{\lambda(i_1, i_2)}, u = (i_1, i_2, 2) \\ \dfrac{\mu_1\delta(i_1)}{\lambda(i_1, i_2)}, u = (i_1 - 1, i_2, 0) \\ \dfrac{\mu_2\delta(i_2)}{\lambda(i_1, i_2)}, u = (i_1, i_2 - 1, 0) \end{cases}.$$

  - Για $s = (i_1, i_2, 1)$ και $\alpha = 1$:

$$p_{su}(1) = \begin{cases} \dfrac{\lambda_1}{\lambda(i_1 + 1, i_2)}, u = (i_1 + 1, i_2, 1) \\ \dfrac{\lambda_2}{\lambda(i_1 + 1, i_2)}, u = (i_1 + 1, i_2, 2) \\ \dfrac{\mu_1}{\lambda(i_1 + 1, i_2)}, u = (i_1, i_2, 0) \\ \dfrac{\mu_2\delta(i_2)}{\lambda(i_1 + 1, i_2)}, u = (i_1 + 1, i_2 - 1, 0) \end{cases}.$$

  - Για $s = (i_1, i_2, 2)$ και $\alpha = 1$:

$$p_{su}(1) = \begin{cases} \dfrac{\lambda_1}{\lambda(i_1, i_2 + 1)}, u = (i_1, i_2 + 1, 1) \\ \dfrac{\lambda_2}{\lambda(i_1, i_2 + 1)}, u = (i_1, i_2 + 1, 2) \\ \dfrac{\mu_2}{\lambda(i_1, i_2 + 1)}, u = (i_1, i_2, 0) \\ \dfrac{\mu_1\delta(i_1)}{\lambda(i_1, i_2 + 1)}, u = (i_1 - 1, i_2 + 1, 0) \end{cases}.$$

- Ο στόχος παραμένει η ελαχιστοποίηση του $\overline{w}_*$. Ωστόσο , κάτι τέτοιο ισοδυναμεί με την ελαχιστοποίηση του ρυθμού μέσου κόστους g(R) ,κάτι που προκύπτει από τον τρόπο με τον οποίο ορίστηκαν τα κόστη αποφάσεων.

[32] Βλ. Κολέτσου Ι., Στογιάννη Δ. (2012). «Εισαγωγή στην Επιχειρησιακή Έρευνα», σελ.545-589.

ΚΕΦΑΛΑΙΟ 2 :

ΟΙ ΜΕΘΟΔΟΙ ΕΠΙΛΥΣΗΣ

Περίληψη: Στο κεφάλαιο αυτό μελετώνται οι τρεις μέθοδοι εύρεσης της βέλτιστης πολιτικής ως προς την ελαχιστοποίηση του μέσου ρυθμού κόστους. Αυτές είναι η μέθοδος βελτίωσης της πολιτικής, η μέθοδος της επαναληπτικής τιμής και η μέθοδος του γραμμικού προγραμματισμού. Τις μεθόδους συνοδεύουν οι αντίστοιχοι αλγόριθμοι, ενώ οι συγκλίσεις των δύο πρώτων από αυτούς αποδεικνύονται σε επόμενο κεφάλαιο (κεφάλαιο 4).

2.1 Η ΜΕΘΟΔΟΣ ΒΕΛΤΙΩΣΗΣ ΤΗΣ ΠΟΛΙΤΙΚΗΣ

Με την μέθοδο βελτίωσης της πολιτικής (Howard,1960) αρχίζοντας από μία οποιαδήποτε πολιτική R,ελέγχουμε αν αυτή είναι βέλτιστη ή όχι (θα δούμε πως θα γίνει κάτι τέτοιο). Σε περίπτωση που δεν είναι, βρίσκουμε μία πολιτική R΄ με g(R΄) ≤ g(R) και ελέγχουμε αν είναι βέλτιστη. Γνωρίζοντας ότι το πλήθος των στάσιμων πολιτικών είναι πεπερασμένο (αφού τόσο ο χώρος καταστάσεων,όσο και ο χώρος αποφάσεων είναι πεπερασμένοι), μπορούμε να συνεχίσουμε την παραπάνω διαδικασία μέχρις ότου καταλήξουμε στην βέλτιστη πολιτική. Το πλεονέκτημα της μεθόδου αυτής είναι ότι αντί να ελέγξουμε όλες τις δυνατές στάσιμες πολιτικές,ελέγχουμε ένα συνήθως μικρό υποσύνολό τους,μεταβαίνοντας από μία πολιτική σε μία καλύτερη– εξ'ου και η ονομασία μέθοδος βελτίωσης της πολιτικής.

Η αρχή γίνεται από την σχέση (1.1.8): g(R) = $\lim_{n\to\infty} \frac{V_n(i,R)}{n}$,$\forall$ $i\in$ S η οποία συνεπάγεται ότι ασυπτωτικά (δηλαδή για $n \to \infty$: μεγάλα n) ισχύει :

$$V_n(i,R) \cong n\mathrm{g(R)} + u_i(R), \forall\ i\in \mathrm{S}\ . \qquad (2.1.1)$$

Οι ποσότητες $u_i(R)$ ονομάζονται σχετικές τιμές των καταστάσεων i όταν εφαρμόζεται μία στάσιμη πολιτική R και η διαφορά τους ισούται με

$$u_i(R) - u_j(R) \cong V_n(i,R) - V_n(j,R), \forall\ i\ ,j\in \mathrm{S}.$$

Η έννοια των σχετικών τιμών είναι θεμελιώδης για την μέθοδο επαναληπτικής τιμής. Αυτές εκφράζουν την παροδική επίδραση των αρχικών καταστάσεων στο αναμενόμενο συνολικό κόστος κάτω από την εφαρμογή της πολιτικής R. Τότε η ποσότητα $u_i(R) - u_j(R)$ εκφράζει[33] την διαφορά στο μέσο συνολικό κόστος,αν η διαδικασία ξεκινήσει από την κατάσταση i σε σύγκριση με το αν ξεκινούσε από την κατάσταση j , όταν εφαρμόζεται η πολιτική R . Ισοδύναμα, η διαφορά αυτή ουσιαστικά αποτελεί το μέγιστο ποσό που συμφέρει κάποιον να διαθέσει,ώστε το σύστημα (η αλυσίδα) να ξεκινήσει από την κατάσταση j παρά από την i κάτω από την πολιτική R .

[33] Βλ. Tijms H.C. (1986). Stochastic Modelling and Analysis: A Computational Approach, σελ. 169.

Μάλιστα , αν υποθέσουμε απεριοδική αλυσίδα X_n(R), τότε τα όρια υπάρχουν[34] $\lim_{n\to\infty} p_{ij}^{(k)}(R) = \pi(j)$ κάτι που θα ισχύει και για τα όρια

$$\lim_{n\to\infty} \mathrm{V_n}(i,R) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \sum_{j\in S} p_{ij}^{(k)}(R) \cdot C\big(j, R_j\big),$$

οπότε και $u_i(R) - u_j(R) = \lim_{n\to\infty} [\mathrm{V}_n(i,R) - \mathrm{V}_n(j,R)]$ που εκφράζει την σε βάθος χρόνου διαφορά στο μέσο συνολικό κόστος,αν η διαδικασία ξεκινήσει από την κατάσταση i παρά από την κατάσταση j,κάτω από την πολιτική R .

Τόσο οι σχετικές τιμές $u_i(R)$, όσο και ο μέσος ρυθμός κόστους g(R) μπορούν να υπολογιστούν ταυτόχρονα από το Θεώρημα I - θα περιγράψουμε την βασική ιδέα του, προτού προχωρήσουμε στην διατύπωση και απόδειξή του.

Ας δούμε το μέσο συνολικό κόστος $\mathrm{V_n}(i,R)$ στα n πρώτα βήματα από την αρχή. Η αλυσίδα ξεκινάει από την κατάσταση i,ενώ εφαρμόζεται η πολιτική R. Τότε στην i επιλέγεται η απόφαση α=R_i ,προκύπτει ένα κόστος $C(i,\alpha) = C(i,R_i)$ και η αλυσίδα μεταφέρεται στην κατάσταση j με πιθανότητα $p_{ij}(R_i)$. Επιπλέον,όταν η αλυσίδα βρεθεί στην κατάσταση j το μέσο συνολικό κόστος στα επόμενα n-1 βήματα είναι $\mathrm{V}_{n-1}(j,R)$. Συνεπώς, ∀ n≥1, i∈ S ισχύει

$$\mathrm{V_n}(i,R) = C(i,R_i) + \sum_{j\in S} p_{ij}(R_i) \cdot \mathrm{V_{n-1}}(j,R). \qquad (2.1.2)$$

Συνδυάζοντας τις (2.1.1) και (2.1.2) δεν είναι δύσκολο να δειχθεί ότι ασυμπτωτικά έχουμε ∀ i∈ S :

$$u_i(R) + \mathrm{g}(R) \cong C(i,R_i) + \sum_{j\in S} p_{ij}(R_i) \cdot u_j(R)$$

(2.1.3)

ή

$$u_i(R) \cong C(i,R_i) - \mathrm{g}(R) + \sum_{j\in S} p_{ij}(R_i) \cdot u_j(R).$$

που αποτελεί το κυρίως θέμα προς απόδειξη του Θεωρήματος I. Στο σημείο αυτό ορίζουμε τις σχετικές τιμές αυστηρά. Πρώτα παραθέτουμε κάποιους απαραίτητους ορισμούς.

- $T_i(R)$: ο μέσος χρόνος 1ης επίσκεψης[35] στην r , δοθέντος ότι $X_0(R) = i$ (η αλυσίδα ξεκίνησε από την i∈ S ,κάτω από την πολιτική R).
- $K_i(R)$: το μέσο κόστος μέχρι την 1η επίσκεψη στην r ,όταν $X_0(R) = i$.

[34] Βλ. Λουλάκη Μ.(2015). «Στοχαστικές Διαδικασίες», σελ. 94.
[35] Βλ. παράγραφο 1.1 του παρόντος.

Τότε για μία γνήσια επαναληπτική κατάσταση r∈S έχουμε από την θεωρία των Ανανεωτικών διαδικασιών[36] ότι g(R) = $\frac{K_r(R)}{T_r(R)}$ (2.1.4).

Με βάση αυτό και την σχέση (2.1.1) ορίζουμε ∀ i∈ S ως *σχετικές τιμές* που αντιστοιχούν στην πολιτική R , τις ποσότητες :

$$u_i(R) = K_i(R) - \mathrm{g}(R)T_i(R) \quad (2.1.5),$$

που για i = r :γνήσια επαναληπτική κατάσταση,η (2.1.5) μέσω της (2.1.4) δίνει

$u_r(R) = 0.$

Διατυπώνουμε και αποδεικνύουμε το Θεώρημα Ι.

Θεώρημα Ι (εύρεσης ρυθμού κόστους και σχετικών τιμών):

Δίνεται στάσιμη πολιτική R τέτοια ώστε η Μαρκοβιανή αλυσίδα $\{X_n(R)\}_{n\in\mathbb{N}_0}$ να έχει μοναδική κλειστή κλάση επικοινωνίας και έστω το γραμμικό σύστημα

$$\Sigma_1: u_i = C(i, R_i) - \mathrm{g} + \sum_{j\in S} p_{ij}(R_i)\cdot u_j, \text{ με } i \in \mathrm{S} \quad (\mathrm{I0})$$

ως προς (g, $\underline{u_i}$ [37]), i ∈ S. Ισχύουν τα ακόλουθα:

Α) Τα (g,$\underline{u_i}$)= (g(R), $\underline{u_i(R)}$), ∀ i ∈ S αποτελούν λύση του Σ_1 (δηλαδή το Σ_1 έχει μία τουλάχιστον λύση).

Β) Αν (g, $\underline{u_i}$), i ∈ S τυχούσα λύση του Σ_1 , τότε ισχύει :

$$\mathrm{g}(R) = \mathrm{g} \quad \text{και} \quad u_i(R) = u_i + c \ , i \in \mathrm{S}$$

όπου c πραγματική σταθερά.

Γ) Το σύστημα

$$\Sigma_2: \begin{cases} u_i = C(i, R_i) - \mathrm{g} + \sum_{j\in S} p_{ij}(R_i)\cdot u_j \ , \ i \in \mathrm{S} \\ u_s = 0 \end{cases}$$

έχει μοναδική λύση, όπου s: αυθαίρετη κατάσταση της Μαρκοβιανής αλυσίδας $\{X_n\}$.

Απόδειξη:

Α) Υπενθυμίζουμε ότι για την γνήσια επαναληπτική κατάσταση r ∈ S, ισχύει

$$T_i(R) = 1 + \sum_{j\neq r} p_{ij}(R_i)\cdot T_j(R) \qquad (\mathrm{I1})^{38}$$

[36] Βλ. Φακίνου Δ.(2012). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 94-95.Εδώ ως κύκλο ορίσαμε τον χρόνο μεταξύ δύο διαδοχικών επισκέψεων στην r.

[37] εδώ για λόγους ευκολίας χρησιμοποιούμε το $\underline{u_i}$ για να δηλώσουμε όλα τα $u_1, u_2, \dots, u_N$ (στοιχεία του χ.κ. S που είναι πεπερασμένος με πληθικό αριθμό $|S| = N$).

[38] Για τις (Ι1) και (Ι2) βλ. Δ. Φακίνου «Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες», 2012, σελ. 254-256.

Όπου $T_i(R)$: ο αναμενόμενος χρόνος πρώτης επίσκεψης στην r , δεδομένου ότι η αλυσίδα ξεκίνησε από την κατάσταση i κάτω από την πολιτική R $[X_0(R) = i]$,

ενώ για το μέσο κόστος $K_i(R)$ της 1ης επίσκεψης στην r , ισχύει αντίστοιχα :

$$K_i(R) = C(i, R_i) + \sum_{j \neq r} p_{ij}(R_i) \cdot K_j(R) \qquad \text{(I2).}$$

Συνδυάζοντας τις (I1) και (I2) προκύπτει:

$K_i(R) - \mathrm{g}(R) \cdot T_i(R) = C(i, R_i) - \mathrm{g}(R) + \sum_{j \neq r} p_{ij}(R_i) \cdot [K_j(R) - \mathrm{g}(R) \cdot T_j(R)]$

$u_i(R) = C(i, R_i) - \mathrm{g}(R) + \sum_{j \neq r} p_{ij}(R_i) \cdot [K_j(R) - \mathrm{g}(R) \cdot T_j(R)] + p_{ir}(R_i) \cdot u_r$

$u_i(R) = C(i, R_i) - \mathrm{g}(R) + \sum_{j \neq r} p_{ij}(R_i) \cdot [K_j(R) - \mathrm{g}(R) \cdot T_j(R)] +$

$+ p_{ir}(R_i) \cdot [K_r(R) - \mathrm{g}(R) \cdot T_r(R)]$ αφού $u_r = 0$,

οπότε ισχύει τελικά:

$u_i(R) = C(i, R_i) - \mathrm{g}(R) + \sum_{j \in S} p_{ij}(R_i) \cdot [K_j(R) - \mathrm{g}(R) \cdot T_j(R)]$

ή αλλιώς $u_i(R) = C(i, R_i) - \mathrm{g}(R) + \sum_{j \in S} p_{ij}(R_i) \cdot u_j$, $i \in S$ (I3) ,

δηλαδή τα (g, $\underline{u_i}$)= (g(R), $\underline{u_i(R)}$) αποτελούν λύση του Σ_1 και το ζητούμενο αποδείχθηκε .

Β) Έστω (g, $\underline{u_i}$) , $i \in S$ τυχούσα λύση του Σ_1 [39]. Θα αποδείξουμε με επαγωγή ότι ισχύει:

$u_i = \sum_{k=0}^{m-1} \sum_{j \in S} p_{ij}^{(k)}(R) C(j, R_j) - m\mathrm{g} + \sum_{j \in S} p_{ij}^{(m)}(R)\, u_j$ (I4).

Αρχικά ισχύει για m=1:

$u_i = \sum_{j \in S} p_{ij}^{(0)}(R) C(j, R_j) - \mathrm{g} + \sum_{j \in S} p_{ij}(R_i)\, u_j$

($P^{(0)} = \{p_{ij}^{(0)}\} = I$:μοναδιαίος πίνακας)

$u_i = 1 \cdot C(i, R_i) - \mathrm{g} + \sum_{j \in S} p_{ij}(R_i)\, u_j$ που ισχύει επειδή τα (g, $\underline{u_i}$) , $i \in S$ αποτελούν λύση του Σ_1 .

Επαγωγικό βήμα:

Υποθέτουμε ότι η (I4) ισχύει για κάποιο m∈ℕ. Τότε για m+1 έχουμε:

$u_i = \sum_{k=0}^{m} \sum_{j \in S} p_{ij}^{(k)}(R) C(j, R_j) - (m+1)\mathrm{g} + \sum_{j \in S} p_{ij}^{(m+1)}(R)\, u_j$

$u_i = \sum_{k=0}^{m-1} \sum_{j \in S} p_{ij}^{(k)}(R) C(j, R_j) + \sum_{j \in S} p_{ij}^{(m)}(R) C(j, R_j) - m\mathrm{g} - \mathrm{g}$
$+ \sum_{j \in S} \sum_{z \in S} p_{iz}^{(m)}(R) p_{zj}(R)\, u_j$.

[39] Το Σ_1 έχει μία τουλάχιστον λύση όπως αποδείχθηκε στο (Α)

Προσθαφαιρούμε τον όρο $\sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j$ και προκύπτει :

$$u_i = \sum_{k=0}^{m-1}\sum_{j\in S} p_{ij}{}^{(k)}\,(R)C\big(\,j,R_j\big) - m\mathrm{g} + \sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j$$

$$-\sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j + \sum_{j\in S} p_{ij}{}^{(m)}(R)C\big(\,j,R_j\big) - \mathrm{g}$$

$$+\sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j.$$

Και λόγω της υπόθεσης του επαγωγικού βήματος έπεται ότι:

$u_i = u_i - \sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j + \sum_{j\in S} p_{ij}{}^{(m)}(R)C\big(\,j,R_j\big) - \mathrm{g}$
$+\sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j$ και αντικαθιστώντας την (I0) έχουμε:

$0 = -\sum_{j\in S} p_{ij}{}^{(m)}(R)\,\big(C\big(j,R_j\big) - \mathrm{g} + \sum_{z\in S} p_{jz}\big(R_j\big)\cdot u_z\big)$

$+\sum_{j\in S} p_{ij}{}^{(m)}(R)C\big(\,j,R_j\big) - \mathrm{g} + \sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j$

$0 = -\sum_{j\in S} p_{ij}{}^{(m)}(R)C\big(j,R_j\big) + \mathrm{g}\sum_{j\in S} p_{ij}{}^{(m)}(R) - \sum_{j\in S} p_{ij}{}^{(m)}(R)\sum_{z\in S} p_{jz}\big(R_j\big)\cdot u_z$

$+\sum_{j\in S} p_{ij}{}^{(m)}(R)C\big(\,j,R_j\big) - \mathrm{g} + \sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j$

και αφού $\sum_{j\in S} p_{ij}{}^{(m)}(R) = 1$ έπεται

$0 = \mathrm{g} - \sum_{j\in S}\sum_{z\in S} p_{ij}{}^{(m)}(R)p_{jz}\big(R_j\big)u_z - \mathrm{g} + \sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j$

και αλλάζοντας την σειρά ολοκλήρωσης

$0 = -\sum_{z\in S}\sum_{j\in S} p_{ij}{}^{(m)}(R)p_{jz}\big(R_j\big)u_z + \sum_{j\in S}\sum_{z\in S} p_{iz}{}^{(m)}(R)p_{zj}\,(R)\, u_j$

που ισχύει[40] και η επαγωγική διαδικασία ολοκληρώθηκε.

Υπενθυμίζοντας ότι το μέσο συνολικό κόστος είναι

$$V_m(i,R) = \sum_{k=0}^{m-1}\sum_{j\in S} p_{ij}{}^{(k)}\,(R)C\big(\,j,R_j\big)$$

η (I4) γίνεται

$u_i = V_m(i,R) - m\mathrm{g} + \sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j\ ,\ i \in S$ (I5) .

Διαιρούμε με m κατά μέλη:

$$\frac{u_i}{m} = \frac{V_m(i,R)}{m} - \mathrm{g} + \frac{\sum_{j\in S} p_{ij}{}^{(m)}(R)u_j}{m}$$

$$\mathrm{g} = \frac{V_m(i,R)}{m} + \frac{\sum_{j\in S} p_{ij}{}^{(m)}(R)u_j}{m} - \frac{u_i}{m}$$

[40] Η αλλαγή συμβολισμού δεν ακυρώνει το γεγονός ότι τα αθροίσματα είναι αντίθετα και άρα έχουν άθροισμα 0.

Και παίρνοντας όρια m→ ∞ κατά μέλη, έχουμε ότι:

$\frac{u_i}{m} \longrightarrow \infty$

$\frac{V_m(i,R)}{m} \longrightarrow$ g(R)

ενώ $\left|\frac{\sum_{j\in S} p_{ij}{}^{(m)}(R)u_j}{m}\right| \leq \frac{\left|\sum_{j\in S} u_j\right|}{m} \leq \frac{h}{m} \longrightarrow \infty$, με $h\in\mathbb{R}$ σταθερά.

Άρα g=g(R) .

Μένει να αποδείξουμε ότι $u_i(R) = u_i + c, i \in S$.

Πράγματι, αν (g,$\underline{u_i}$) και (g΄,$\underline{u_i{}'}$) είναι δύο λύσεις του Σ_1 ,τότε ικανοποιούν και την (I5) και επομένως g =g΄=g(R) όπως αποδείχθηκε παραπάνω. Δηλαδή έχουμε

$u_i = V_n(i,R) - m\text{g} + \sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j$,

$u_i{}' = V_n(i,R) - m\text{g} + \sum_{j\in S} p_{ij}{}^{(m)}(R)\, u_j{}'$

και αφαιρώντας κατά μέλη τις δύο προηγούμενες σχέσεις

$u_i - u_i{}' = \sum_{j\in S} p_{ij}{}^{(m)}(R_i) \cdot (u_j - u_j{}')$ (I6) με το δεύτερο μέλος να εξαρτάται από το i.

Αθροίζοντας για *m=1,2,…, n*,στην (I6) και αλλάζοντας την σειρά άθροισης έχουμε:

$$n(u_i - u_i{}') = \sum_{k=1}^{n}\sum_{j\in S} p_{ij}{}^{(m)}(R_i) \cdot (u_j - u_j{}')$$

$$u_i - u_i{}' = \frac{1}{n}\sum_{k=1}^{n}\sum_{j\in S} p_{ij}{}^{(m)}(R_i) \cdot (u_j - u_j{}')$$

$$u_i - u_i{}' = \sum_{j\in S}\left[\frac{1}{n}\sum_{k=1}^{n} p_{ij}{}^{(m)}(R_i)\right] (u_j - u_j{}')$$

Ωστόσο, από τις υποθέσεις για την $\{X_n(R)\}_{n\in\mathbb{N}_0}$(μοναδική κλειστή κλάση και πεπεράσμενος χώρος καταστάσεων) έχουμε[41] ότι για m→ ∞ ισχύει

$\frac{1}{n}\sum_{k=1}^{n} p_{ij}{}^{(m)}(R_i) \rightarrow \pi_j(R_i)\ \forall\ i\in$ S.

Μάλιστα η παραπάνω σχέση ισχύει ακόμη και για παροδικές καταστάσεις i, αφού μετά από πεπερασμένο αριθμό βημάτων θα βρεθούν στην μοναδική κλειστή κλάση[42] (η οποία ως πεπερασμένη είναι γνησίως επαναληπτική).

[41] βλ. Δ. Φακίνου, «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», 2007, σελ. 180-181.
[42] Βλ. Μ. Λουλάκη(2015) «Στοχαστικές Διαδικασίες», σελ. 26-27.

Ας σημειωθεί εδώ ότι $\pi_j(R_i)=0 \ \forall \ j$[43] που δεν ανήκει στην μοναδική θετικά επαναληπτική κλάση $S(R_i)$ του S . Εξάλλου $p_{ij}{}^{(m)}(R_i) \to 0 \ \forall \ i \in S$ και $j \notin S(R_i)$[44] .

Επομένως , παίρνοντας όρια $m \to \infty$ κατά μέλη στην (I6), έχουμε :

$$u_i - u_i{}' = \sum_{j \in S} \pi_j(R_i) \cdot (u_j - u_j{}') = c$$

c : σταθερά.

Δηλαδή , $u_i = u_i{}' + c$. Αφού κάθε $u_i(R)$, $i \in S$ είναι λύση του Σ_1 , τελικά προκύπτει $u_i(R) = u_i + c$, $i \in S$ με $\underline{u_i}$ τυχούσα λύση του Σ_1.

Γ)Αρχικά θα αποδείξουμε ότι το Σ_2 έχει μία τουλάχιστον λύση , την οποία και θα κατασκευάσουμε.

Έστω $(g, \underline{u_i})$, $i \in S$ τυχούσα λύση του Σ_1 , τότε και η $(g, \underline{u_i}')$ με $u_i{}' = u_i + k$, με $k \in \mathbb{R}$ είναι λύση του Σ_1. Πράγματι, αφού ισχύει

$u_i = C(i, R_i) - g + \sum_{j \in S} p_{ij}(R_i) \cdot u_j$ (I0)

Τότε

$$u_i{}' = C(i, R_i) - g + \sum_{j \in S} p_{ij}(R_i) \cdot u_j{}'$$

$$u_i + k = C(i, R_i) - g + \sum_{j \in S} p_{ij}(R_i) \cdot (u_j + k)$$

$$u_i + k = C(i, R_i) - g + \sum_{j \in S} p_{ij}(R_i) \cdot u_j + \sum_{j \in S} p_{ij}(R_i) \cdot k$$

Όμως αφού $\sum_{j \in S} p_{ij}(R_i) = 1$ έπεται ότι

$$u_i + k = C(i, R_i) - g + \sum_{j \in S} (p_{ij}(R_i) \cdot u_j) + k$$

$$u_i = C(i, R_i) - g + \sum_{j \in S} (p_{ij}(R_i) \cdot u_j)$$

που ισχύει από (I0).

Επιλέγουμε $k = -u_s \in \mathbb{R}$. Σε αυτήν την περίπτωση $u_i{}' = u_i - u_s$ ικανοποιεί την (I0), δηλαδή αποτελεί λύση του Σ_1 όπως αποδείχθηκε. Ταυτόχρονα ικανοποιεί και την

[43] Βλ. Μ. Λουλάκη(2015) «Στοχαστικές Διαδικασίες», σελ. 72-76.
[44] Βλ.Henk C. Tijms. (1986) Stochastic Modelling and Analysis: A Computational Approach, σελ. 178.

συνθήκη u_s'= 0. Επομένως, η u_i'= $u_i - u_s$ αποτελεί λύση του Σ_2 . Αυτό σημαίνει ότι το Σ_2 έχει μία τουλάχιστον λύση.

Απόδειξη της μοναδικότητας:

Έστω u_i'' δεύτερη λύση του Σ_2. Τότε η u_i'' επαληθεύει την (I0) και λόγω του ερωτήματος (B) ισχύει

$$u_i'' = u_i' + c,\ c \in \mathbb{R}\ .$$

Όμως , οι u_i', u_i'' αποτελούν λύσεις του Σ_2 και επομένως $u_s' = u_s'' = 0$. Με βάση αυτό η προηγούμενη σχέση δίνει c=0 και $u_i' = u_i''$ $\forall\ i \in$ S που αποδεικνύει το ζητούμενο.

Το Θεώρημα I αν και θεμελιώδες μας υποδεικνύει τον τρόπο εύρεσης των σχετικών τιμών και του μέσου ρυθμού κόστους, αλλά δεν μας παρέχει κάποιο τρόπο για να βελτιώσουμε την πολιτική που έχουμε επιλέξει. Κάτι τέτοιο θα γίνει με το Θεώρημα II. Ουσιαστικά θέλουμε να πειραματιστούμε ως εξής: όταν η αρχική κατάσταση είναι η i επιλέγουμε μία απόφαση α και εν συνεχεία εφαρμόζουμε την πολιτική R. Τότε η (2.1.2) γίνεται

$$\mathrm{V_n}(i,\alpha,R) = C(i,\alpha) + \sum_{j\in S} p_{ij}(\alpha)\cdot \mathrm{V_{n-1}}(j,R)$$

και η (2.1.1) σε συνδυασμό με την (2.1.3) παίρνει την μορφή

$$C(i,\alpha) + \sum_{j\in S} p_{ij}(\alpha)\cdot \mathrm{V_{n-1}}(j,R) \cong n\mathrm{g}(R) + C(i,\alpha) - \mathrm{g}(R) + \sum_{j\in S} p_{ij}(\alpha)\cdot u_j(R).$$

Καθώς το n μεγαλώνει , για να ισχύει η ισότητα πρέπει να επιλέξουμε μία κατάλληλη απόφαση α=$\bar{R}_\iota$ που να ελαχιστοποιεί την ποσότητα:

$$C(i,\alpha) - \mathrm{g}(R) + \sum_{j\in S} p_{ij}(\alpha)\cdot u_j(R)$$

την καλούμενη και ποσότητα βελτίωσης της πολιτικής.

Θεώρημα II (εύρεσης βελτιωμένης πολιτικής):

Δίνονται g, u_i , $i \in$ S και πολιτική $\bar{R}$ τέτοια ώστε:

$$C(i,\bar{R}_\iota) - \mathrm{g} + \sum_{j\in S} p_{ij}(\bar{R}_\iota)\cdot u_j \le u_i\ ,\text{(II1)}$$

$i \in$ S . Τότε για τον ρυθμό κόστους g($\bar{R}$) που αντιστοιχεί στην πολιτική $\bar{R}$ ισχύει

$$\mathrm{g}(\bar{R}) \le \mathrm{g}\ \ \text{(II2)}.$$

Ομοίως, αν $C(i,\bar{R}_\iota) - \mathrm{g} + \sum_{j\in S} p_{ij}(\bar{R}_\iota)\cdot u_j \ge u_i$, (II3) , τότε g($\bar{R}$) $\ge$ g (II4).

Απόδειξη:

Η απόδειξη είναι παραπλήσια με αυτήν του Θεωρήματος I (B).

Αφού ισχύει η (ΙΙ1), με επαγωγή[45] αποδεικνύεται ότι

$$V_m(i,\bar{R}) - mg + \sum_{j\in S} p_{ij}{}^{(m)}(\bar{R})\, u_j \le u_i \quad \text{(ΙΙ5α)}$$

ή

$$V_m(i,\bar{R}) - u_i + \sum_{j\in S} p_{ij}{}^{(m)}(\bar{R})\, u_j \le mg \quad \text{(ΙΙ5β)}$$

Τελικά , διαιρώντας κατά μέλη με m και παίρνοντας όρια με m→∞ , κατ' αναλογία με το Θεώρημα ΙΙ (Β) έπεται η (ΙΙ2).

Απολύτως όμοια είναι και η απόδειξη της (ΙΙ4) δεδομένης της (ΙΙ3) , η οποία αφήνεται ως άσκηση στον αναγνώστη.

Διαισθητικά ,το Θεώρημα ΙΙ θα μπορούσε να εξηγηθεί ως εξής[46]:

Δεδομένου ενός κόστους ελέγχου (control cost) $C(i,a_i) - g$ που προκαλείται όταν ευρισκόμενοι στην κατάσταση i λαμβάνουμε την απόφαση a_i,

και ενός τελικού κόστους u_j που προκαλείται όταν ακολουθούμενης της πολιτικής $\bar{R}$, η αλυσίδα σταματήσει στην κατάσταση j ,

η (ΙΙ1) καταδεικνύει ότι μας συμφέρει να συνεχίσουμε στην κατάσταση j παρά να παραμείνουμε στην κατάσταση i . Αυτό θα συνεχίσει να ισχύει αν επαναλάβουμε την τακτική αυτή για m βήματα· οπότε προκύπτει η (ΙΙ5α) και αφού ο χώρος καταστάσεων είναι πεπερασμένος (το ίδιο θα συμβαίνει και με το πλήθος των στάσιμων πολιτικών[47]), έπεται τελικά η (ΙΙ2). Δηλαδή ο τελικός ρυθμός κόστους g($\bar{R}$) είναι μικρότερος από τον αρχικό g , που με την σειρά του συνεπάγεται βελτίωση της πολιτικής.

Συνεπώς , έχουμε βρεί το κριτήριο με το οποίο μπορούμε να βελτιώνουμε την πολιτική μας - αρκεί να λαμβάνουμε αποφάσεις $a_i = \bar{R}_\iota$ που ικανοποιούν την (ΙΙ1). Τότε, η νέα πολιτική θα έχει μικρότερο ή ίσο ρυθμό μέσου κόστους από την υπάρχουσα.

Εν συντομία, η μέθοδος βελτίωσης της πολιτικής συνίσταται στα εξής:

- Για να βελτιώσουμε μία δεδομένη πολιτική R, χρησιμοποιούμε το Θεώρημα Ι ώστε να βρούμε τις σχετικές τιμές $u_i(R)$, i∈S και τον μέσο ρυθμό κόστους g(R).
- Ακολούθως εφαρμόζουμε το Θεώρημα ΙΙ με g= g(R) και $u_i = u_i(R)$, i∈S .
- Κατ' αυτόν τον τρόπο ορίζουμε μία νέα πολιτική $\bar{R}$ για την οποία ισχύει $C(i,\bar{R}_\iota) - g + \sum_{j\in S} p_{ij}(\bar{R}_\iota)\cdot u_j \le u_i$, i∈S οπότε g($\bar{R}$) ≤ g , που σημαίνει ότι η $\bar{R}$ είναι καλύτερη ή (το λιγότερο) ισοδύναμη με την R.

Για κάθε i∈S,η απόφαση α=$\bar{R}_\iota$ επιλέγεται έτσι ώστε να ελαχιστοποιείται ως προς α∈A(i) η ποσότητα βελτίωσης της πολιτικής $C(i,\alpha) - g(R) + \sum_{j\in S} p_{ij}(\alpha)\cdot u_j(R)$ δηλαδή α=$\underset{\alpha\in A(i)}{\operatorname{argmin}}\{C(i,\alpha) - g(R) + \sum_{j\in S} p_{ij}(\alpha)\cdot u_j(R)\}$.

Σε αυτό το σημείο παραθέτουμε τον αλγόριθμο βελτίωσης της πολιτικής,ο οποίος κωδικοποιεί την παραπάνω διαδικασία.

---

[45] Η απόδειξη είναι πανομοιότυπη με αυτήν της σχέσης (Ι4)/(Ι5) του Θεωρήματος Ι.

[46] βλ.Henk C. Tijms, Stochastic Modelling and Analysis: A Computational Approach, 1986, σελ. 170.

[47]βλ. Δ. Φακίνου, «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», 2007, σελ. 366.

Αλγόριθμος Βελτίωσης της πολιτικής:

Βήμα 1ο : Ορίζουμε μία αρχική στάσιμη πολιτική $R = (R_i)_{i \in S}$.

Βήμα 2ο : Υπολογίζουμε τα g(R) , $u_i(R), i \in$S λύνοντας το γραμμικό σύστημα:

$$\Sigma : \begin{cases} u_i = C(i, R_i) - \mathrm{g} + \sum_{j \in S} p_{ij}(R_i) \cdot u_j \ , \ i \in \mathrm{S} \\ u_s = 0 \end{cases} .$$

Βήμα 3ο : Βελτιώνουμε την πολιτική,επιλέγοντας για κάθε i εκείνες τις αποφάσεις $\alpha_i = \bar{R}_i$ ώστε να ελαχιστοποιείται η ποσότητα βελτίωσης της πολιτικής:

$$\min_{\alpha \in A(i)} [\, C(i, a) - \mathrm{g}(R) + \sum_{j \in S} p_{ij}(a) \cdot u_j(R)] \ .$$

Η νέα στάσιμη πολιτική είναι η $\bar{R} = (\bar{R}_i)_{i \in S}$.

Βήμα 4ο : Αν $\bar{R} = R$, τότε ο αλγόριθμος τερματίζει και η R είναι η βέλτιστη πολιτική με ελάχιστο ρυθμό μέσου κόστους g(R) .

Αν $\bar{R} \neq R$, τότε πηγαίνουμε στο Βήμα 2 αντικαθιστώντας την R με την $\bar{R}$.

Τελικά , μετά την εφαρμογή του αλγορίθμου ουσιαστικά ισχύει

$$u_j(R^*) = \min_{\alpha \in A(i)} [\, C(i, a) - \mathrm{g}(R^*) + \sum_{j \in S} p_{ij}(a) \cdot u_j(R^*)] \quad (2.1.6)$$

με R^* την βέλτιστη πολιτική . Η (2.1.6) καλείται εξίσωση βελτιστοποίησης αφού από αυτήν προκύπτει η βέλτιστη πολιτική R^*.

2.2 Η ΜΕΘΟΔΟΣ ΕΠΑΝΑΛΗΠΤΙΚΗΣ ΤΙΜΗΣ

Η μέθοδος επαναληπτικής τιμής ή μέθοδος των επιτυχών προσεγγίσεων (Lloyd Shapley, 1953) βασίζεται στον δυναμικό προγραμματισμό. Ουσιαστικά με την μέθοδο αυτή υπολογίζεται σύμφωνα με έναν αναδρομικό τύπο η ακολουθία $\mathrm{V_n}(i)$ η οποία αποτελεί μία πολύ καλή προσέγγιση του ελάχιστου ρυθμού κόστους g(R). Αφετηρία της μεθόδου αυτής είναι η σχέση (2.1.2) με την διαφορά ότι στην συγκεκριμένη περίπτωση θέλουμε να επιλέξουμε την απόφαση α ώστε να ελαχιστοποιείται το μέσο συνολικό κόστος. Συνεπώς, αρχίζοντας από μία αρχική συνάρτηση $\mathrm{V_0}(i)$,$i\in$ S υπολογίζουμε αναδρομικά για n=1,2,… την

$$\mathrm{V_n}(i) = \min_{\alpha\in A(i)}\{\, C(i,\alpha) + \sum_{j\in S} p_{ij}(\alpha)\cdot \mathrm{V_{n-1}}(j)\},$$

για κάθε $i\in$ S.

Η ποσότητα $\mathrm{V_n}(i)$ εκφράζει το ελάχιστο μέσο συνολικό κόστος για τις n περιόδους που απομένουν,όταν η παρούσα κατάσταση είναι i και το σύστημά μας επιβαρύνεται με ένα τερματικό κόστος $\mathrm{V_0}(i)$, για να τερματίσει στην κατάσταση j. Υπό αυτήν την οπτική γωνία, η διαφορά $\mathrm{V_n}(i) - \mathrm{V_{n-1}}(j)$ προσεγγίζει σε βάθος χρόνου (για μεγάλο n) τον ελάχιστο ρυθμό μέσου κόστους . Με βάση αυτή την ιδέα προκύπτει το παρακάτω θεώρημα.

Θεώρημα ΙΙΙ (μεθόδου επαναληπτικής τιμής):

Υποθέτουμε ότι η Μαρκοβιανή αλυσίδα $\{X_n(R)\}_{n\in\mathbb{N}_0}$ έχει μοναδική κλειστή κλάση επικοινωνίας για κάθε στάσιμη πολιτική R (ή για κάθε βέλτιστη στάσιμη πολιτική[48]) και έστω στάσιμη πολιτική R(n) ∀ n∈ℕ της οποίας οι αποφάσεις R_i(n),$i\in$S, ελαχιστοποιούν το δεύτερο μέλος της ακολουθίας τιμών για κάθε $i\in$S:

$V_n(i) = \min_{a\in A(i)}\{C(i,a) + \sum_{j\in S} p_{ij}(a)\cdot V_{n-1}(j)\}$, (ΙΙΙ0) με $V_0(i)$ αυθαίρετη, για κάθε $i\in$S.

Ορίζουμε:

$\mathrm{g}^* = \min_R g(R)$: ο ελάχιστος ρυθμός κόστους

$m_n = \min_{j\in S}\{V_n(j) - V_{n-1}(j)\}$, n∈ℕ και

$M_n = \max_{j\in S}\{V_n(j) - V_{n-1}(j)\}$, n∈ℕ .

Τότε ισχύει

$$m_n \le \mathrm{g}^* \le \mathrm{g}\big(R(n)\big) \le M_n \qquad (i\in\mathrm{S},\ \mathrm{n}\in\mathbb{N}),$$

ενώ η ακολουθία $\{m_n\}_{n\in\mathbb{N}}$ είναι αύξουσα και η ακολουθία $\{M_n\}_{n\in\mathbb{N}}$ είναι φθίνουσα.

---

[48] Είναι μια ασθενέστερη συνθήκη η οποία ωστόσο μπορεί να υποτεθεί χωρίς να αλλάξει το αποτέλεσμα του Θεωρήματος ΙΙΙ. Αναλυτικότερα βλ. υποπαράγραφο 2.3.1 του παρόντος. βλ. Δ. Φακίνου, «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», 2007, σελ. 374.

Απόδειξη:

Εξ' ορισμού για την $V_n(i)$ ισχύει :

$V_n(i) \leq C(i,a) + \sum_{j\in S} p_{ij}(a) \cdot V_{n-1}(j)$, $a \in A(i)$ (III1) ,

ενώ από τον ορισμό της R(n) έπεται ότι :

$V_n(i) = C(i, \mathrm{R}_i(\mathrm{n})) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{n})) \cdot V_{n-1}(j)$ (III2),

i∈S.

Ωστόσο, $V_n(i) = V_{n-1}(i) + V_n(i) - V_{n-1}(i)$ (III3) ,

με $V_n(i) - V_{n-1}(i) \leq M_n$ (III4) και $V_n(i) - V_{n-1}(i) \geq m_n$ (III5), ∀ n∈ℕ.

Απόδειξη της ανισότητας g(R(n))≤ M_n:

Ισχύει:

$$(\mathrm{III2}) \xRightarrow{(\mathrm{III3}),(\mathrm{III4})} C(i, \mathrm{R}_i(\mathrm{n})) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{n})) \cdot V_{n-1}(j) \leq V_{n-1}(i) + M_n$$

$$C(i, \mathrm{R}_i(\mathrm{n})) - M_n + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{n})) \cdot V_{n-1}(j) \leq V_{n-1}(i)$$

Συνεπώς, ισχύει η σχέση (II1) και από το Θεώρημα II έπεται ότι g(R(n))≤ M_n .

Επίσης είναι προφανές ότι g^* ≤ g(R(n)) (εξ' ορισμού του g^* ως ελαχίστου ρυθμού κόστους).

Απόδειξη της ανισότητας $m_n \leq g^*$:

Από την (III1) αν επιλέξουμε απόφαση α=R_i και χρησιμοποιώντας τις σχέσεις (III3) και (III5) έχουμε :

$C(i, \mathrm{R}_i) + \sum_{j\in S} p_{ij}(\mathrm{R}_i) \cdot V_{n-1}(j) \geq V_n(i) = V_{n-1}(i) + V_n(i) - V_{n-1}(i)$

$C(i, \mathrm{R}_i) + \sum_{j\in S} p_{ij}(\mathrm{R}_i) \cdot V_{n-1}(j) \geq V_{n-1}(i) + m_n$

$C(i, \mathrm{R}_i) - m_n + \sum_{j\in S} p_{ij}(\mathrm{R}_i) \cdot V_{n-1}(j) \geq V_{n-1}(i)$

Επομένως, ισχύει η σχέση (II3) και από το Θεώρημα II έπεται ότι g(R)≥ m_n .

Όμως g(R)≥ m_n ⇒ $\min_R g(R) \geq m_n$ ⇒ $g^* \geq m_n$.

Απόδειξη ότι η $\{m_n\}_{n\in\mathbb{N}}$ είναι αύξουσα:

Αρκεί να δείξουμε ότι ισχύει: $m_k \leq m_{k+1}$, ∀ k∈ℕ .

Αν αντικαταστήσουμε στην (III1) όπου n=k και α=$\mathrm{R}_i(k+1)$ έχουμε :

$V_k(i) \leq C(i, \mathrm{R}_i(k+1)) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(k+1)) \cdot V_{k-1}(j)$ ή

$-V_k(i) \geq -C\big(i, \mathrm{R}_i(k+1)\big) - \sum_{j\in S} p_{ij}(\mathrm{R}_i(k+1)) \cdot V_{k-1}(j)$ (III6)

Ενώ για n=k+1 η (III2) γίνεται :

$V_{k+1}(i) = C(i, \mathrm{R}_i(\mathrm{k}+1)) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1)) \cdot V_k(j)$ (III7)

Προσθέτωντας στην (ΙΙΙ6) την (ΙΙΙ7) έχουμε:

$V_{k+1}(i) - V_k(i) \geq C(i, \mathrm{R}_i(\mathrm{k}+1)\,) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1)) \cdot V_k(j) \quad -$
$-C\big(i, \mathrm{R}_i(k+1)\big) - \sum_{j\in S} p_{ij}(\mathrm{R}_i(k+1)) \cdot V_{k-1}(j)$

$V_{k+1}(i) - V_k(i) \ \geq \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1)) \cdot \{V_k(j) - V_{k-1}(j)\}$

$V_{k+1}(i) - V_k(i) \geq \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1)) \cdot \min_{j'\in S}\{V_k(j') - V_{k-1}(j')\}$

$V_{k+1}(i) - V_k(i) \geq \min_{j'\in S}\{V_k(j') - V_{k-1}(j')\} \cdot \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1))$

$V_{k+1}(i) - V_k(i) \geq \min_{j\in S}\{V_k(j) - V_{k-1}(j)\}$,αφού $\sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k}+1)) = 1.$

$V_{k+1}(i) - V_k(i) \geq m_k$, άρα και $\min_{j\in S}\{\, V_{k+1}(i) - V_k(i)\} \geq \ m_k$,

δηλαδή $m_{k+1} \geq m_k$ για κάθε k∈ℕ (η επιλογή του k ήταν τυχαία).

Απόδειξη ότι η $\{M_n\}_{n\in\mathbb{N}}$ είναι φθίνουσα:

Αρκεί να δείξουμε ότι ισχύει: $M_{k+1} \leq M_k$, ∀ k∈ℕ .

Αν αντικαταστήσουμε στην (ΙΙΙ1) όπου n=k+1 και α=$\mathrm{R}_i(k)$ έχουμε :

$V_{k+1}(i) \ \leq C(i, \mathrm{R}_i(k)) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(k)) \cdot V_k(j)$ (ΙΙΙ8)

Ενώ για n=k η (ΙΙΙ2) γίνεται :

$V_k(i) = C(i, \mathrm{R}_i(\mathrm{k})\,) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k})) \cdot V_{k-1}(j)$

$-V_k(i) = -C\big(i, \mathrm{R}_i(\mathrm{k})\big) - \sum_{j\in S} p_{ij}(\mathrm{R}_i(\mathrm{k})) \cdot V_{k-1}(j)$ (ΙΙΙ9) .

Προσθέτωντας στην (ΙΙΙ8) την (ΙΙΙ9) έπεται ότι:

$V_{k+1}(i) - V_k(i) \ \leq C(i, \mathrm{R}_i(k)) + \sum_{j\in S} p_{ij}(\mathrm{R}_i(k)) \cdot V_k(j) - C\big(i, \mathrm{R}_i(k)\big) -$
$\sum_{j\in S} p_{ij}(\mathrm{R}_i(k)) \cdot V_{k-1}(j)$

$V_{k+1}(i) - V_k(i) \ \leq \sum_{j\in S} p_{ij}(\mathrm{R}_i(k)) \cdot \{V_k(j) - V_{k-1}(j)\}$

$V_{k+1}(i) - V_k(i) \ \leq \sum_{j\in S} p_{ij}(\mathrm{R}_i(k)) \cdot \max_{j'\in S}\{V_k(j') - V_{k-1}(j')\}$

$V_{k+1}(i) - V_k(i) \ \leq \max_{j'\in S}\{V_k(j') - V_{k-1}(j')\} \cdot \sum_{j\in S} p_{ij}(\mathrm{R}_i(k))$

$V_{k+1}(i) - V_k(i) \ \leq \max_{j'\in S}\{V_k(j') - V_{k-1}(j')\}$

$V_{k+1}(i) - V_k(i) \ \leq M_k$ και άρα $\max_{i\in S}\{\, V_{k+1}(i) - V_k(i)\} \leq M_k$.

δηλαδή $M_{k+1} \leq M_k$ για κάθε k∈ℕ και η απόδειξη ολοκληρώθηκε.

Από το Θεώρημα ΙΙΙ,θα προσπαθήσουμε να προσεγγίσουμε το μέσο ρυθμό κόστους g* από την ακολουθία $\mathrm{g}\big(R(n)\big)$ για *n* μεγάλο . Για να επιτύχουμε κάτι τέτοιο,

αρχικά υποθέτουμε ότι $C(i,\alpha) \geq 0$. Αν δεν συμβαίνει κάτι τέτοιο,προσθέτουμε κατάλληλη σταθερά ώστε να το επιτύχουμε − τότε ο μέσος ρυθμός κόστους μεταβάλλεται κατά αυτήν την σταθερά C, αφού

$$g'(R)= \lim_{m\to\infty} \frac{V'_m(i,R)}{m} = \lim_{m\to\infty} \frac{\sum_{k=0}^{m-1}\sum_{j\in S} p_{ij}^{(k)}(R)[C(j,R_j)+C]}{m} =$$

$$\lim_{m\to\infty} \frac{mC+\sum_{k=0}^{m-1}\sum_{j\in S} p_{ij}^{(k)}(R)C(j,R_j)}{m} = g(R) + C.$$

Επιλέγουμε $V_0(i) : 0 \leq V_0(i) \leq \min_{a\in A(i)} C(i,\alpha)\ \ \forall\, i \in S$.

Τότε ,

$$C(i,a) + \sum_{j\in S} p_{ij}(a)\cdot V_0(j) \geq V_0(i)\ \ \forall\, i \in S, a \in A(i)$$

και $$V_1(i) = \min_{a\in A(i)}\{C(i,a) + \sum_{j\in S} p_{ij}(a)\cdot V_0(j)\} \geq V_0(i)$$

δηλαδή $V_1(i) \geq V_0(i) \quad \forall\, i \in S$

$$V_1(i) - V_0(i) \geq 0\ \forall\, i \in S$$

οπότε $m_0 = \min_{i\in S}\{V_1(i) - V_0(i)\} \geq 0$ και αφού η m_k είναι αύξουσα ισχύει

$$m_k \geq 0\ \forall\, k \in \mathbb{N}.$$

Συνεπώς, έχουμε πλέον την ελαφρά αλλά σημαντικά αλλαγμένη σχέση:

$$0 \leq m_n \leq g^* \leq g(R(n)) \leq M_n \qquad (i\in S,\ n\in\mathbb{N})\ \text{(III)}.$$

Επιπλέον υποθέτουμε ότι η μεταβολή $\frac{M_n - m_n}{m_n} \leq \varepsilon$ (*) είναι μικρότερη από ένα θετικό ε , το οποίο μπορεί να είναι οσοδήποτε μικρό.

Τότε (*) $\Leftrightarrow M_n \leq (\varepsilon+1)m_n$ και από την (III) έχουμε :

$$0 \leq g^* \leq g(R(n)) \leq (\varepsilon+1)m_n$$

$$0 \leq \frac{g(R(n)) - g^*}{g^*} \leq \frac{(\varepsilon+1)m_n - g^*}{g^*} = (\varepsilon+1)\frac{m_n}{g^*} - 1 \leq \varepsilon$$

με την τελευταία ανισότητα να ισχύει καθώς από την (III) προκύπτει $\frac{m_n}{g^*} \leq 1$

και άρα $(\varepsilon+1)\frac{m_n}{g^*} \leq \varepsilon + 1 \Leftrightarrow (\varepsilon+1)\frac{m_n}{g^*} - 1 \leq \varepsilon$.

Τελικά , αποδείχθηκε ότι αν $\frac{M_n - m_n}{m_n} \leq \varepsilon$ τότε $0 \leq \frac{g(R(n)) - g^*}{g^*} \leq \varepsilon$.

Αφού $g(R(n))$ ο ρυθμός κόστους στην ν-οστή επανάληψη και g^* ο ελάχιστος ρυθμός κόστους,συμπεραίνουμε ότι αν η μεταβολή της m_n από την M_n είναι μικρότερη από ένα θετικό ε,τότε και η μεταβολή του g^* από τον $g(R(n))$ είναι το πολύ της τάξεως του 100ε% . Ισοδύναμα, αν το ε είναι πολύ μικρό έχουμε μία πολύ καλή προσέγγιση του μέσου ρυθμού κόστους g^* από την $g(R(n))$.

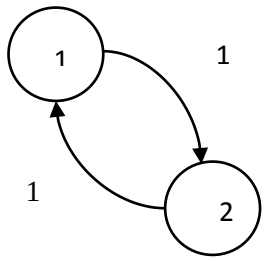
Φαίνεται λογικό να θεωρήσουμε ότι οι ακολουθίες m_n,M_n συγκλίνουν και μάλιστα στον ίδιο αριθμό. Σε αυτήν την περίπτωση,η (ΙΙΙ) δίνει από το κριτήριο παρεμβολής $\lim_{n\to\infty} g(R(n)) = g^*$ και παίρνουμε άμεσα ότι η επανάληψη της (ΙΙΙ) για κάποιο (ίσως μεγάλο) αριθμό βημάτων n δίνει μια καλή προσέγγιση του μέσου ρυθμού κόστους. Δυστυχώς,το θέμα δεν είναι τόσο απλό – οι ακολουθίες δεν συγκλίνουν απαραίτητα στον ίδιο αριθμό. Το επόμενο παράδειγμα το επιβεβαιώνει.

Παράδειγμα 2.1:

Δίνεται η μαρκοβιανή διαδικασία αποφάσεων $\{X_n(R)\}_{n\in\mathbb{N}_0}$ με χώρο καταστάσεων S={1,2},χώρο δυνατών αποφάσεων A={a_0},κόστη αποφάσεων C(1, a_0)=1, C(2, a_0)=0 και πίνακα πιθανοτήτων μετάβασης πρώτης τάξης

$$\mathbb{P}(\mathrm{R})= \mathbb{P}(a_0) = \begin{array}{cc} \begin{array}{cc} 1 & 2 \end{array} & \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \begin{array}{c} 1 \\ 2 \end{array} \end{array}$$

ενώ στο σχήμα 2.1 φαίνεται ο γράφος[49] των δυνατών μεταβάσεων της αλυσίδας.



Σχήμα 2.1: Ο γράφος των δυνατών μεταβάσεων για την αλυσίδα κάτω από την μοναδική απόφαση a_0 του παραδείγματος 2.1

Στην συγκεκριμένη περίπτωση,η αλυσίδα έχει μία μοναδική κλειστή κλάση (για την ακρίβεια είναι μη-υποβιβάσιμη) κάτω από την μοναδική απόφαση a_0 .

Τότε, ορίζουμε αυθαίρετα $V_n(i) = 0$, και

$$V_n(i) = \min_{a\in A(i)}\{C(i,a) + \sum_{j\in S} p_{ij}(a)\cdot V_{n-1}(j)\}\,, i = 1,2$$

οπότε :

$V_1(1) = C(1,a_0) + 0 = 1 \quad , V_1(2) = C(2,a_0) + 0 = 0$

$V_2(1) = C(1,a_0) + \sum_{j\in S} p_{1j}(a_0)\cdot V_1(j) = 1 + p_{12}(a_0)\cdot V_1(2) = 1$

$V_2(2) = C(2,a_0) + \sum_{j\in S} p_{2j}(a_0)\cdot V_1(j) = p_{21}(a_0)\cdot V_1(1) = 1$

[49] Βλ. Μ. Λουλάκη(2015) «Στοχαστικές Διαδικασίες», σελ. 19-20.

$V_3(1) = C(1, a_0) + \sum_{j \in S} p_{1j}(a_0) \cdot V_2(j) = 1 + p_{12}(a_0) \cdot V_2(2) = 1 + 1 = 2$

$V_3(2) = 0 + p_{21}(a_0) \cdot V_2(1) = 1$

$\vdots$

και τελικά $\forall\, k \in \mathbb{N}$ ισχύει:

$V_{2k}(1) = k$, $k \geq 0$ $\qquad$ $V_{2k}(2) = k$, $k \geq 0$

$V_{2k-1}(1) = k$, $k \geq 1$ $\qquad$ $V_{2k-1}(2) = k - 1$, $k \geq 1$

και

$$m_n = \min_{j \in S}\{V_n(j) - V_{n-1}(j)\} = V_n(1) - V_{n-1}(1) = 0 \xrightarrow{n \to \infty} 0$$

$$M_n = \max_{j \in S}\{V_n(j) - V_{n-1}(j)\} = V_n(2) - V_{n-1}(2) = 1 \xrightarrow{n \to \infty} 1$$

και ως εκ τούτου οι ακολουθίες m_n , M_n δεν συγκλίνουν στο ίδιο όριο.

Στην προκείμενη περίπτωση, γίνεται εύκολα αντιληπτό ότι η αλυσίδα είναι περιοδική με περίοδο δύο (2) – αυτό είναι και ένα από τα πιο συνηθισμένα εμπόδια στα περισσότερα θεωρήματα συγκλίσεων στις στοχαστικές ανελίξεις[50] (και όχι μόνο) .

Γενικά, αν υποτεθεί η ιδιότητα της απεριοδικότητας για το σύστημά μας , τότε οι m_n, M_n συγκλίνουν στο ίδιο όριο [από την (ΙΙΙ) αυτό θα είναι το g^*] και άρα η κεντρική ιδέα της προσέγγισης του αναμενόμενου ρυθμού κόστους από τις m_n και M_n λειτουργεί. Ενδεικτικά ισχύει το παρακάτω θεώρημα[51].

Θεώρημα ΙΙΙ-β:

Υποθέτουμε ότι για κάθε βέλτιστη στάσιμη πολιτική R η μαρκοβιανή αλυσίδα $\{X_n(\mathrm{R})\}_{n \in \mathbb{N}_0}$ έχει μοναδική κλειστή κλάση επικοινωνίας και είναι απεριοδική. Αν g^* ο ελάχιστος ρυθμός μέσου κόστους, τότε υπάρχουν α>0 και 0<β<1 :

i) $|M_n - m_n| \leq \alpha\beta^n$, $n \in \mathbb{N}$ και

ii) $\lim_{n \to \infty} M_n = g^* = \lim_{n \to \infty} m_n$.

Ωστόσο, με τις συγκλίσεις των μεθόδων θα ασχοληθούμε σε επόμενο κεφάλαιο.

Η υπόθεση της απεριοδικότητας στο προηγούμενο θεώρημα είναι ρεαλιστική, αφού μία περιοδική μαρκοβιανή αλυσίδα μπορεί να μετασχηματιστεί σε μία απεριοδική με την ίδια στάσιμη κατανομή και άρα τον ίδιο ρυθμό μέσου κόστους .

[50] Βλ. Μ. Λουλάκη(2015) «Στοχαστικές Διαδικασίες», σελ. 87-89.

[51] Βλ. Tijms, H.C. (1994). Stochastic Models, An Algorithmic Approach, John Wiley and Sons, New York, σελ. 190-192.

Κάτι τέτοιο επιτυγχάνεται[52] αν η νέα μαρκοβιανή αλυσίδα $\overline{X_n}$ έχει πιθανότητες μετάβασης $\bar{p}_{ij}=\begin{cases} \tau p_{ij}\,,\ i \neq j \\ \tau p_{ii} + 1 - \tau\,,\ i = j \end{cases}$,0< τ <1. Πράγματι, $\overline{p_{ii}} > 0$ και η $\overline{X_n}$ είναι απεριοδική, ενώ για να βρούμε την στάσιμη κατανομή της αρκεί να λύσουμε το σύστημα (1.1.5):

$$\pi_j = \sum_{i\in S} \pi_i\, \bar{p}_{ij} \Leftrightarrow \pi_j - \pi_j \bar{p}_{jj} = \sum_{i\neq j} \pi_i\, \bar{p}_{ij} \Leftrightarrow \pi_j(\tau - \tau p_{jj}) = \sum_{i\neq j} \pi_i\, \tau p_{ij} \Leftrightarrow$$

$$\pi_j = \pi_j \tau p_{jj} + \sum_{i\neq j} \pi_i\, \tau p_{ij} \overset{\tau>0}{\Longleftrightarrow} \pi_j = \sum_{i\in S} \pi_i\, p_{ij}, \forall\, j \in S,$$

που είναι ακριβώς το σύστημα (1.1.5) για την αρχική κατάσταση X_n.

Κωδικοποιούμε το Θεώρημα ΙΙΙ και παίρνουμε τον παρακάτω αλγόριθμο:

Αλγόριθμος επαναληπτικής τιμής:

Βήμα 1ο : Για κάθε i ∈ S επιλέγουμε αυθαίρετα $V_0(i)$ τέτοιο ώστε

$$0 \leq V_0(i) \leq minC(i,a)\ ,\ \forall\ i \in S\ .$$

Επίσης n=1 και επιλέγω ε > 0 την επιθυμητή ακρίβεια.

Βήμα 2ο : Υπολογίζουμε τις τιμές $V_n(i) = \min_{a\in A(i)} \{C(i,a) + \sum_{j\in S} p_{ij}(a) \cdot V_{n-1}(i)\}$

και ορίζουμε την αντίστοιχη πολιτική R(n)= $\{R_i(n)\}_{i\in S}$, οι αποφάσεις της οποίας ελαχιστοποιούν το 2ο μέλος της παραπάνω ισότητας.

Βήμα 3ο : Υπολογίζουμε τα φράγματα $m_n = \min_{j\in S}\{V_n(j) - V_{n-1}(j)\}$ και

$M_n = \max_{j\in S}\{V_n(j) - V_{n-1}(j)\}$. Ο αλγόριθμος τερματίζει με την στάσιμη πολιτική R(n) όταν $0 \leq M_n - m_n \leq m_n \cdot \varepsilon$. Ο ρυθμός μέσου κόστους g(R) δίνεται προσεγγιστικά από τις ποσότητες M_n, m_n.

Αν ο αλγόριθμος δεν τερματίσει, πηγαίνουμε στο Βήμα 4.

Βήμα 4ο : Θέτουμε n:=n+1 και πηγαίνουμε στο Βήμα 2.

Αξίζει να σημειωθεί ότι με την μέθοδο επαναληπτικής τιμής υπολογίζεται μία σχεδόν βέλτιστη πολιτική , κι αυτό λόγω της προσέγγισης του ελάχιστου ρυθμού μέσου κόστους από τον αλγόριθμο. Βέβαια, η επιλογή ενός πολύ μικρού ε>0 καθιστά την προσέγγιση αυτή όσο καλή θέλουμε.

[52]Βλ. Φακίνου Δ.(2007). «Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα», σελ. 460.

2.3 Η ΕΠΙΛΥΣΗ ΜΕ ΓΡΑΜΜΙΚΟ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ

2.3.1 Η περίπτωση του απλού μοντέλου

Στις προηγούμενες παραγράφους η εύρεση βέλτιστης πολιτικής για την ελαχιστοποίηση του μέσου κόστους επιλύθηκε με την μέθοδο βελτίωσης πολιτικής ή την μέθοδο επαναληπτικής τιμής. Ωστόσο, αυτές οι μέθοδοι δεν είναι οι μοναδικές για την αντιμετώπιση του εν λόγω προβλήματος.Κι αυτό, διότι η εξίσωση βελτιστοποίησης δεν λύνεται μόνο με τις παραπάνω μεθόδους. Ένας τρίτος τρόπος (De Ghellinck και Manne, 1960) είναι η θεώρησή του ως πρόβλημα γραμμικού προγραμματισμού (π.γ.π). Το πρώτο βήμα είναι η μορφοποίηση του π.γ.π που φαίνεται παρακάτω[53]:

Μεταβλητές απόφασης: x_{ia} , $i \in S, a \in A(i)$

Αντικειμενική συνάρτηση:

$$min\, z = min \sum_{i \in S} \sum_{a \in A(i)} c_i(a) x_{ia}$$

Περιορισμοί:

$$\sum_{a \in A(j)} x_{ja} - \sum_{i \in S} \sum_{a \in A(i)} p_{ij}(a) x_{ia} = 0 \,, j \in S$$

$$\sum_{i \in S} \sum_{a \in A(i)} x_{ia} = 1$$

$$x_{ia} \geq 0 \,, i \in S, a \in A(i) \,.$$

Χάριν συντομίας,θα ονομάσουμε την παραπάνω μορφοποίηση ως (π.γ.π-1),ενώ η πρώτη ομάδα περιορισμών θα καλείται "πρώτος περιορισμός". Πριν περάσουμε στο επόμενο βήμα είναι απαραίτητο να εξηγήσουμε τους παραπάνω συμβολισμούς.

Κατ' αρχάς, οι μεταβλητές απόφασεις x_{ia} εκφράζουν το ποσοστό σε βάθος χρόνου των χρόνων απόφασης που το σύστημα (η μαρκοβιανή αλυσίδα) βρίσκεται στην κατάσταση $i \in S$ κάτω από την απόφαση $a \in A(i)$. Συνεπώς, το διπλό άθροισμα

$$\sum_{i \in S} \sum_{a \in A(i)} c_i(a) x_{ia}$$

εκφράζει το συνολικό μέσο ρυθμό κόστους από το γεγονός ότι η αλυσίδα βρέθηκε σε όλες τις δυνατές καταστάσεις κάτω από όλες τις δυνατές αποφάσεις. Ως εκ τούτου ο περιορισμός της μη-αρνητικότητας είναι προφανής,ενώ το ίδιο ισχύει και για τον δεύτερο περιορισμό (το σύνολο των ποσοστών οφείλει να αθροίζει στην μονάδα).

Τέλος, ο πρώτος περιορισμός γράφεται και ως εξής:

[53] Tijms, H.C. (1986). Stochastic Modeling and Analysis, A Computational Approach, σελ. 181.

$$\sum_{a \in A(j)} x_{ja} = \sum_{i \in S} \sum_{a \in A(i)} p_{ij}(a) x_{ia}$$

για κάθε $j \in S$, που σημαίνει ότι σε βάθος χρόνου ο μέσος ρυθμός μετακινήσεων από την j πρέπει να είναι ίσος με τον μέσο ρυθμό μετακινήσεων προς την j, για κάθε $j \in S$.

Αλγόριθμος γραμμικού προγραμματισμού (LPA[54]-1)

Βήμα 1ο : Εύρεση μιας βέλτιστης βασικής λύσης $\{x_{ia}^*\}$ του (π.γ.π-1) με χρήση της μεθόδου Simplex .

Βήμα 2ο : Για κάθε $i \in S^*$ θέτουμε $R_i^* :=$ α με α: $x_{ia}^* > 0$,

όπου

$$S^* = \{i | \sum_{a \in A(i)} x_{ia}^* > 0\}$$

μη-κενό σύνολο.

Βήμα 3ο : Αν $S=S^*$,η βέλτιστη πολιτική μέσου ρυθμού κόστους είναι η $R^*=\{R_i^*\}$ και ο αλγόριθμος τερματίζει.

Αλλιώς,βρίσκουμε $i \notin S^*$ και κατάσταση $a \in A(i)$ τέτοια ώστε $p_{ij}(a) > 0$ για κάποιο $j \in S^*$,θέτουμε $R_i^* :=$α και $S^* := S^* \cup \{i\}$ και επαναλαμβάνουμε το 3ο βήμα.

Προτού σκιαγραφηθεί η απόδειξη ότι ο παραπάνω αλγόριθμος επιλύει την εξίσωση βελτιστοποίησης,πρέπει να αναφέρουμε ότι με την μέθοδο του γραμμικού προγραμματισμού,επιτρέπεται η χαλάρωση της υπόθεσης ότι *για κάθε* πολιτική R, η αλυσίδα $\{X_n(R)\}_{n \in \mathbb{N}_0}$ έχει μοναδική κλειστή κλάση ή σύνολο. Εν προκειμένω, υποθέτουμε ότι *για κάθε βέλτιστη* πολιτική R,η $\{X_n(R)\}_{n \in \mathbb{N}_0}$ έχει μοναδική κλειστή κλάση ή σύνολο[55].Αυτό συνεπάγεται ότι για κάθε βέλτιστη πολιτική R υπάρχουν καταστάσεις $i \in S$ που είναι επισκέψιμες από κάθε άλλη κατάσταση $j \in S$. Λόγω του πεπερασμένου χώρου καταστάσεων S, το μοναδικό κλειστό σύνολο περιέχει όλες τις γνησίως επαναληπτικές καταστάσεις[56], οι οποίες και είναι τελικά επισκέψιμες από κάθε άλλη κατάσταση[57]-οι υπόλοιπες καταστάσεις είναι παροδικές. Επισημαίνουμε ότι πλέον δεν απαιτούμε τα παραπάνω να ισχύουν απαραίτητα για κάθε πολιτική R,αλλά μόνο για τις βέλτιστες πολιτικές-που σημαίνει ότι η υπόθεση είναι πιο ασθενής από την αρχική. Όσον αφορά τις μη-βέλτιστες πολιτικές, αυτές επιτρέπεται να ορίζουν και πολλαπλές κλειστές κλάσεις οι οποίες θα είναι ξένες μεταξύ τους. Βέβαια, αν αλλάξει η πολιτική τότε ενδεχομένως να αλλάξουν και οι ανοικτές ή και οι κλειστές κλάσεις της νέας αλυσίδας $\{X_n(R)\}_{n \in \mathbb{N}_0}$. Ωστόσο, αν η R είναι βέλτιστη, η κλειστή κλάση θα είναι αναγκαστικά μοναδική.

[54] LPA: Linear Programming Algorithm.
[55] Tijms, H.C. (1986). Stochastic Modeling and Analysis, A Computational Approach, σελ. 180-181.
[56] Λουλάκης Μ. (2015). Στοχαστικές Διαδικασίες, σελ. 75-76.
[57] Λουλάκης Μ. (2015). Στοχαστικές Διαδικασίες, σελ. 26-27.

Κάτω από την προναφερθείσα υπόθεση,η εξίσωση βελτιστοποίησης δίνει μοναδική λύση τον ελάχιστο αναμενόμενο ρυθμό κόστους,ο οποίος μάλιστα προκύπτει ανεξάρτητος[58] της αρχικής κατάστασης i.

Θεώρημα IV:

Έστω η μαρκοβιανή αλυσίδα $\{X_n\}_{n\in\mathbb{N}_0}$,για την οποία ισχύει ότι για κάθε βέλτιστη πολιτική R, η $\{X_n(R)\}_{n\in\mathbb{N}_0}$έχει μοναδική κλειστή κλάση ή σύνολο. Τότε ο αλγόριθμος LPA-1 που αναφέρθηκε παραπάνω δίνει μία πολιτική που είναι βέλτιστη ως προς τον μέσο ρυθμό κόστους.

Απόδειξη:

Αρχικά υποθέτουμε ότι η εξίσωση βελτιστοποίησης έχει λύση την $(g^*, \underline{u_i^*})$, $i\in S$. Δηλαδή ισχύει

$u_i^* = \min_{\alpha\in A(i)}\{c_i(a) - g^* + \sum_{j\in S} p_{ij}(a)u_j^*\}$ (IV1), $i\in S$ και

$g^* = \min_R g(R)$ (IV2)

όπου βέλτιστη είναι κάθε στάσιμη πολιτική R^* τέτοια ώστε για κάθε $i\in S$ η δραστηριότητα R_i ελαχιστοποιεί το δεύτερο μέλος της (IV1).

Από την (IV1) έπεται ότι οι ανισώσεις

$$u_i \le c_i(a) - g + \sum_{j\in S} p_{ij}(a)u_j \ , \forall\, i \in S \text{ και } a \in A(i)$$

έχουν λύση (IV3). Οπότε χρησιμοποιώντας το Θεώρημα II για κάθε $i \in S$ και $\alpha = R_i$ από την (II3) έπεται ότι ισχύει η (II4),δηλαδή:

$$g_i(R) \ge g \ \ \forall\, i \in S,$$

με $g_i(R)$ τον μέσο ρυθμό κόστους σε βάθος χρόνου κάτω από την στάσιμη πολιτική R και με αρχική κατάσταση i.

Τότε $\min_i g_i(R) \ge g$ και λόγω της (IV2) έχουμε $g^* \ge g$ ή $g \le g^*$ (IV4).

Με λίγα λόγια,οι (IV3) και οι (IV4) καταδεικνύουν ότι για να λύσουμε τις (IV1)-(IV2)αρκεί να βρούμε την μέγιστη τιμή g τέτοια ώστε οι ανισώσεις

$$u_i \le c_i(a) - g + \sum_{j\in S} p_{ij}(a)u_j \ , \forall\, i \in S \text{ και } a \in A(i)$$

να έχουν λύση.

Με λίγα λόγια αρκεί να λύσουμε το (π.γ.π-2) :

[58] Για απόδειξη βλ. Denardo and Fox (1968). “Multichain Markov renewal programs”, SIAM Journal of Applied Mathematics, 16, 468-487.

max z = max g

με περιορισμούς :

$$g + u_i - \sum_{j \in S} p_{ij}(a) u_j \leq c_i(a) \text{ , } \forall\, i \in S \text{ και } a \in A(i).$$

g , u_i : χωρίς περιορισμό.

Παρατηρούμε ότι το (π.γ.π-2) είναι το δυϊκό πρόβλημα του (π.γ.π-1). Πράγματι, το πρωτεύον (π.γ.π-1) είναι πρόβλημα ελαχιστοποίησης , άρα το δυϊκό θα είναι πρόβλημα μεγιστοποίησης. Στους περιορισμούς του (π.γ.π-1) αντιστοιχίζουμε τις μεταβλητές απόφασης του δυϊκού όπως φαίνεται παρακάτω:

Πίνακας 1: Αντιστοιχία πρωτεύοντος (π.γ.π-1) και δυϊκού (π.γ.π-2) .

Περιορισμός πρωτεύοντος	Μεταβλητή δυϊκού	Συντελεστής αντικειμενικής συνάρτησης δυϊκού
$\sum_{a \in A(j)} x_{ja} - \sum_{i \in S} \sum_{a \in A(i)} p_{ij}(a) x_{ia} = 0$, $j \in$ S	u_j , $j \in$ S	0 $j \in$ S
$\sum_{i \in S} \sum_{a \in A(i)} x_{ia} = 1$	g	1

Άρα η αντικειμενική συνάρτηση του δυϊκού είναι z =1·g + 0 · $\sum_{j \in S} u_j$= g και

αφού έχουμε πρόβλημα μεγιστοποίησης,οι περιορισμοί θα έχουν την μορφή ανισώσεων με φορά $\leq$ $c_i(a)$ $\forall\, i \in$ S και $a \in A(i)$. Αφήνεται ως άσκηση στον αναγνώστη η απόδειξη ότι οι περιορισμοί του (π.γ.π-2) παίρνουν τελικά την μορφή:

$$g + u_i - \sum_{j \in S} p_{ij}(a) u_j \leq c_i(a)$$

$\forall\, i \in$ S και $a \in A(i)$.

Εφόσον τα (π.γ.π-1) και (π.γ.π-2) έχουν βέλτιστη βασική εφικτή λύση, θα έχουν την ίδια βέλτιστη τιμή g^* της αντικειμενικής συνάρτησης που αποτελεί και τον ελάχιστο μέσο ρυθμό κόστους. Η σχέση αυτή (δυϊκού-πρωτεύοντος) φαινομενικά ελάχιστα σχετική με την απόδειξη, θα μας χρειαστεί στη συνέχεια όπως θα δούμε παρακάτω.

Ο αναγνώστης ενδεχομένως να έχει την εύλογη απορία κατά πόσον μία βέλτιστη βασική εφικτή λύση $\{x_{ia}^*\}$ του αλγορίθμου LP-1 ορίζει μία βέλτιστη πολιτική μέσου ρυθμού κόστους. Κάτι τετοιο είναι αληθές και η απόδειξή του θα γίνει σε τρία βήματα.

Βήμα 1ο : Το σύνολο $S^* = \{i | \sum_{a \in A(i)} x_{ia}^* > 0\}$ είναι κλειστό (κλειστό σύνολο επικοινωνίας) κάτω από οποιαδήποτε στάσιμη πολιτική R.

Απόδειξη:

Ο δεύτερος περιορισμός του (π.γ.π-1) εξασφαλίζει ότι το σύνολο S^* είναι μη-κενό.

Έστω ότι υπάρχει κατάσταση $i \in S^*$ και απόφαση $a' \in A(i)$, τέτοια ώστε $p_{ij}(a') > 0$ για κάποια $j \notin S^*$. Αν δεν ισχύει κάτι τέτοιο,τότε δεν έχουμε τίποτα να αποδείξουμε. Επιλέγουμε την κατάσταση $i \in S^*$ ώστε $x^*_{ia'} > 0$. Είναι εύλογο να αναρωτηθούμε αν έχουμε την δυνατότητα μιας τέτοιας επιλογής. Ενδεχομένως για την κατάσταση i να ισχύει μεν $p_{ij}(a') > 0$ αλλά όχι και $x^*_{ia'} > 0$. Ωστόσο και σε αυτήν την περίπτωση, μπορούμε να βρούμε μία κατάσταση $l \in S^*$ τέτοια ώστε $x^*_{la''} > 0$ και $p_{li}(a'') > 0$. Αν η j είναι προσβάσιμη από την i και η i είναι προσβάσιμη από την l ,τότε και η j είναι προσβάσιμη από την l [59]. Ας βρούμε μία τέτοια κατάσταση $l \in S^*$:

Από τον πρώτο περιορισμό του (π.γ.π-1),βάζοντας όπου j το $i \in S^*$ και όπου i το $l \in S^*$, λαμβάνοντας υπόψη ότι $x_{la} = 0$ για κάθε $l \notin S^*$ και $a \in A(l)$ έχουμε :

$$\sum_{l \in S^*} \sum_{a \in A(l)} p_{li}(a) x^*_{la} = \sum_{a \in A(i)} x^*_{ia} > 0$$

Επομένως υπάρχει $l \in S^*$ και $a \in A(l)$ τέτοια ώστε $p_{li}(a) x^*_{la} > 0$ δηλαδή η l έχει πρόσβαση στην $j \notin S^*$ και επιπλέον $x^*_{la} > 0$.

Τέλος, χρησιμοποιώντας ξανά τον πρώτο περιορισμό του (π.γ.π-1) για το $l \in S^*$ και το $j \notin S^*$ έχουμε :

$$\sum_{a \in A(j)} x^*_{ja} = \sum_{l \in S} \sum_{a \in A(l)} p_{lj}(a) x^*_{la} > 0$$

που συνεπάγεται $j \in S^*$: Άτοπο. Άρα , η κλάση S^* είναι κλειστή.

Βήμα 2ο : Το 3ο βήμα του αλγορίθμου LP-1 ολοκληρώνεται σε πεπερασμένα βήματα. Αυτό σημαίνει ότι μπορούμε να βρούμε μια πολιτική R^*.

Απόδειξη:

Υπενθυμίζουμε ότι από την υπόθεση ισχύει ότι *για κάθε βέλτιστη* πολιτική R, η $\{X_n(R)\}_{n \in \mathbb{N}_0}$ έχει μοναδική κλειστή κλάση . Ωστόσο, από το 1ο βήμα έπεται ότι το σύνολο S^*είναι κλειστό , άρα είναι και το μοναδικό κλειστό σύνολο και αφού είναι πεπερασμένο λόγω του πεπερασμένου χώρου καταστάσεων S , όλες οι επαναληπτικές καταστάσεις θα ανήκουν απαραίτητα σ' αυτό. Άρα αναγκαστικά αν $i \notin S^*$, η κατάσταση i είναι παροδική κάτω από μια οποιαδήποτε βέλτιστη πολιτική μέσου ρυθμού κόστους. Συνεπώς αν υπάρχει παροδική κατάσταση i αυτή δεν θα ανήκει στην S^*, άρα μπορούμε να επιλέξουμε $i \notin S^*$ και κατάσταση $a \in A(i)$ τέτοια ώστε $p_{ij}(a) > 0$ για κάποιο $j \in S^*$και να ολοκληρώσουμε το βήμα 3.Αν δεν υπάρχει παροδική κατάσταση, όλες οι καταστάσεις είναι επαναληπτικές. Επομένως, S^*=S και ο αλγόριθμος τερματίζει άμεσα.

[59] Λουλάκης Μ. (2015) , Στοχαστικές Διαδικασίες , σελ. 18.

Βήμα 3ο : Η πολιτική R^* που κατασκευάστηκε με τον τερματισμό του αλγορίθμου είναι βέλτιστη ως προς τον μέσο ρυθμό κόστους.

Απόδειξη:

Έστω $\{x_{ia}^*\}$, $i\in$ S, α=$R_i{}^*$,η βέλτιστη βασική (εφικτή) λύση του πρωτεύοντος (π.γ.π-1) και (g^*,$\underline{u_i^*}$), $i\in$ S η αντίστοιχη λύση του δυϊκού (π.γ.π-2).

Τότε από το Θεώρημα συμπληρωματικότητας [60] για τα δύο προβλήματα, έπεται ότι για κάθε $i\in S^*$ ισχύει:

$$[c_i(R_i^*) - (\,g^* + u_i^* - \sum_{j\in S} p_{ij}(R_i^*)u_j^*)]x_{ia}^* = 0$$

με α=R_i^* να έχει επιλεχθεί έτσι ώστε $x_{ia}^* > 0$ (βήμα 2), οπότε έχουμε για κάθε $i\in S^*$:

$$c_i(R_i^*) - (\,g^* + u_i^* - \sum_{j\in S} p_{ij}(R_i^*)u_j^*) = 0$$

$$c_i(R_i^*) = \,g^* + u_i^* - \sum_{j\in S} p_{ij}(R_i^*)u_j^*.$$

Αν η αρχική κατάσταση $i\in S^*$ είναι επαναληπτική,τότε η αλυσίδα θα μείνει σε επαναληπτικές καταστάσεις και συνεπώς $j \in S^*$: επαναληπτική (προφανώς κάτω από την R^*).Συμβολίζοντας ως $S(R^*)$ το σύνολο των επαναληπτικών καταστάσεων του S κάτω από την πολιτική R^*, έχουμε την (IV5):

$$c_i(R_i^*) = \,g^* + u_i^* - \sum_{j\in S(R^*)} p_{ij}(R_i^*)u_j^*$$

για κάθε $i\in S^*$.

Συνεχίζοντας όπως στην απόδειξη του Θεωρήματος I (B) , η (IV5) δίνει ότι η ποσότητα g^* είναι ο ελάχιστος ρυθμός μέσου κόστους κάτω από την πολιτική R^* για κάθε αρχική κατάσταση i που είναι επαναληπτική. Άρα η R^* είναι βέλτιστη .

Στην περίπτωση που η αλυσίδα μας αρχίσει από παροδική κατάσταση , τότε μετά από πεπερασμένα βήματα θα βρεθεί σε επαναληπτική κατάσταση[61],οπότε θα παραμείνει για πάντα στην $S(R^*)$ και η (IV5) συνεχίζει να ισχύει και σε αυτήν την περίπτωση. Επομένως, ισχύουν απαράλλακτα τα προαναφερθέντα και η απόδειξη ολοκληρώθηκε.

[60] Complementary slackness property (Linear Programming Theory)
βλ. Φακίνου Δ. Οικονόμου Α. (2003). Εισαγωγή στην Επιχειρησιακή Έρευνα, σελ. 61.
[61]Λουλάκης Μ. (2015). Στοχαστικές Διαδικασίες, σελ. 26-27.

2.3.2 Η περίπτωση των στοχαστικών περιορισμών

Στο παρόν κεφάλαιο γίνεται μία σύντομη αναφορά στην περίπτωση των στοχαστικών περιορισμών ή περιορισμών τυχαιότητας ή περιορισμών πιθανοτήτων . Το ερώτημα που τίθεται εδώ είναι κατά πόσο θα αλλάξει ο αλγόριθμος αν προστεθούν περιορισμοί πιθανοτήτων.Δεν είναι καθόλου σπάνιο να υπάρχουν περιορισμοί που επηρεάζουν συγκεκριμένες συχνότητες για κάποιες καταστάσεις της αλυσίδας. Αυτό σημαίνει ότι κάποιοι περιορισμοί δεν απασχολούν αιτιοκρατικά το σύστημά μας αλλά ασκούν επίδραση σε κάποιες πιθανότητες αυτού. Για παράδειγμα κάτι τέτοιο είναι δυνατό να συμβεί όταν:

- σε ένα πρόβλημα συντήρησης (maintenance problem) , με τυχαίες αλλαγές στην κατάσταση του υπό συντήρηση μηχανήματος, υπάρχει περιορισμός στην συχνότητα εμφάνισης μίας βλάβης.
- σε πρόβλημα ελέγχου αποθεμάτων (inventory control), με δυσκολία στον ακριβή υπολογισμό του κόστους έλλειψης μία μονάδος (shortage cost), υπάρχουν περιορισμοί όσον αφορά τις πιθανότητες έλλειψης ή περιορισμοί στο ποσοστό των περιόδων στις οποίες η ζήτηση δεν καλύφθηκε από το διαθέσιμο απόθεμα.

Φυσικά τέτοιου είδους περιορισμοί επηρεάζουν την λήψη της βέλτιστης απόφασης και κατά συνέπεια αλλάζουν δραστικά το αποτέλεσμα . Το ακόλουθο παράδειγμα[62] είναι αρκετά διαφωτιστικό.

Παράδειγμα 2.2:

Η ζήτηση D ενός προϊόντος Π είναι τυχαία μεταβλητή με κατανομή

$$P(D = x) = \begin{cases} 1/6, & x = 0 \\ 1/6, & x = 1 \\ 2/3, & x = 2 \end{cases}$$

Οι τ.μ D_i είναι ανεξάρτητες για διαδοχικές μέρες i. Το θέμα που ανακύπτει είναι το πόσο θα πρέπει να παραγγείλουμε από το προϊόν Π ώστε να καλύψουμε την ζήτηση. Η παραγγελία μας αυτή γίνεται στην αρχή κάθε ημέρας και η παράδοση των μονάδων που παραγγείλαμε γίνεται άμεσα (ακαριαία-ο χρόνος παράδοσης είναι αμελητέος). Το κόστος παραγγελίας κάθε μονάδας του Π είναι c χρηματικές μονάδες. Επίσης οι μονάδες που μένουν απούλητες καταστρέφονται ως παρωχημένες.

Ο στόχος μας είναι η ελαχιστοποίηση του μέσου ημερίσιου κόστους παραγγελίας με περιορισμό η ζήτηση να ικανοποιείται με ποσοστό (πιθανότητα) τουλάχιστον 1/3 .

Έχουμε δύο τρόπους να αντιμετωπίσουμε το συγκεκριμένο πρόβλημα: ο 1ος είναι ο αιτιοκρατικός (ντετερμινιστικός έλεγχος),ενώ ο 2ος είναι ο στοχαστικός (τυχαιοποιημένος έλεγχος).

Στα παρακάτω θεωρούμε ως S το ημερίσιο απόθεμα μας για την κάλυψη της ζήτησης. Στο συγκεκριμένο παράδειγμα αυτό ισοδυναμεί με την ποσότητα της παραγγελίας που έχουμε κάνει .

[62]Tijms, H.C. (1986). Stochastic Modeling and Analysis, A Computational Approach, σελ. 185 .

$1^{ος}$ τρόπος (ντετερμινιστικός έλεγχος): παραγγέλνουμε μία (1) μονάδα ημερισίως.

Ο περιορισμός απαιτεί $P(S \geq D) \geq 1/3$. Πράγματι, S=1 και

$P(S \geq D) = P(1 \geq D)= 1- P(D >1) = 1- P(D=2) = 1- 2/3 =1/3$.

Εύκολα θα υποστηριζόταν ότι αυτός είναι ο βέλτιστος τρόπος αφού φαινομενικά παραγγέλνουμε τον ελάχιστο απαιτούμενο αριθμό προϊόντων Π για να καλύψουμε την ζήτηση. Μάλιστα, το ημερίσιο κόστος είναι c χρηματικές μονάδες- φαινομενικά το ελάχιστο δυνατό.

$2^{ος}$ τρόπος (τυχαιοποιημένος έλεγχος) : παραγγέλνουμε 0 μονάδες με πιθανότητα 4/5 και 2 μονάδες με πιθανότητα 1/5.

Τότε εφαρμόζοντας το Θεώρημα Ολικής Πιθανότητας έχουμε:

$P(S \geq D) = P(S \geq D|S=0)\cdot P(S =0) + P(S \geq D|S=2)\cdot P(S =2)$

$= P(D=0)\cdot P(S =0) + P(D \leq 2)\cdot P(S =2)$

$= \frac{1}{6}\cdot\frac{4}{5} + 1\cdot\frac{1}{5} = \frac{1}{3}$

που σημαίνει ότι ο περιορισμός ικανοποιείται. Το αναμενόμενο (μέσο) ημερίσιο κόστος παραγγελίας είναι

$E(c\cdot S) = \sum_{s\in R_s} c\cdot s\cdot P(S=s) = c\cdot 0\cdot\frac{4}{5} + c\cdot 2\cdot\frac{1}{5} = 0{,}4c$.

Συγκρίνοντας τους δύο τρόπους ελέγχου συμπεραίνουμε ότι ο δεύτερος από αυτούς είναι καλύτερος . Αυτό σημαίνει ότι ο τυχαιοποιημένος έλεγχος είναι προτιμότερος και αναιρεί το πρώτο φαινομενικό συμπέρασμα ότι η ημερίσια παραγγελία μίας μονάδας επιτυγχάνει το ελάχιστο ημερίσιο κόστος παραγγελίας.

Υπενθυμίζουμε τον ορισμό μιας στάσιμης τυχαιοποιημένης πολιτικής :

Μία πολιτική π καλείται στάσιμη τυχαιοποιημένη πολιτική αν μπορεί να περιγραφεί από μία κατανομή πιθανοτήτων $\{ \pi_\alpha(i), \alpha\in A(i)\}$ για κάθε $i\in S$, τέτοια ώστε αν η αλυσίδα βρίσκεται στην κατάσταση i,η απόφαση $\alpha\in A(i)$ θα επιλεγεί με πιθανότητα $\pi_\alpha(i)$.

Προφανώς $\pi_\alpha(i) \geq 0$ και $\sum_{\alpha\in A(i)} \pi_\alpha(i)=1\ \forall\ i\in S$.

Ειδική περίπτωση: Αν οι $\pi_\alpha(i)$ παίρνουν τις τιμές 0 ή 1 $\forall\ i\in S$ και $\alpha \in A(i)$, τότε η στάσιμη τυχαιοποιημένη πολιτική π εκφυλίζεται στην γνωστή μας στάσιμη πολιτική, όπου οι επιλογές των αποφάσεων γίνονται αιτιοκρατικά.

Ορισμός:

Για κάθε πολιτική π, ορίζουμε ως $f_{i\alpha}$ (π) το ποσοστό των χρόνων απόφασης σε βάθος χρόνου στις οποίες η αλυσίδα βρισκόταν στην κατάσταση i,χρησιμοποιούνταν η πολιτική π και *υπό αυτήν την πολιτική* είχε ληφθεί η απόφαση α .

Υπενθυμίζουμε ότι ο στόχος παραμένει η ελαχιστοποίηση του μέσου ρυθμού κόστους. Ωστόσο, στην προκειμένη περίπτωση υπάρχουν γραμμικοί περιορισμοί στις ποσότητες f_{ia} (π), οι οποίοι μορφοποιούνται ως εξής :

$$\sum_{i \in S} \sum_{a \in A(i)} \gamma_{ia}(k) \cdot f_{ia}\ (\pi) \leq \beta(k)$$

για $k = 1,2,\ldots, L$ με $\gamma_{ia}(k)$, $\beta(k)$: δοσμένες σταθερές.

Τελικά, το (π.γ.π – 1) παίρνει την μορφή (π.γ.π – 3):

Μεταβλητές απόφασης: x_{ia} , $i \in S, a \in A(i)$

Αντικειμενική συνάρτηση:

$$min\, z = min \sum_{i \in S} \sum_{a \in A(i)} c_i(a) x_{ia}$$

Περιορισμοί:

$$\sum_{a \in A(j)} x_{ja} - \sum_{i \in S} \sum_{a \in A(i)} p_{ij}(a) x_{ia} = 0\ , j \in S$$

$$\sum_{i \in S} \sum_{a \in A(i)} x_{ia} = 1$$

$$\sum_{i \in S} \sum_{a \in A(i)} \gamma_{ia}(k) \cdot x_{ia} \leq \beta(k)\ , k\ =\ 1,2, \ldots, L$$

$$x_{ia} \geq 0\ , i \in S, a \in A(i)\ .$$

με $\gamma_{ia}(k)$, $\beta(k)$: δοσμένες σταθερές.

Στην συγκεκριμένη περίπτωση, οι μεταβλητές απόφασης x_{ia} υποκαθιστούν τις f_{ia} (π) με π την ζητούμενη βέλτιστη τυχαιοποιημένη πολιτική , η οποία υπολογίζεται με βάση το Θεώρημα V[63] .

[63] Για μία απόδειξη αυτού βλ. Derman, C. (1970). Finite State Markovian Decision Processes, Academic Press, New York.

Θεώρημα V:

Έστω η μαρκοβιανή αλυσίδα $\{X_n\}_{n\in\mathbb{N}_0}$, για την οποία ισχύει ότι για κάθε[64] πολιτική R, η $\{X_n(R)\}_{n\in\mathbb{N}_0}$ έχει μοναδική κλειστή κλάση ή σύνολο. Έστω $\{x_{ia}^*\}$ η βέλτιστη βασική εφικτή λύση[65] του (π.γ.π – 3) και $S^* = \{i | \sum_{a\in A(i)} x_{ia}^* > 0\}$. Τότε μία τυχαιοποιημένη πολιτική που είναι βέλτιστη ως προς τον μέσο ρυθμό κόστους είναι η

$$\pi_a^*(i) = \begin{cases} x_{ia}^* \Big/ \sum_{a\in A(i)} x_{ia}^* & , \ a \in A(i) \text{ και } i \in S^* \\ c & , \text{αλλιώς} \end{cases}$$

όπου c μία αυθαίρετη τιμή .

Προφανώς η ελάχιστη τιμή του μέσου κόστους δίνεται από την βέλτιστη τιμή της αντικειμενικής συνάρτησης του (π.γ.π – 3) .

Αξίζει να αναφέρουμε ότι άλλη μία μέθοδος αντιμετώπισης των περιορισμών τυχαιότητας είναι με χρήση των πολλαπλασιαστών Lagrange,η μελέτη των οποίων ξεφεύγει από τους στόχους της παρούσας εργασίας. Σχετικά μα την μέθοδο αυτή παραπέμπουμε τον αναγνώστη στην μελέτη του Tijms[66] (1986).

2.4 Σύγκριση των μεθόδων επίλυσης

Έχοντας αναφέρει και τους τρεις τρόπους επίλυσης έχουμε πλέον την δυνατότητα να κάνουμε κάποιες συγκρίσεις,αναφέροντας τα πλεονεκτήματα και τα μειονεκτήματα καθενός από αυτούς.

Όσον αφορά τον αλγόριθμο γραμμικού προγραμματισμού ,αυτός βασίζεται στην μέθοδο Simplex, οπότε τόσο οι κώδικες επίλυσης όσο και η ανάλυση ευαισθησίας είναι εύκολα διαθέσιμα. Ωστόσο, η μέθοδος Simplex εξαρτάται σε μεγάλο βαθμό από το μέγεθος του προβλήματος, ενώ η μέθοδος βελτίωσης της πολιτικής απαιτεί έναν γενικά μικρό αριθμό επαναλήψεων για να οδηγηθεί στην βέλτιστη λύση ανεξαρτήτως των διαστάσεων του προβλήματος. Οπότε, η μέθοδος βελτίωσης πολιτικής χρειάζεται μικρότερο αριθμό επαναλήψεων από την μέθοδο γραμμικού προγραμματισμού. Όμως, από την άλλη πλευρά η μέθοδος βελτίωσης της πολιτικής έχει αυξημένη πολυπλοκότητα με την έννοια ότι σε κάθε επανάληψη αυτής, είναι αναγκαίο να λυθεί ένα ακόμη γραμμικό σύστημα - σε αντίθεση με την μέθοδο Simplex στην οποία αλλάζει μόνο μία απόφαση κάθε φορά. Αν και υπάρχουν τροποποιήσεις στις παραπάνω μεθόδους που τις βελτιώνουν αρκετά, είναι δύσκολο να αποφανθούμε ποια από τις δύο είναι πιο αποτελεσματική.

Για την ακρίβεια,το μεγάλο μειονέκτημα των προαναφερθέντων τρόπων επίλυσης είναι ότι απαιτούν την επίλυση γραμμικών συστημάτων διάστασης ίσης με

---

[64] Εδώ δεν μπορούμε να κάνουμε την ασθενή υπόθεση ότι για κάθε βέλτιστη πολιτική η $\{X_n(R)\}$ έχει μοναδική κλειστή κλάση .

[65] Έχουμε υποθέσει ότι οι περιορισμοί του (π.γ.π – 3) επιτρέπουν την ύπαρξη εφικτής λύσης.

[66] Tijms, H.C. (1986). Stochastic Modeling and Analysis, A Computational Approach, σελ. 186-187.

το πλήθος των καταστάσεων. Ως εκ τούτου, οι συγκεκριμένοι μέθοδοι δεν προτιμώνται για μεγάλης κλίμακας προβλήματα,αφού παρουσιάζουν όλο και μεγαλύτερη πολυπλοκότητα,όσο αυξάνονται οι καταστάσεις. Στην περίπτωση αυτή,η καταλληλότερη μέθοδος είναι αυτή της επαναληπτικής τιμής. Όπως αναφέρεται και στην υποπαράγραφο 2.2 του παρόντος,με τον αλγόριθμο επαναληπτικής τιμής υπολογίζεται αναδρομικά μία ακολουθία τιμών,η οποία (κάτω από την υπόθεση της απεριοδικότητας της αλυσίδας) συγκλίνει στην βέλτιστη τιμή του μέσου ρυθμού κόστους – μάλιστα η εν λόγω προσέγγιση μπορεί να είναι όσο καλή θέλουμε. Κατ' αυτόν τον τρόπο αποφεύγεται η επίλυση μεγάλων γραμμικών συστημάτων,ενώ ταυτόχρονα παρέχονται πολύ καλά όρια μέσα στα οποία κινείται το μέσο κόστος ανά μονάδα χρόνου. Τα παραπάνω καθιστούν την μέθοδο επαναληπτικής τιμής ως την πλέον αποτελεσματική και ελκυστική για προβλήματα μεγάλης διάστασης.

ΚΕΦΑΛΑΙΟ 3 :

ΛΥΜΕΝΕΣ ΕΦΑΡΜΟΓΕΣ

Περίληψη: Σε αυτό το κεφάλαιο επιλύονται εφαρμογές των Μ.Δ.Α από το κεφάλαιο 1. Συγκεκριμένα παρατίθενται οι λύσεις των παραδειγμάτων 1.3 , 1.4 και 1.5. Η επίλυση του 1.3 γίνεται και με τις τρεις μεθόδους,του 1.4 με την μέθοδο βελτίωσης της πολιτικής και του 1.5 με την μέθοδο του γραμμικού προγραμματισμού.

Παράδειγμα 3.1:

Υποθέτουμε ότι στο παράδειγμα 1.3 έχουμε N=5 και οι πιθανότητες $p_{ij}(0) = q_{ij}$ φαίνονται στον πίνακα

$$\mathbb{Q} = (q_{ij}) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0.75 & 0.20 & 0.05 & 0 & 0 \\ 0 & 0.50 & 0.20 & 0.20 & 0.10 \\ 0 & 0 & 0.50 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0.30 & 0.70 \end{pmatrix} \end{array}.$$

Το ζητούμενο παραμένει το ίδιο: να βρεθεί η βέλτιστη πολιτική συντήρησης – δηλαδή εκείνη η πολιτική που ελαχιστοποιεί το ρυθμό μέσου κόστους (ποσοστό ημερών μη – λειτουργίας ή επισκευής του μηχανήματος) $g(R) = \sum_{i \in S} C(i, R_i) \cdot \pi_i(R)$.

Λύση 1:

Εφαρμόζουμε τον αλγόριθμο βελτίωσης της πολιτικής. Η πολιτική που θα χρησιμοποιείται στην k-οστή επανάληψη θα συμβολίζεται με

$$R^{(k)} = \left(R_1^{(k)}, \dots, R_N^{(k)}, R_{N+1}^{(k)}\right).$$

1η ΕΠΑΝΑΛΗΨΗ :

Βήμα 1ο : Ορίζουμε ως αρχική πολιτική την $R^{(1)}$= (0,1,1,1,1,1) που σημαίνει ότι

$$R_1^{(1)} = 0, \quad R_i^{(1)} = 1, \quad i = 2,3,4,5,6 \,.$$

Βήμα 2ο : Υπολογίζουμε τα $g(R^{(1)})$, $u_i\left(R^{(1)}\right), i = 1,2, \dots ,6$ λύνοντας το σύστημα:

$$\Sigma_1 : \begin{cases} u_i = C\left(i, R_i^{(1)}\right) - g + \sum_{j \in S} p_{ij}\left(R_i^{(1)}\right) \cdot u_j \ , \ i = 1,2, \dots ,6 \\ u_6 = 0 \end{cases} .$$

Ας βρούμε πρώτα τις εξισώσεις του συστήματος:

$$i = 1 : u_1 = C\left(1, R_1^{(1)}\right) - g + \sum_{j \in S} p_{1j}\left(R_1^{(1)}\right) \cdot u_j$$

$$u_1 = 0 - g + p_{11}\left(R_1^{(1)}\right) \cdot u_1 + p_{12}\left(R_1^{(1)}\right) \cdot u_2 + p_{13}\left(R_1^{(1)}\right) \cdot u_3 + 0 + 0$$

$$u_1 = -g + 0.75 \cdot u_1 + 0.20 \cdot u_2 + 0.05 \cdot u_3 \quad (\text{αφού } R_1^{(1)} = 0)$$

$i = 2 : u_2 = C\left(2, R_2^{(1)}\right) - \mathrm{g} + \sum_{j \in S} p_{2j}\left(R_2^{(1)}\right) \cdot u_j$

$u_2 = 1 - \mathrm{g} + p_{21}\left(R_2^{(1)}\right) \cdot u_1$ (γιατί $R_2^{(1)} = 1$, οπότε $p_{2j}(1) = 0, j \neq 1$)

$u_2 = 1 - \mathrm{g} + u_1$

Ομοίως για $i = 3, 4$:

$u_3 = 1 - \mathrm{g} + u_1$

$u_4 = 1 - \mathrm{g} + u_1$

$i = 5 : u_5 = 1 - \mathrm{g} + u_6$ (γιατί $p_{56}\left(R_5^{(1)}\right) = p_{56}(1) = 1$)

$i = 6 : u_6 = 1 - \mathrm{g} + u_1$ ($i = 6$: 2η ημέρα επισκευής $p_{61}\left(R_6^{(1)}\right) = p_{61}(1) = 1$).

Επομένως έχουμε να λύσουμε το σύστημα:

$$\Sigma_1: \begin{cases} u_1 = -\mathrm{g} + 0.75 \cdot u_1 + 0.20 \cdot u_2 + 0.05 \cdot u_3 \\ u_2 = 1 - \mathrm{g} + u_1 \\ u_3 = 1 - \mathrm{g} + u_1 \\ u_4 = 1 - \mathrm{g} + u_1 \\ u_5 = 1 - \mathrm{g} + u_6 \\ u_6 = 1 - \mathrm{g} + u_1 \\ u_6 = 0 \end{cases} .$$

Εδώ, η κατάσταση i=6 έχει επιλεγεί αυθαίρετα ώστε $u_6 = 0$ για να ικανοποιηθεί το Θεώρημα Ι (Γ).

Η επίλυση του συγκεκριμένου συστήματος απαιτεί λίγες πράξεις και μάλιστα μικρής πολυπλοκότητας . Παρ'όλα αυτά, μπορούμε να χρησιμοποιήσουμε εναλλακτικά την on-line εφαρμογή WolframAlpha. Πληκτρολογώντας στο πεδίο εισαγωγής της εφαρμογής τα παρακάτω:

solve { -u1 -g+0.75*u1+0.20*u2+0.05*u3=0, u1-g-u2+1=0, u1-g-u3+1=0, u1-g-u4+1=0, 1-g-u5=0,1-g+ u1=0}

η εφαρμογή εμφανίζει ως output τους δύο παρακάτω πίνακες:

solve	$-u1 - g + 0.75\ u1 + 0.2\ u2 + 0.05\ u3 = 0$
	$u1 - g - u2 + 1 = 0$
	$u1 - g - u3 + 1 = 0$
	$u1 - g - u4 + 1 = 0$
	$1 - g - u5 = 0$
	$1 - g + u1 = 0$

και

$$g = \frac{1}{5} \text{ and } u1 = -\frac{4}{5} \text{ and } u2 = 0 \text{ and } u3 = 0 \text{ and } u4 = 0 \text{ and } u5 = \frac{4}{5}$$

Στον πρώτο πίνακα φαίνεται το σύστημα που επιλύθηκε , ενώ στον δεύτερο η λύση του, που είναι η :

$g(R^{(1)}) = 0.2$ και

$u_1(R^{(1)}) = -0.8,$ $u_2(R^{(1)}) = u_3(R^{(1)}) = u_4(R^{(1)}) = 0,$ $u_5(R^{(1)}) = 0.8,$ $u_6(R^{(1)}) = 0$.

Βήμα 3° : Βελτιώνουμε την πολιτική , επιλέγοντας για κάθε i εκείνες τις αποφάσεις α ώστε να ελαχιστοποιείται η ποσότητα ελέγχου:

$$\min_{\alpha \in A(i)} [\, C(i,a) - g(R^{(1)}) + \sum_{j \in S} p_{ij}(a) \cdot u_j(R^{(1)})] \,.$$

i=1 : A(1)={0} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=0. Άρα

$$R_1^{(2)} = 0.$$

*i=2 :*A(2)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(2)} [\, C(2,a) - 0.2 - 0.8 p_{21}(a) + 0.8 p_{25}(a)]$$

απόφαση α=0: $0 - 0.2 - 0.8 \cdot 0 + 0.8 \cdot 0.10 = -0.12 \quad \leftarrow min$
απόφαση α=1: $1 - 0.2 - 0.8 \cdot 1 + 0.8 \cdot 0 = 0$

επιλογή απόφασης που ελαχιστοποιεί : α=0 → $R_2^{(2)} = 0$.

i=3 : A(3)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(3)} [\, C(3,a) - 0.2 - 0.8 p_{31}(a) + 0.8 p_{35}(a)]$$

απόφαση α=0: $0 - 0.2 - 0.8 \cdot 0 + 0.8 \cdot 0.25 = 0$
απόφαση α=1: $1 - 0.2 - 0.8 \cdot 1 + 0.8 \cdot 0 = 0$

αυθαίρετη επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_3^{(2)} = 1$.

i=4 : A(4)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(4)} [\, C(4,a) - 0.2 - 0.8 p_{41}(a) + 0.8 p_{45}(a)]$$

απόφαση α=0: $0 - 0.2 - 0.8 \cdot 0 + 0.8 \cdot 0.70 = 0.36$
απόφαση α=1: $1 - 0.2 - 0.8 \cdot 1 + 0.8 \cdot 0 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_4^{(2)} = 1$.

i=5 : A(5)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$$R_5^{(2)} = 1.$$

i=6 : A(6)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$$R_6^{(2)} = 1.$$

Συνεπώς, η νέα πολιτική είναι η $R^{(2)} = (0,0,1,1,1,1)$. Αφού

$$R^{(2)} = (0,0,1,1,1,1) \neq R^{(1)} = (0,1,1,1,1,1)$$

Ο αλγόριθμος δεν τερματίζει ,οπότε ακολουθεί δεύτερη επανάληψη.

2η ΕΠΑΝΑΛΗΨΗ :

Βήμα 1ο : Ορίζουμε ως αρχική πολιτική την $R^{(2)} = (0,0,1,1,1,1)$ που σημαίνει ότι

$$R_1^{(2)} = R_2^{(2)} = 0, \quad R_i^{(1)} = 1, \quad i = 3,4,5,6 .$$

Βήμα 2ο : Υπολογίζουμε τα $g(R^{(2)})$, $u_i(R^{(2)}), i = 1,2,\dots,6$ λύνοντας το σύστημα:

$$\Sigma_2: \begin{cases} u_i = C\left(i, R_i^{(2)}\right) - g + \sum_{j\in S} p_{ij}\left(R_i^{(2)}\right) \cdot u_j \ , \ i = 1,2,\dots,6 \\ u_6 = 0 \end{cases} .$$

Ας βρούμε πρώτα τις εξισώσεις του συστήματος:

$i = 1 : u_1 = C\left(1, R_1^{(2)}\right) - g + \sum_{j\in S} p_{1j}\left(R_1^{(2)}\right) \cdot u_j$

$$u_1 = 0 - g + p_{11}\left(R_1^{(2)}\right) \cdot u_1 + p_{12}\left(R_1^{(2)}\right) \cdot u_2 + p_{13}\left(R_1^{(2)}\right) \cdot u_3 + 0 + 0$$

$u_1 = -g + 0.75 \cdot u_1 + 0.20 \cdot u_2 + 0.05 \cdot u_3$ (αφού $R_1^{(1)} = 0$ οι πιθανότητες δεν άλλαξαν)

$i = 2 : u_2 = C\left(2, R_2^{(2)}\right) - g + \sum_{j\in S} p_{2j}\left(R_2^{(2)}\right) \cdot u_j$

$u_2 = 0 - g + p_{21}\left(R_2^{(2)}\right) \cdot u_1 + p_{22}\left(R_2^{(2)}\right) \cdot u_2 + p_{23}\left(R_2^{(2)}\right) \cdot u_3 + p_{24}\left(R_2^{(2)}\right) \cdot u_4 + p_{25}\left(R_2^{(2)}\right) \cdot u_5$ (γιατί $R_2^{(2)} = 0$, οπότε οι πιθανότητες άλλαξαν)

$u_2 = -g + p_{21}(0) \cdot u_1 + p_{22}(0) \cdot u_2 + p_{23}(0) \cdot u_3 + p_{24}(0) \cdot u_4 + p_{25}(0) \cdot u_5$

$u_2 = -g + 0 \cdot u_1 + 0.50 \cdot u_2 + 0.20 \cdot u_3 + 0.20 \cdot u_4 + 0.10 \cdot u_5$ (αφού $p_{ij}(0) = q_{ij}$)

$u_2 = -g + 0.50 \cdot u_2 + 0.20 \cdot u_3 + 0.20 \cdot u_4 + 0.10 \cdot u_5$.

Για $i = 3, 4, 5, 6$ ισχύουν τα ίδια με την 1η επανάληψη αφού $R_i^{(2)} = 1$, οπότε οι πιθανότητες μετάβασης δεν άλλαξαν:

$$u_3 = 1 - \mathrm{g} + u_1$$

$$u_4 = 1 - \mathrm{g} + u_1$$

$$u_5 = 1 - \mathrm{g} + u_6$$

$$u_6 = 1 - \mathrm{g} + u_1$$

Επομένως έχουμε να λύσουμε το σύστημα:

$$\Sigma_2: \begin{cases} u_1 = -\mathrm{g} + 0.75 \cdot u_1 + 0.20 \cdot u_2 + 0.05 \cdot u_3 \\ u_2 = -\mathrm{g} + 0.50 \cdot u_2 + 0.20 \cdot u_3 + 0.20 \cdot u_4 + 0.10 \cdot u_5 \\ u_3 = 1 - \mathrm{g} + u_1 \\ u_4 = 1 - \mathrm{g} + u_1 \\ u_5 = 1 - \mathrm{g} + u_6 \\ u_6 = 1 - \mathrm{g} + u_1 \\ u_6 = 0 \end{cases} .$$

Όπου πάλι η κατάσταση i=6 έχει επιλεγεί αυθαίρετα ώστε $u_6 = 0$ και να ικανοποιείται το Θεώρημα Ι (Γ).

Για την επίλυση του απλού αυτού συστήματος θα χρησιμοποιήσουμε ξανά την on-line εφαρμογή WolframAlpha. Πληκτρολογώντας στο πεδίο εισαγωγής της εφαρμογής τα παρακάτω:

solve{-u1-g+0.75*u1+0.20*u2+0.05*u3=0,-g-0.50*u2+0.20*u3+0.20*u4 +0.10*u5=0, u1-g-u3+1=0, u1-g-u4+1=0, 1-g-u5=0,1-g+ u1=0}

η εφαρμογή εμφανίζει ως output τους δύο παρακάτω πίνακες:

solve	$-u1 - g + 0.75\ u1 + 0.2\ u2 + 0.05\ u3 = 0$
	$-g - 0.5\ u2 + 0.2\ u3 + 0.2\ u4 + 0.1\ u5 = 0$
	$u1 - g - u3 + 1 = 0$
	$u1 - g - u4 + 1 = 0$
	$1 - g - u5 = 0$
	$1 - g + u1 = 0$

και

$g \approx 0.17160$ and $u1 \approx -0.82840$ and
$u2 \approx -0.17751$ and $u3 = 0$ and $u4 = 0$ and $u5 \approx 0.82840$

Συνεπώς, η λύση του συστήματος είναι η:

$g(R^{(2)}) = 0.172$ και

$u_1(R^{(2)}) = -0.828, \quad u_2(R^{(2)}) = -0.178,$

$u_3(R^{(2)}) = u_4(R^{(2)}) = 0, \quad u_5(R^{(2)}) = 0.828, \quad u_6(R^{(2)}) = 0$.

Βήμα 3ο : Βελτιώνουμε την πολιτική, επιλέγοντας για κάθε i εκείνες τις αποφάσεις α ώστε να ελαχιστοποιείται η ποσότητα ελέγχου:

$$\min_{\alpha \in A(i)} [\, C(i,a) - g(R^{(2)}) + \sum_{j \in S} p_{ij}(a) \cdot u_j(R^{(2)})] \,.$$

i=1 : A(1)={0} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=0. Άρα

$$R_1^{(3)} = 0.$$

*i=2 :*A(2)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(2)} [\, C(2,a) - 0.172 - 0.828 p_{21}(a) - 0.178 p_{22}(a) + 0.828 p_{25}(a)]$$

απόφαση α=0: $0 - 0.172 - 0.828 \cdot 0 - 0.178 \cdot 0.5 + 0.828 \cdot 0.1 = -0.178 \quad \leftarrow min$
απόφαση α=1: $\quad 1 - 0.172 - 0.828 \cdot 1 - 0 + 0 = 0$

επιλογή απόφασης που ελαχιστοποιεί : α=0 → $R_2^{(3)} = 0$.

i=3 : A(3)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(3)} [C(3,a) - 0.172 - 0.828 p_{31}(a) - 0.178 p_{32}(a) + 0.828 p_{35}(a)]$$

απόφαση α=0: $\quad 0 - 0.172 - 0 - 0 + 0.828 \cdot 0.25 = 0.035$
απόφαση α=1: $\quad 1 - 0.172 - 0.8282 = 0 \leftarrow min$

αυθαίρετη επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_3^{(3)} = 1$.

i=4 : A(4)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(4)} [\, C(4,a) - 0.172 - 0.828 p_{41}(a) - 0.178 p_{42}(a) + 0.828 p_{45}(a)]$$

απόφαση α=0: $\quad 0 - 0.172 - 0 - 0 + 0.828 \cdot 0.70 = 0.4076$
απόφαση α=1: $\quad 1 - 0.172 - 0.828 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_4^{(3)} = 1$.

i=5 : A(5)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$$R_5^{(3)} = 1.$$

i=6 : A(6)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$$R_6^{(3)} = 1.$$

Συνεπώς, η νέα πολιτική είναι η $R^{(3)} = (0,0,1,1,1,1)$. Αφού

$$R^{(3)} = (0,0,1,1,1,1) \neq R^{(2)} = (0,0,1,1,1,1)$$

ο αλγόριθμος τερματίζει και ως εκ τούτου η βέλτιστη πολιτική είναι η

$$R = R^{(3)} = (0,0,1,1,1,1).$$

Αυτό σημαίνει ότι αν το μηχάνημα βρίσκεται στις καταστάσεις 1 ή 2,δεν επισκευάζεται.

Αντιθέτως αν βρίσκεται σε οποιαδήποτε από τις υπόλοιπες καταστάσεις,τότε πρέπει να επισκευαστεί. Με αυτήν την πολιτική ελαχιστοποιείται το αναμενόμενο ποσοστό ημερών που το μηχάνημα βρίσκεται εκτός λειτουργίας (δηλαδή σε επισκευή) με ελάχιστη τιμή:

g = 0.172 ή 17.2% .

Λύση 2:

Σε αυτήν την λύση εφαρμόζουμε την μέθοδο επαναληπτικής τιμής.

1η ΕΠΑΝΑΛΗΨΗ:

Βήμα 1ο : Για κάθε i ∈ S επιλέγουμε αυθαίρετα $V_0(i)$ τέτοιο ώστε

$$0 \leq V_0(i) \leq minC(i,a).$$

Συγκεκριμένα ορίζουμε : $V_0(i) = 0, i = 1,2,3,4$ και $V_0(5) = V_0(6) = 1$.

Εναλλακτικά θα μπορούσαμε να αρχικοποιήσουμε με μηδενικά για όλα τα *i* αλλά αυτό θα άυξανε τις επαναλήψεις κατά μία και θα οδηγούσε στην αρχικοποίηση που τελικά επιλέξαμε.

Επίσης n=1.

Βήμα 2ο : Υπολογίζουμε τις τιμές $V_1(i) = \min_{a \in A(i)} \{C(i,a) + \sum_{j \in S} p_{ij}(a) \cdot V_0(i)\}$

ή ισοδύναμα της $V_1(i) = \min_{a \in A(i)} \{C(i,a) + p_{i5}(a) + p_{i6}(a)\}$ $\forall\, i \in S$ και την αντίστοιχη πολιτική R(n) που ελαχιστοποιεί την ποσότητα $C(i,a) + p_{i5}(a) + p_{i6}(a)$.

i=1 : A(1)={0},άρα $R_1(1) = 0$ και $V_1(1) = C(1,0) + p_{15}(0) + p_{16}(0) = 0.$

i=2 : A(2)={0, 1}

απόφαση α=0: $0 + 0.10 + 0 = 0.10 \leftarrow min$
απόφαση α=1: $1 + 0 + 0 = 1$

άρα $V_1(2) = 0.10$ και $R_2(1) = 0$.

i=3 : A(3)={0, 1}

απόφαση α=0: $0 + 0.25 + 0 = 0.25 \leftarrow min$
απόφαση α=1: $1 + 0 + 0 = 1$

άρα $V_1(3) = 0.25$ και $R_3(1) = 0$.

i=4 : A(4)={0, 1}

απόφαση α=0: $0 + 0.70 + 0 = 0.70 \leftarrow min$
απόφαση α=1: $1 + 0 + 0 = 1$

άρα $V_1(4) = 0.70$ και $R_4(1) = 0$.

i=5 : A(5)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_1(5) = 1 + 0 + 1 = 2$ και $R_5(1) = 1$.

i=6 : A(6)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_1(6) = 1 + 0 + 0 = 1$ και $R_6(1) = 1$.

Οπότε έχουν προκύψει οι ακολουθίες :

$\{V_1(i)\}_{i\in S} = (0, 0.10, 0.25, 0.70\ , 2, 1)$ και $\{R_i(1)\}_{i\in S} = (0,0,0,0,1,1)$.

Βήμα 3ο : Υπολογισμός φραγμάτων $m_1 = \min_{j\in S}\{V_1(j) - V_0(j)\}$ και

$M_1 = \max_{j\in S}\{V_1(j) - V_0(j)\}$.

Αρχικά έχουμε : $V_1(j) - V_0(j)$

j=1 : 0 ← min

j=2 : 0.10

j=3 : 0.25

j=4 : 0.70

j=5 : 2 ← max

j=6 : 1

οπότε $m_1 = 0$ και $M_1 = 2$ και

$0 \leq M_1 - m_1 \leq m_1 \cdot \varepsilon \iff 0 \leq 2 - 0 \leq 0 \cdot 0.01$ που δεν ισχύει . Αυτό σημαίνει ότι ο αλγόριθμος δεν τερματίζει !

Βήμα 4ο : Θέτουμε n= 2 και πηγαίνουμε στο Βήμα 2.

2η ΕΠΑΝΑΛΗΨΗ:

Βήμα 2ο : Υπολογίζουμε τις τιμές $V_2(i) = \min_{a \in A(i)} \{C(i,a) + \sum_{j \in S} p_{ij}(a) \cdot V_1(i)\}$

ή ισοδύναμα της

$V_2(i) = \min_{a \in A(i)} \{C(i,a) + p_{i2}(a) \cdot 0.10 + p_{i3}(a) \cdot 0.25 + p_{i4}(a) \cdot 0.70 + p_{i5}(a) \cdot 2 + p_{i6}(a) \cdot 1\}$ $\forall\, i \in$ S και την αντίστοιχη πολιτική R(n) που ελαχιστοποιεί την ποσότητα του δεξιού μέλους.

i=1 : A(1)={0} , άρα $R_1(2) = 0$ και $V_2(1) = C(1,0) + 0.20 \cdot 0.10 + 0.05 \cdot 0.25 + 0 = 0.0325$.

i=2 : A(2)={0, 1}

απόφαση α=0: $0 + 0.5 \cdot 0.10 + 0.2 \cdot 0.25 + 0.20 \cdot 0.70 + 0.10 \cdot 2 = 0.44 \leftarrow min$
απόφαση α=1: $1 + 0 = 1$

άρα $V_2(2) = 0.44$ και $R_2(2) = 0$.

i=3 : A(3)={0, 1}

απόφαση α=0: $0 + 0 + 0.5 \cdot 0.25 + 0.25 \cdot 0.70 + 0.25 \cdot 2 + 0 = 0.8 \leftarrow min$
απόφαση α=1: $1 + 0 = 1$

άρα $V_2(3) = 0.8$ και $R_3(2) = 0$.

i=4 : A(4)={0, 1}

απόφαση α=0: $0 + 0.30 \cdot 0.70 + 0.70 \cdot 2 + 0 = 1.61$
απόφαση α=1: $1 + 0 = 1 \leftarrow min$

άρα $V_2(4) = 1$ και $R_4(2) = 1$.

i=5 : A(5)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_2(5) = 1 + 0 + 1 = 2$ και $R_5(2) = 1$.

i=6 : A(6)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_2(6) = 1 + 0 + 0 = 1$ και $R_6(2) = 1$.

Οπότε έχουν προκύψει οι ακολουθίες :

$\{V_2(i)\}_{i\in S} = (0.0325, 0.44,\ 0.80\ , 1\ , 2,\ 1)$ και $\{R_i(2)\}_{i\in S} = (0,0,0,1,1,1)$.

Βήμα 3ο : Υπολογισμός φραγμάτων $m_2 = \min_{j\in S}\{V_2(j) - V_1(j)\}$ και

$M_2 = \max_{j\in S}\{V_2(j) - V_1(j)\}$.

Αρχικά έχουμε : $V_2(j) - V_1(j)$

j=1 : 0.0325

j=2 : 0.34

j=3 : 0.55 ← max

j=4 : 0.30

j=5 : 0 ← min

j=6 : 0

οπότε $m_2 = 0$ και $M_2 = 0.55$ και

$0 \leq M_2 - m_2 \leq m_2 \cdot \varepsilon \iff 0 \leq 0.55 - 0 \leq 0 \cdot 0.01$ που δεν ισχύει . Αυτό σημαίνει ότι ο αλγόριθμος δεν τερματίζει !

Βήμα 4ο : Θέτουμε n= 3 και πηγαίνουμε στο Βήμα 2.

3η ΕΠΑΝΑΛΗΨΗ:

Βήμα 2ο : Υπολογίζουμε τις τιμές $V_3(i) = \min_{a\in A(i)}\{C(i,a) + \sum_{j\in S} p_{ij}(a)\cdot V_2(i)\}$

ή ισοδύναμα της

$V_3(i) = \min_{a\in A(i)}\{C(i,a) + p_{i1}(a)\cdot 0.0325 + p_{i2}(a)\cdot 0.44 + p_{i3}(a)\cdot 0.8 + p_{i4}(a) + p_{i5}(a)\cdot 2 + p_{i6}(a)\}\ \forall\ i \in S$ και την αντίστοιχη πολιτική R(n) που ελαχιστοποιεί την ποσότητα του δεξιού μέλους.

i=1 : A(1)={0} , άρα $R_1(3) = 0$ και $V_3(1) = C(1,0) + 0.75\cdot 0.0325 + 0.20\cdot 0.44 + 0.05\cdot 0.8 + 0 = 0.1524$.

i=2 : A(2)={0, 1}

απόφαση α=0: $0 + 0 + 0.50\cdot 0.44 + 0.20\cdot 0.8 + 0.20 + 0.10\cdot 2 = 0.78 \leftarrow min$
απόφαση α=1: $1 + 1\cdot 0.0325 = 1.0325$

άρα $V_3(2) = 0.78$ και $R_2(3) = 0$.

i=3 : A(3)={0, 1}

απόφαση α=0: $0 + 0 + 0 + 0.50 \cdot 0.8 + 0.25 + 0.25 \cdot 2 = 1.15$
απόφαση α=1: $1 + 1 \cdot 0.0325 = 1.0325 \leftarrow min$

άρα $V_3(3) = 1.0325$ και $R_3(3) = 1$.

i=4 : A(4)={0, 1}

απόφαση α=0: $0 + 0 + 0 + 0 + 0.30 + 0.70 \cdot 2 = 1.7$
απόφαση α=1: $1 + 0.0325 = 1.0325 \leftarrow min$

άρα $V_3(4) = 1.0325$ και $R_4(3) = 1$.

i=5 : A(5)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_3(5) = 1 + 1 = 2$ και $R_5(3) = 1$.

i=6 : A(6)={1} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=1. Άρα

$V_3(6) = 1 + 0.0325 = 1.0325$ και $R_6(3) = 1$.

Οπότε έχουν προκύψει οι ακολουθίες :

$\{V_3(i)\}_{i\in S} = (0.1524,\ 0.78,\ 1.0325\ , 1.0325\ , 2,\ 1.0325)$ και

$\{R_i(3)\}_{i\in S} = (0,0,1,1,1,1)$.

Βήμα 3ο : Υπολογισμός φραγμάτων $m_3 = \min_{j\in S}\{V_3(j) - V_2(j)\}$ και

$M_3 = \max_{j\in S}\{V_3(j) - V_2(j)\}$.

Αρχικά έχουμε : $V_3(j) - V_2(j)$

j=1 : 0.1199

j=2 : 0.34 ← max

j=3 : 0.2325

j=4 : 0.0325

j=5 : 0 ← min

j=6 : 0.0325

οπότε $m_3 = 0$ και $M_3 = 0.34$ και

$0 \leq M_3 - m_3 \leq m_3 \cdot \varepsilon \iff 0 \leq 0.34 - 0 \leq 0 \cdot 0.01$ που δεν ισχύει . Αυτό σημαίνει ότι ο αλγόριθμος δεν τερματίζει για ακόμη μία φορά !

Βήμα 4ο : Θέτουμε n= 4 και πηγαίνουμε στο Βήμα 2.

4η ΕΠΑΝΑΛΗΨΗ:

Εδώ θα προκύψουν οι ακολουθίες:

$\{V_4(i)\}_{i\in S} = (0.322,\ 1.003,\ 1.1524\ , 1.1524\ , 2.0325,\ 1.1524)$ και

$\{R_i(4)\}_{i\in S} = (0,0,1,1,1,1)$,

ενώ $m_4 = 0.0325$ και $M_4 = 0.223$ και το κριτήριο τερματισμού δεν ικανοποιείται για ακόμη μία φορά.

Παρατηρούμε ότι ενώ έχει βρεθεί η βέλτιστη πολιτική (κάτι που γνωρίζουμε αφού έχουμε εφαρμόσει τον αλγόριθμο βελτίωσης της πολιτικής),ο αλγόριθμος επαναληπτικής τιμής δεν έχει τερματίσει στην 4η επανάληψη (σε αντίθεση με τον αλγόριθμο βελτίωσης της πολιτικής που χρειάστηκε μόλις δύο επαναλήψεις στο συγκεκριμένο παράδειγμα). Αυτό συμβαίνει γιατί ο αλγόριθμος επαναληπτικής τιμής χρειάζεται έναν μεγαλύτερο αριθμό επαναλήψεων για να συγκλίνει στον ελάχιστο ρυθμό κόστους . Βέβαια, όπως έχει σημειωθεί, ο συγκεκριμένος αλγόριθμος αποτελεί ιδιαίτερα εύχρηστο εργαλείο για προβλήματα μεγάλης διάστασης,αλλά δεν είναι καθόλου πρακτικός για περιπτώσεις μικρού χώρου καταστάσεων όπως στο παράδειγμα 3.1 όπου το πλήθος των καταστάσεων είναι πολύ μικρό (N(S)=6). Φυσικά,σε κάθε επανάληψη η διαφορά $M_n - m_n$μειώνεται (η ακολουθία είναι φθίνουσα),αλλά θα απαιτηθούν πολλές ακόμη επαναλήψεις για να τερματίσει τελικά ο αλγόριθμος επαναληπτικής τιμής στο συγκεκριμένο παράδειγμα και να δώσει τελικά τον ελάχιστο ρυθμό μέσου κόστους. Σε αυτό το σημείο είναι σημαντικό να τονίσουμε την μικρή πολυπλοκότητα στις πράξεις που παρουσιάζει ο εν λόγω αλγόριθμος,σε αντίθεση με τον αλγόριθμο βελτίωσης της πολιτικής,όπου σε κάθε επανάληψη λύνεται ένα νέο γραμμικό σύστημα (s+1)x(s+1) με s:= N(S).

Λύση 3:

Τέλος εφαρμόζουμε την λύση του γραμμικού προγραμματισμού.

Βήμα 1ο : Επιλύουμε το π.γ.π – 1 :

Μεταβλητές απόφασης: x_{ia} , $i \in S, a \in A(i)$

Αντικειμενική συνάρτηση:

$$min\, z = min \sum_{i\in S} \sum_{a\in A(i)} c_i(a) x_{ia}$$

Περιορισμοί:

$$\sum_{a\in A(j)} x_{ja} - \sum_{i\in S} \sum_{a\in A(i)} p_{ij}(a) x_{ia} = 0\ , j \in S$$

$$\sum_{i\in S} \sum_{a\in A(i)} x_{ia} = 1$$

$$x_{ia} \geq 0\ , i \in S, a \in A(i)\ .$$

Αναλυτικότερα οι περιορισμοί γράφονται ως εξής :

$$j = 1: \qquad x_{10} - \sum_{i \in S} \sum_{a \in A(i)} p_{i1}(a) x_{ia} = 0$$

$$x_{10} - p_{11}(0)x_{10} - p_{21}(0)x_{20} - p_{21}(1)x_{21} - p_{31}(0)x_{30} - p_{31}(1)x_{31} - p_{41}(0)x_{40} - p_{41}(1)x_{41} - p_{61}(1)x_{61} = 0$$

$$x_{10} - 0.75x_{10} - x_{21} - x_{31} - x_{41} - x_{61} = 0$$

$$j = 2: \qquad x_{20} + x_{21} - p_{12}(0)x_{10} - p_{22}(0)x_{20} - p_{32}(0)x_{30} - p_{42}(0)x_{40} = 0$$

$$x_{20} + x_{21} - 0.20x_{10} - 0.50x_{20} \qquad = 0$$

και ομοίως για τα υπόλοιπα.

Τελικά έχουμε να λύσουμε το π.γ.π – 1 :

$$min\, z = min\ x_{21} + x_{31} + x_{41} + x_{51} + x_{61}$$

$$x_{10} - 0.75x_{10} - x_{21} - x_{31} - x_{41} - x_{61} = 0$$

$$x_{20} + x_{21} - 0.20x_{10} - 0.50x_{20} \qquad = 0$$

$$x_{30} + x_{31} - 0.05x_{10} - 0.20x_{20} - 0.50x_{30} = 0$$

$$x_{40} + x_{41} - 0.20x_{20} - 0.25x_{30} - 0.30x_{40} = 0$$

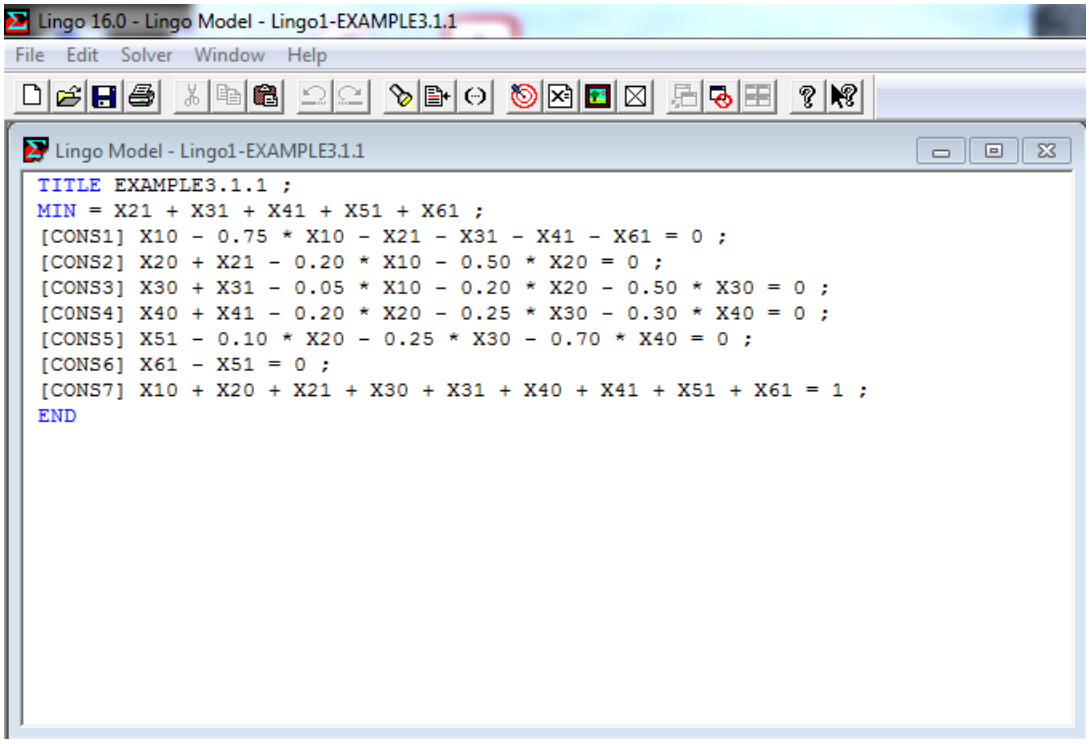
$$x_{51} - 0.10x_{20} - 0.25x_{30} - 0.70x_{40} \qquad = 0$$

$$x_{61} - x_{51} = 0$$

$$x_{10} + x_{20} + x_{21} + x_{30} + x_{31} + x_{40} + x_{41} + x_{51} + x_{61} = 1$$

$$x_{10}, x_{20}, x_{21}, x_{30}, x_{31}, x_{40}, x_{41}, x_{51}, x_{61} \geq 0$$

Για την επίλυση του θα χρησιμοποιήσουμε το πρόγραμμα LINGO της LINDO Systems (έκδοση 16.0). Στην εικόνα 3.1 φαίνεται η μορφή του π.γ.π – 1 στην γλώσσα του LINGO. Γίνεται εύκολα αντιληπτή η ευκολία της μεταφοράς του προβλήματός μας σε αυτήν την μορφή. Αυτό είναι και ένα πολύ θετικό στοιχείο του LINGO που το καθιστά πολύ εύχρηστο τόσο στον αρχάριο όσο και στον προχωρημένο χρήστη. Ο περιορισμός της μη αρνητικότητας των μεταβλητών είναι προεπιλεγμένος στο LINGO (default), οπότε είναι περιττό να τον εισάγουμε .

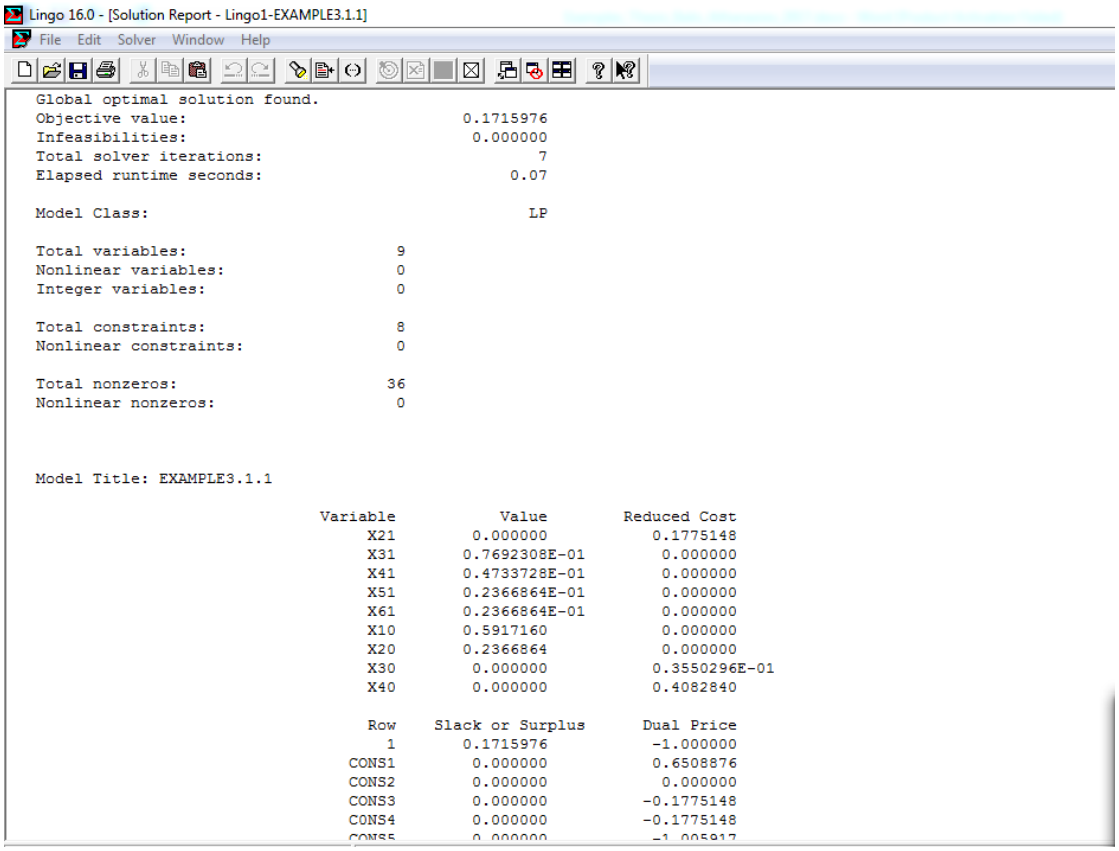


Εικόνα 3.1: Υπολογιστικό φύλλο του LINGO για την επίλυση του προβλήματος 3.1

Πατώντας την επιλογή Solver πάνω αριστερά και εν συνεχεία Solve έχουμε την αναφορά της λύσης (Εικόνα 3.2). Άρα η λύση του π.γ.π – 1 είναι :

$x_{10} = 0.5917$, $x_{20} = 0.2367$, $x_{31} = 0.0769$, $x_{41} = 0.0473$, $x_{51} = x_{61} = 0.0237$

$x_{21} = 0$, $x_{30} = 0$, $x_{40} = 0$ και $min\, z = 0.1716$.



```
Lingo 16.0 - [Solution Report - Lingo1-EXAMPLE3.1.1]
File  Edit  Solver  Window  Help

Global optimal solution found.
Objective value:                              0.1715976
Infeasibilities:                              0.000000
Total solver iterations:                             7
Elapsed runtime seconds:                          0.07

Model Class:                                        LP

Total variables:                     9
Nonlinear variables:                 0
Integer variables:                   0

Total constraints:                   8
Nonlinear constraints:               0

Total nonzeros:                     36
Nonlinear nonzeros:                  0


Model Title: EXAMPLE3.1.1

                              Variable           Value        Reduced Cost
                                   X21        0.000000           0.1775148
                                   X31       0.7692308E-01       0.000000
                                   X41       0.4733728E-01       0.000000
                                   X51       0.2366864E-01       0.000000
                                   X61       0.2366864E-01       0.000000
                                   X10       0.5917160           0.000000
                                   X20       0.2366864           0.000000
                                   X30        0.000000          0.3550296E-01
                                   X40        0.000000          0.4082840

                                   Row    Slack or Surplus      Dual Price
                                     1       0.1715976          -1.000000
                                 CONS1        0.000000           0.6508876
                                 CONS2        0.000000           0.000000
                                 CONS3        0.000000          -0.1775148
                                 CONS4        0.000000          -0.1775148
```

Εικόνα 3.2 : Αναφορά επίλυσης του LINGO για το πρόβλημα 3.1

Βήμα 2ο : Το σύνολο $S^* = \{i | \sum_{a \in A(i)} x_{ia}^* > 0\}$ ={ 1, 2, 3, 4, 5, 6}= S και αφού

$x_{10}^* = 0.5917$, $x_{20}^* = 0.2367$, $x_{31}^* = 0.0769$, $x_{41}^* = 0.0473$, $x_{51}^* = x_{61}^* = 0.0237$

θέτουμε $R_1{}^* := 0$, $R_2{}^* := 0$, $R_3{}^* := 1$, $R_4{}^* := 1$, $R_5{}^* := 1$, $R_6{}^* := 1$.

Βήμα 3ο : Αφού S = S^*, η βέλτιστη πολιτική μέσου ρυθμού κόστους είναι η $R^*=\{R_i{}^*\}$ = (0, 0, 1, 1, 1, 1) με ελάχιστο αναμενόμενο ποσοστό ημερών επισκευής του μηχανήματος 0.1716 ή 17,16% και ο αλγόριθμος τερματίζει.

Παράδειγμα 3.2:

Λύνουμε το παράδειγμα 1.4 χρησιμοποιώντας τον αλγόριθμο βελτίωσης της πολιτικής.

1η ΕΠΑΝΑΛΗΨΗ :

Βήμα 1ο : Ορίζουμε ως αρχική πολιτική την $R^{(1)}$= (0,1,1,1,1) που σημαίνει ότι

$$R_1^{(1)} = 0, \quad R_i^{(1)} = 1, \quad i = 2,3,4,5 .$$

Βήμα 2ο : Υπολογίζουμε τα $g(R^{(1)})$, $u_i(R^{(1)}), i = 1,2, \dots ,5$ λύνοντας το σύστημα:

$$\Sigma_1 : \begin{cases} u_i = C\left(i, R_i^{(1)}\right) - g + \sum_{j \in S} p_{ij}\left(R_i^{(1)}\right) \cdot u_j \ , \ i = 1,2, \dots ,5 \\ u_5 = 0 \end{cases} .$$

Πρώτα βρίσκουμε τις εξισώσεις του συστήματος:

$$i = 1 : u_1 = C\left(1, R_1^{(1)}\right) - \mathrm{g} + \sum_{j \in S} p_{1j}\left(R_1^{(1)}\right) \cdot u_j$$

$$u_1 = 1 - \mathrm{g} + p_{11}(0) \cdot u_1 + p_{12}(0) \cdot u_2 + 0$$

$$u_1 = 1 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2$$

$$i = 2 : u_2 = C\left(2, R_2^{(1)}\right) - \mathrm{g} + \sum_{j \in S} p_{2j}\left(R_2^{(1)}\right) \cdot u_j$$

$$u_2 = 11 - \mathrm{g} + p_{21}(1) \cdot u_1 + p_{21}(1) \cdot u_1 \qquad (\text{γιατί } p_{2j}(1) = q_{1j})$$

$$u_2 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2$$

Ομοίως για $i = 3, 4,5$: (γιατί $p_{ij}(1) = q_{1j}, i = 2,3,4,5$)

$$u_3 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2$$

$$u_4 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2$$

$$u_5 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2$$

Επομένως έχουμε να λύσουμε το σύστημα:

$$\Sigma_1 : \begin{cases} u_1 = 1 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2 \\ u_2 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2 \\ u_3 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2 \\ u_4 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2 \\ u_5 = 11 - \mathrm{g} + \frac{3}{4} \cdot u_1 + \frac{1}{4} \cdot u_2 \\ u_5 = 0 \end{cases}.$$

Εδώ, επιλέγεται αυθαίρετα η κατάσταση i=5 ώστε $u_5 = 0$ για να ικανοποιηθεί το Θεώρημα Ι (Γ).

Η επίλυση του συγκεκριμένου συστήματος γίνεται πολύ εύκολα αφού $u_i = 0$, $\forall$ i=2,3,4,5 και σε συνδυασμό με τις δύο πρώτες εξισώσεις έχουμε g=3.5, $u_1 = -10$.

Συνεπώς, η λύση του Σ_1 , είναι η :

$\mathrm{g}(R^{(1)}) = 3.5$ και

$u_1\left(R^{(1)}\right) = -10, u_2\left(R^{(1)}\right) = u_3\left(R^{(1)}\right) = u_4\left(R^{(1)}\right) = u_5\left(R^{(1)}\right) = 0.$

Βήμα 3° : Βελτιώνουμε την πολιτική , επιλέγοντας για κάθε i εκείνες τις αποφάσεις α ώστε να ελαχιστοποιείται η ποσότητα ελέγχου:

$$\min_{\alpha \in A(i)} [\, C(i,a) - \mathrm{g}(R^{(1)}) + \textstyle\sum_{j \in S} \, p_{ij}(a) \cdot u_j(R^{(1)})] \, .$$

i=1 : A(1)={0} , οπότε η μοναδική απόφαση που μπορούμε να πάρουμε είναι α=0. Άρα

$$R_1^{(2)} = 0.$$

*i=2 :*A(2)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(2)} [\, C(2,a) - 3.5 - 10 p_{21}(a)]$$

απόφαση α=0: $6 - 3.5 - 10 \cdot 0 = 2.5$
απόφαση α=1: $11 - 3.5 - 10 \cdot 3/4 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_2^{(2)} = 1$.

i=3 : A(3)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(3)} [\, C(3,a) - 3.5 - 10 p_{31}(a)]$$

απόφαση α=0: $9 - 3.5 - 10 \cdot 0 = 5.5$
απόφαση α=1: $11 - 3.5 - 10 \cdot 3/4 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_3^{(2)} = 1$.

i=4 : A(4)={0, 1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(4)} [\, C(4,a) - 3.5 - 10 p_{41}(a)]$$

απόφαση α=0: $10 - 3.5 - 10 \cdot 0 = 6.5$
απόφαση α=1: $11 - 3.5 - 10 \cdot 3/4 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_4^{(2)} = 1$.

i=5 : A(5)={0,1}

ποσότητα ελέγχου

$$\min_{\alpha \in A(5)} [\, C(5,a) - 3.5 - 10 p_{51}(a)]$$

απόφαση α=0: $12 - 3.5 - 10 \cdot 0 = 8.5$
απόφαση α=1: $11 - 3.5 - 10 \cdot 3/4 = 0 \leftarrow min$

επιλογή απόφασης που ελαχιστοποιεί : α=1 → $R_5^{(2)} = 1.$

Συνεπώς, η νέα πολιτική είναι η $R^{(2)} = (0,1,1,1,1)$.

Βήμα 4° : Αφού

$$R^{(2)} = (0,1,1,1,1) = R^{(1)},$$

ο αλγόριθμος τερματίζει,οπότε η βέλτιστη πολιτική είναι η $R^{(1)}$= (0,1,1,1,1) που σημαίνει ότι αν το μηχάνημα έχει έστω και ένα ζεύγος απορυθμισμένο, το ρυθμίζουμε, διαφορετικά το αφήνουμε ως έχει. Ο ελάχιστος ρυθμός μέσου κόστους είναι g=3.5 χρηματικές μονάδες ανά μονάδα χρόνου.

Παράδειγμα 3.3:

Λύνουμε το παράδειγμα 1.5 εφαρμόζοντας την λύση του γραμμικού προγραμματισμού.

Βήμα 1° : Επιλύουμε το π.γ.π – 1 :

Μεταβλητές απόφασης: x_{ia} , $i \in S, a \in A(i)$

Αντικειμενική συνάρτηση:

$$min\, z = min \sum_{i \in S} \sum_{a \in A(i)} c_i(a) x_{ia}$$

Περιορισμοί:

$$\sum_{a \in A(j)} x_{ja} - \sum_{i \in S} \sum_{a \in A(i)} p_{ij}(a) x_{ia} = 0 \,, j \in S$$

$$\sum_{i \in S} \sum_{a \in A(i)} x_{ia} = 1$$

$$x_{ia} \geq 0 \,, i \in S, a \in A(i) \,.$$

Αντικαθιστώντας τους αριθμούς έχουμε να λύσουμε το π.γ.π – 1:

$$min\, z = min\ 0x_{00} + 0x_{10} + 3.75x_{20} + 15x_{30} + 41.25x_{40} + 30x_{11} + 35x_{21} + 40x_{31} + 45x_{41}$$

$$x_{00} - \frac{1}{8}x_{00} - \frac{1}{8}x_{11} - \frac{1}{8}x_{21} - \frac{1}{8}x_{31} - \frac{1}{8}x_{41} = 0$$

$$x_{10} + x_{11} - \frac{1}{2}x_{00} - \frac{1}{8}x_{10} - \frac{1}{2}x_{11} - \frac{1}{2}x_{21} - \frac{1}{2}x_{31} - \frac{1}{2}x_{41} = 0$$

$$x_{20} + x_{21} - \frac{1}{4}x_{00} - \frac{1}{2}x_{10} - \frac{1}{4}x_{11} - \frac{1}{8}x_{20} - \frac{1}{4}x_{21} - \frac{1}{4}x_{31} - \frac{1}{4}x_{41} = 0$$

$$x_{30} + x_{31} - \frac{1}{8}x_{00} - \frac{1}{4}x_{10} - \frac{1}{8}x_{11} - \frac{1}{2}x_{20} - \frac{1}{8}x_{21} - \frac{1}{8}x_{30} - \frac{1}{8}x_{31} - \frac{1}{8}x_{41} = 0$$

$$x_{40} + x_{41} - \frac{1}{8}x_{10} - \frac{3}{8}x_{20} - \frac{7}{8}x_{30} - x_{40} = 0$$

$$x_{00} + x_{10} + x_{11} + x_{20} + x_{21} + x_{30} + x_{31} + x_{40} + x_{41} = 1$$

$$x_{00}, x_{10}, x_{11}, x_{20}, x_{21}, x_{30}, x_{31}, x_{40}, x_{41} \geq 0$$

ή ισοδύναμα:

$$min\, z = min\ \ 0x_{00} + 0x_{10} + 3.75x_{20} + 15x_{30} + 41.25x_{40} + 30x_{11} + 35x_{21} + 40x_{31} + 45x_{41}$$

$$7x_{00} - x_{11} - x_{21} - x_{31} - x_{41} \ = 0$$

$$7x_{10} + 4x_{11} - 4x_{00} - 4x_{21} - 4x_{31} - 4x_{41} \ = 0$$

$$7x_{20} + 6x_{21} - 2x_{00} - 4x_{10} - 2x_{11} - 2x_{31} - 2x_{41} \ = 0$$

$$7x_{30} + 7x_{31} - x_{00} - 2x_{10} - x_{11} - 4x_{20} - x_{21} - x_{41} \ = 0$$

$$8x_{41} - x_{10} - 3x_{20} - 7x_{30} = 0$$

$$x_{00} + x_{10} + x_{11} + x_{20} + x_{21} + x_{30} + x_{31} + x_{40} + x_{41} = 1$$

$$x_{00}, x_{10}, x_{11}, x_{20}, x_{21}, x_{30}, x_{31}, x_{40}, x_{41} \geq 0$$

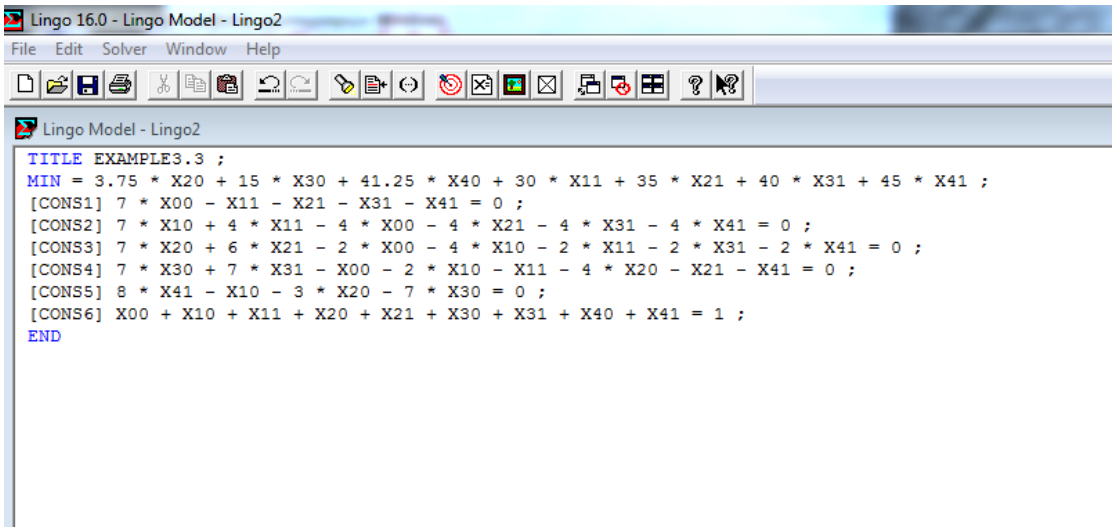
Για την επίλυση του θα χρησιμοποιήσουμε το πρόγραμμα LINGO της LINDO Systems (έκδοση 16.0). Στην εικόνα 3.3 φαίνεται η μορφή του π.γ.π – 1 στην γλώσσα του LINGO, ενώ στην εικόνα 3.4 φαίνεται η αναφορά της λύσης. Υπενθυμίζουμε ότι ο περιορισμός της μη αρνητικότητας των μεταβλητών είναι προεπιλεγμένος στο LINGO (default) , οπότε είναι περιττό να τον εισάγουμε .

Άρα η λύση του π.γ.π – 1 είναι :

$x_{00} = 0.057$, $x_{10} = 0.262$, $x_{11} = 0$, $x_{20} = 0.280$, $x_{21} = 0$, $x_{30} = 0$,

$x_{31} = 0.263$, $x_{40} = 0$, $x_{41} = 0.138$,

και $min\, z = 17.7687$.



Εικόνα 3.3: Η μορφή του π.γ.π – 1 του παραδείγματος 3.3 όπως εμφανίζεται στο φύλλο εργασίας του LINGO.

```
File  Edit  Solver  Window  Help

Global optimal solution found.
Objective value:                              17.76869
Infeasibilities:                              0.000000
Total solver iterations:                             6
Elapsed runtime seconds:                          0.42

Model Class:                                        LP

Total variables:                     9
Nonlinear variables:                 0
Integer variables:                   0

Total constraints:                   7
Nonlinear constraints:               0

Total nonzeros:                     46
Nonlinear nonzeros:                  0


Model Title: EXAMPLE3.3

                              Variable           Value        Reduced Cost
                                   X20       0.2803738            0.000000
                                   X30        0.000000            1.606308
                                   X40        0.000000            23.48131
                                   X11        0.000000            17.52336
                                   X21        0.000000            8.878505
                                   X31       0.2628505            0.000000
                                   X41       0.1378505            0.000000
                                   X00       0.5724299E-01        0.000000
                                   X10       0.2616822            0.000000
```

Εικόνα 3.4: Η αναφορά της λύσης του π.γ.π -1 του παραδείγματος 3.3 όπως εμφανίζεται στο LINGO.

Βήμα 2° : Το σύνολο $S^* = \{i | \sum_{a \in A(i)} x_{ia}^* > 0\}$ ={0, 1, 2, 3, 4}= S και αφού

$x_{00}^* = 0.057, x_{10}^* = 0.262$, $x_{20}^* = 0.280, x_{31}^* = 0.263$, $x_{41}^* = 0.138$,

θέτουμε $R_0{}^* := 0$, $R_1{}^* := 0$, $R_2{}^* := 0$, $R_3{}^* := 1$, $R_4{}^* := 1$.

Βήμα 3° : Αφού S = S^*, ο αλγόριθμος τερματίζει. Άρα η βέλτιστη πολιτική μέσου ρυθμού κόστους είναι η $R^*=\{R_i{}^*\}$ = (0, 0, 0, 1, 1) που σημαίνει ότι αν υπάρχουν πάνω από δύο μονάδες στην αποθήκη τις μεταφέρουμε σε άλλο χώρο. Ο ελάχιστος ρυθμός μέσου κόστους είναι 17.7687 χρηματικές μονάδες ανά μονάδα χρόνου .

Τελικά , η πολιτική που προτάθηκε είναι βέλτιστη.

ΚΕΦΑΛΑΙΟ 4:

ΣΥΓΚΛΙΣΕΙΣ ΤΩΝ ΜΕΘΟΔΩΝ

Περίληψη: Στο κεφάλαιο 4 παρουσιάζονται εν συντομία οι αποδείξεις των συγκλίσεων δύο αλγορίθμων˙ αυτού της βελτίωσης της πολιτικής και του αντίστοιχου της επαναληπτικής τιμής.Επιπλέον,δίνονται και οι σχετικοί ρυθμοί σύγκλισης. Η απόδειξη της σύγκλισης του αλγορίθμου του γραμμικού προγραμματισμού έχει δοθεί στο κεφάλαιο 2.

4.1 ΣΥΓΚΛΙΣΗ ΤΗΣ ΜΕΘΟΔΟΥ ΒΕΛΤΙΩΣΗΣ ΤΗΣ ΠΟΛΙΤΙΚΗΣ

Υπενθυμίζουμε ότι δίνεται μαρκοβιανή αλυσίδα $\{X_n\}_{n\in\mathbb{N}_0}$, τέτοια ώστε για κάθε πολιτική R, η $\{X_n(R)\}_{n\in\mathbb{N}_0}$ έχει μοναδική κλειστή κλάση ή σύνολο. Επιπλέον δοθείσης μίας πολιτικής R οι σχετικές τιμές $u_i(R)$ των καταστάσεων που αντιστοιχούν στην πολιτική αυτή είναι μονοσήμαντα ορισμένες (εκτός ίσως από μία προστιθέμενη σταθερά) και μάλιστα η ακολουθία των πολιτικών R που παράγεται από τον αλγόριθμο βελτίωσης της πολιτικής δεν εξαρτάται από κάποια συγκεκριμένη επιλογή των σχετικών τιμών.

Έστω R ≠ R΄ δύο διαδοχικές πολιτικές όπως αυτές παράγονται από τον αλγόριθμο.

Θεώρημα VI (σύγκλισης αλγορίθμου βελτίωσης πολιτικής) :

Από τα παρακάτω ισχύει ένα ακριβώς :

1) g(R΄) < g(R) ή

2) g(R΄) = g(R) και $u_i(R') \le u_i(R)$ ∀ i∈S , με την ισότητα να ισχύει ∀ i∈S(R΄)[67] και την ανισότητα να ισχύει γνήσια (αυστηρά) για τουλάχιστον ένα i∉S(R΄).

Απόδειξη:

Από το 3ο βήμα του αλγορίθμου έχουμε την ανισότητα (VI1)

$$C(i, R_i') - g(R) + \sum_{j\in S} p_{ij}(R_i') \cdot u_j(R) \le u_i(R)\,, \quad i \in S$$

η οποία ισχύει γνήσια για τα i εκείνα για τα οποία $R_i' \neq R_i$, κάτι που ισχύει εκ κατασκευής της νέας πολιτικής R_i'. Ως εκ τούτου,εφαρμόζοντας το Θεώρημα II προκύπτει g(R΄) ≤ g(R) από την (II2).

Με βάση αυτό θα αποδείξουμε ότι g(R΄) < g(R) μόνο αν η ανισότητα (VI1) ισχύει γνήσια για καταστάσεις που είναι θετικά επαναληπτικές κάτω από την πολιτική R΄, δηλαδή για i∈S(R΄).

[67] Υπενθυμίζουμε ότι S(R) είναι το σύνολο των θετικά επαναληπτικών καταστάσεων κάτω από την πολιτική R.

Οπότε υποθέτουμε ότι για i∈S(R΄) ισχύει

$$C(i, R_i') - \mathrm{g}(R) + \sum_{j \in S} p_{ij}(R_i') \cdot u_j(R) \leq u_i(R)$$

Πολλαπλασιάζουμε κατά μέλη με $\pi_i(\mathrm{R}') \geq 0$ [η γνήσια ανισότητα ισχύει μόνο για i∈S(R΄)[68]] :

$$\pi_i(\mathrm{R}')C(i, R_i') - \pi_i(\mathrm{R}')\mathrm{g}(R) + \sum_{j \in S} \pi_i(\mathrm{R}')p_{ij}(R_i')u_j(R) \leq \pi_i(\mathrm{R}')u_i(R)$$

Προσθέτοντας τις ανισότητες κατά μέλη για όλα τα i∈S:

$\sum_{i \in \mathrm{S}} \pi_i(\mathrm{R}')C(i, R_i') - \mathrm{g}(R) + \sum_{i \in \mathrm{S}} \sum_{j \in S} \pi_i(\mathrm{R}')p_{ij}(R_i')u_j(R) \leq \sum_{i \in \mathrm{S}} \pi_i(\mathrm{R}')u_i(R)$

[προφανώς έχει χρησιμοποιηθεί η σχέση $\sum_{i \in \mathrm{S}} \pi_i(\mathrm{R}') = 1$]

Αλλάζοντας την σειρά άθροισης

$\mathrm{g}(R') - \mathrm{g}(R) + \sum_{j \in \mathrm{S}} \sum_{i \in S} \pi_i(\mathrm{R}')p_{ij}(R_i')u_j(R) \leq \sum_{i \in \mathrm{S}} \pi_i(\mathrm{R}')u_i(R)$

$\mathrm{g}(R') - \mathrm{g}(R) + \sum_{j \in \mathrm{S}} \pi_j(\mathrm{R}')u_j(R) \leq \sum_{i \in \mathrm{S}} \pi_i(\mathrm{R}')u_i(R)$

[αφού εξ' ορισμού $\pi_j(\mathrm{R}') = \sum_{i \in S} \pi_i(\mathrm{R}')p_{ij}(R_i')$, ∀ j∈S]

Επομένως $\mathrm{g}(\mathrm{R}') \leq \mathrm{g}(\mathrm{R})$ (VI2) και η ανισότητα ισχύει γνήσια μόνο αν η (VI1) ισχύει γνήσια για κάποια i∈S(R΄). Εν συντομία, αποδείξαμε ότι ισχύει:

$\mathrm{g}(\mathrm{R}') < \mathrm{g}(\mathrm{R})$ ή

$\mathrm{g}(\mathrm{R}') = \mathrm{g}(\mathrm{R})$ εφόσον η (VI1) ισχύει ως ισότητα ∀ i∈S(R΄) [αν υπήρχε κάποια κατάσταση i∈S(R΄) για την οποία δεν ίσχυε η ισότητα,τότε αναγκαστικά θα ίσχυει η $\mathrm{g}(\mathrm{R}') < \mathrm{g}(\mathrm{R})$ λόγω της (VI2)] .

Μένει να ολοκληρώσουμε την απόδειξη του (2).

Σε αυτήν την περίπτωση έχουμε $\mathrm{g}(\mathrm{R}') = \mathrm{g}(\mathrm{R})$, που σημαίνει ότι η (VI1) ισχύει ως ισότητα ∀ i∈S(R΄) και από το βήμα 3 του αλγορίθμου έχουμε

$$\mathrm{R}_i' = \mathrm{R}_i \text{ , } \forall\, i \in \mathrm{S}(\mathrm{R}') \quad \text{(VI3).}$$

Επιπλέον, αφού $p_{ij}(R) = 0$, ∀ i ∈ S(R) και ∀ j ∉ S(R),

$p_{ij}(R) = 0$ ∀ i, j ∈ S(R) και ομοίως για την R΄

τελικά

$$\mathrm{S}(\mathrm{R}) = \mathrm{S}(\mathrm{R}') \quad \text{(VI4),}$$

δηλαδή οι κλάσεις των θετικά επαναληπτικών καταστάσεων κάτω από τις πολιτικές R και R΄ ταυτίζονται.

[68]Βλ. Μ. Λουλάκη. (2015) «Στοχαστικές Διαδικασίες», σελ. 75.

Εξ' ορισμού, για κάθε πολιτική R οι σχετικές τιμές

$$u_i(R) = K_i(R) - \mathrm{g}(R) \cdot T_i(R), \text{ για } i \in \mathrm{S(R)}$$

δεν εξαρτώνται από τις αποφάσεις R_i , $i \notin$ S(R). Κι αυτό διότι η αλυσίδα θα παραμείνει για πάντα στην κλάση των θετικά επαναληπτικών καταστάσεων, αν βρεθεί σ' αυτήν.

Επομένως, $u_i(R) = u_i(R')$ για κάθε $i \in$ S(R) (VI5) .

Αυτό που τελικά μένει προς απόδειξη είναι ότι $u_i(R') < u_i(R)$ για κάποιο $i \notin$ S(R′).

Από την (VI1) η οποία ισχύει γνήσια για $i : R_i' \neq R_i$ έπεται[69] ότι

$V_m(i, R') - m\mathrm{g} + \sum_{j \in S} p_{ij}{}^{(m)}(R')\, u_j \leq u_i(R)$ η οποία πάλι ισχύει γνήσια για για i : $R_i' \neq R_i$ (VI6).

Από την (I5), αντικαθιστώντας όπου R την R΄ έχουμε

$u_i(R) = V_m(i, R') - m\mathrm{g}(R) + \sum_{j \in S} p_{ij}{}^{(m)}(R')\, u_j(R)$

[ας μην ξεχνάμε ότι είμαστε στην περίπτωση (2) , όπου g(R′) = g(R)]

$u_i(R) - u_i(R') + u_i(R') = V_m(i, R') - m\mathrm{g}(R) + \sum_{j \in S} p_{ij}{}^{(m)}(R')\, u_j(R)$

και από την (I6) προκύπτει η μορφή:

$\sum_{j \in S} p_{ij}{}^{(m)}(R_i)[u_j(R) - u_j(R')] + u_i(R') = V_m(i, R') - m\mathrm{g}(R) + \sum_{j \in S} p_{ij}{}^{(m)}(R')\, u_j(R)$.

Οπότε η (VI6) γίνεται $u_i(R) \geq \sum_{j \in S} p_{ij}{}^{(m)}(R_i)[u_j(R) - u_j(R')] + u_i(R')$ που ισχύει γνήσια μόνο για $i : R_i' \neq R_i$ (VI7).

Παίρνοντας όρια m → ∞ κατά μέλη :

$p_{ij}{}^{(m)}(R_i) \to 0 \ \forall \ i \in$ S και $j \notin$ S(R_i)[70]

$p_{ij}{}^{(m)}(R_i) \to \pi_j(R_i) \ \forall \ j \in$ S(R_i)[71] αλλά $u_j(R) = u_j(R')$ για κάθε $j \in$ S(R) , οπότε η (VI7) γίνεται

$u_i(R) \geq u_i(R')$ που ισχύει γνήσια μόνο για $i : R_i' \neq R_i$.

Λόγω της (VI3) το τελευταίο σημαίνει ότι η γνήσια ανισότητα $u_i(R) > u_i(R')$ ισχύει για $i \notin$ S(R_i) και η απόδειξη ολοκληρώθηκε.

Με το Θεώρημα VI,ουσιαστικά αποδείξαμε ότι σε κάθε επανάληψη του αλγορίθμου βελτίωσης της πολιτικής, είτε θα μειώνεται ο μέσος ρυθμός κόστους , είτε θα μειώνονται οι σχετικές τιμές u_i των παροδικών καταστάσεων $i \notin$ S(R_i). Ως εκ τούτου, αφού το πλήθος των στάσιμων πολιτικών είναι πεπερασμένο,το Θεώρημα VI καταδεικνύει ότι ο αλγόριθμος συγκλίνει κατόπιν πεπερασμένου πλήθους βημάτων.

---

[69] η σχέση προέρχεται από την συνεπαγωγή (II1) → (II5α) .

[70] βλ.Henk C. Tijms. (1986) Stochastic Modelling and Analysis: A Computational Approach, σελ. 178.

[71] Αν $i \notin$ S(R_i) , μετά από πεπερασμένο αριθμό βημάτων η αλυσίδα θα βρεθεί στην κλάση S(R_i) οπότε ισχύουν τα ίδια όρια.

Τέλος, αποδεικνύεται ότι η τάξη της σύγκλισης του αλγορίθμου βελτίωσης της πολιτικής είναι τουλάχιστον δύο[72] (τετραγωνική τάξη σύγκλισης).Αυτό καθιστά την σύγκλισή της πολύ γρήγορη στην πράξη.

4.2 ΣΥΓΚΛΙΣΗ ΤΗΣ ΜΕΘΟΔΟΥ ΕΠΑΝΑΛΗΠΤΙΚΗΣ ΤΙΜΗΣ

Η μέθοδος επαναληπτικής τιμής υπολογίζει μία πολύ καλή προσέγγιση του ελάχιστου ρυθμού μέσου κόστους αφού συγκλίνει με $O(\lambda^n)$, με 0< λ <1 τον παράγοντα προεξόφλησης. Ωστόσο,η σύγκλιση επιτυγχάνεται αργά[73] (όπως είδαμε στο παράδειγμα 3.1- λύση 2) και μάλιστα είναι γραμμική 1ης τάξης με ρυθμό λ. Όσο ο ρυθμός πλησιάζει την μονάδα,η σύγκλιση επιβραδύνεται. Λαμβάνοντας τις παρατηρήσεις αυτές αποδεικνύουμε ένα πιο γενικό θεώρημα σχετικά με την σύγκλιση του εν λόγω αλγορίθμου με κριτήριο την μεγιστοποίηση του ρυθμού αναμενόμενης αμοιβής με παράγοντα προεξόφλησης λ∈(0,1).

Θεώρημα VII:

α) Αν για κάθε i∈S $V_0(i) = 0$, τότε $|V(i) - V_n(i)| \leq \frac{\lambda^{n+1}}{1-\lambda} C$ για κάθε i∈S.

β) για κάθε φραγμένο $V_0(i)$,ισχύει η ομοιόμορφη σύγκλιση ως προς i : $\lim_{n\to\infty} V_n(i) = V(i)$.

Απόδειξη:

Αρχικά, υπενθυμίζουμε ότι

$$\mathrm{V_n}(i,R) = \mathrm{E}_R[\sum_{k=0}^{n-1} \lambda^k C(X_k, a_k)|X_0 = i\ ]$$

0<λ<1 και στόχος είναι η μεγιστοποίηση του g(R) = $\lim_{n\to\infty} \frac{\mathrm{V}_n(i,R)}{n}$, που προσπαθούμε να επιτύχουμε μέσω της ακολουθίας

$$\mathrm{V_n}(i) = \max_{\alpha\in A(i)}\{\ C(i,\alpha) + \sum_{j\in S} p_{ij}(\alpha)\cdot \mathrm{V_{n-1}}(j)\},$$

για κάθε i∈ S.

Συμβολίζουμε με $V(i) := \max_a\ \mathrm{V}_\alpha(i)$.

Επίσης,$|E_R[\alpha\mu o\iota\beta\acute{\eta}\ \gamma\iota\alpha\ \tau\iota\varsigma\ \varepsilon\pi\acute{o}\mu\varepsilon\nu\varepsilon\varsigma\ n+1\ \chi\rho o\nu\iota\kappa\acute{\varepsilon}\varsigma\ \sigma\tau\iota\gamma\mu\acute{\varepsilon}\varsigma|X_0 = j]|$=

$|E_R[\sum_{k=n+1}^{\infty} \lambda^k C(X_k, a_k)\ |X_0 = j]| \leq \frac{\lambda^{n+1}}{1-\lambda} C$ (4.2.1) αφού οι αμοιβές είναι φραγμένες από το C και 0<λ<1 , άρα $\sum_{k=0}^{\infty} \lambda^k = \frac{1}{1-\lambda}$.

[72] Βλ.Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 182.

[73] Βλ.Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming, σελ. 160-165.

α) Αφού $V_0(i) = 0$, τότε αρχίζοντας από την κατάσταση i η ποσότητα $\mathrm{V_n}(i)$ εκφράζει την μέγιστη αναμενόμενη αμοιβή στα n βήματα. Οπότε για την βέλτιστη πολιτική R ισχύει εξ ‘ ορισμού της $\mathrm{V_n}(i)$ σε συνδυασμό με την (4.2.1):

$V(i) = E_R(\text{αμοιβή στα πρώτα n στάδια}) + E_R(\text{υπόλοιπη αμοιβή}) \leq \mathrm{V_n}(i) + \frac{\lambda^{n+1}}{1-\lambda} C$ (4.2.2).

Αντίθετα, $V(i) \geq E_R(\text{αμοιβή στα πρώτα n στάδια}) +$

$E(\textit{επιπλέον αμοιβές από τα } n+1 \textit{ επόμενα βήματα}) \geq \mathrm{V_n}(i) - \frac{\lambda^{n+1}}{1-\lambda} C$ (4.2.3).

αφού εξ’ ορισμού του $V(i)$, είναι μεγαλύτερο της αναμενόμενης αμοιβής που θα προέκυπτε αν εφαρμοζόταν η βέλτιστη πολιτική R στα πρώτα n στάδια και κατόπιν μία οποιαδήποτε πολιτική.

Οι (4.2.2) και (4.2.3) αποδεικνύουν το ζητούμενο.

β) Έστω $V_n^0(i) := V_n(i)$ για εκείνα τα i για τα οποία $V_0(i) = 0$. Τότε σχετικά εύκολα προκύπτει ότι

$$|V_n(i) - V_n^0(i)| \leq \lambda^n \max_j |V_n^0(i)|$$

η οποία σε συνδυασμό με το (α) αν πάρουμε όρια n→∞ έπεται ομοιόμορφη σύγκλιση

$V_n(i) \to V(i)$,οπότε η μέθοδος επαναληπτικής τιμής συγκλίνει .

ΒΙΒΛΙΟΓΡΑΦΙΑ

1.ΕΛΛΗΝΙΚΗ

1.1) Κολέτσος, Ι. και Στογιάννης, Δ. (2012). Εισαγωγή στην Επιχειρησιακή Έρευνα. Εκδόσεις Συμεών, Αθήνα.

1.2)Λουλάκης, Μ.(2015). Στοχαστικές Διαδικασίες. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, Εθνικό Μετσόβιο Πολυτεχνείο, Ζωγράφου.

1.3) Κούτρα , Μ.(2004). Εισαγωγή στις Πιθανότητες , Θεωρία και Εφαρμογές, μέρος Ι. Εκδόσεις Σταμούλη, Αθήνα.

1.4) Φακίνου, Δ. (2007). Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα, Θεωρία και Ασκήσεις. Εκδόσεις Συμμετρία, Αθήνα .

1.5) Φακίνου, Δ. (2012). Εισαγωγή στις Πιθανότητες και τις Στοχαστικές Διαδικασίες. Εκδόσεις Συμμετρία, Αθήνα .

1.6) Φακίνου, Δ. και Οικονόμου, Α. (2003). Εισαγωγή στην Επιχειρησιακή Έρευνα, Θεωρία και Ασκήσεις. Εκδόσεις Συμμετρία, Αθήνα.

2.ΞΕΝΟΓΛΩΣΣΗ

2.1) Hoel, P., Port, S. , Stone,C. (2002). «Εισαγωγή στη Θεωρία Πιθανοτήτων», Πανεπιστημιακές Εκδόσεις Κρήτης.

2.2) Puterman, M. L. (2005). Markov Decision Processes, Discrete Stochastic Dynamic Programming. John Wiley and Sons, New Jersey.

2.3)Ravidran, R.A.(2007). Operations Research and Management Science Handbook.CRC Press.

2.4) Ross, S. M. (1983). Introduction to Stochastic Dynamic Programming. Academic Press, San Diego, California.

2.5) Tijms, H. C. (1986). Stochastic Modelling and Analysis: A Computational Approach. John Wiley and Sons, New York.