ΚΕΦΑΛΑΙΟ 5

Επαναληπτική μέθοδος πολιτικής για προβλήματα POMDP σε άπειρο χρονικό ορίζοντα

Περίληψη

Στο προηγούμενο κεφάλαιο μελετήσαμε προσεγγίσεις της άριστης συνάρτησης τιμών και της άριστης πολιτικής για ένα πρόβλημα POMDP με την επαναληπτική μέθοδο τιμών (value iteration). Στο κεφάλαιο αυτό, παρουσιάζεται η επαναληπτική μέθοδος πολιτικής (policy-iteration),με την οποία σε κάθε επανάληψη επέρχεται βελτίωση της πολιτικής.

Στην ενότητα 5.1 παρουσιάζουμε την θεωρητική βάση της μεθόδου policy-iteration.Έχοντας υπόψη ότι για το πρόβλημα POMDP σε άπειρο χρονικά ορίζοντα υπάρχει μη τυχαιοποιημένη στάσιμη (nonrandomized stationary) άριστη πολιτική, θεωρούμε μη τυχαιοποιημένες στάσιμες πολιτικές. Μία εγγενής δυσκολία στην υπολογιστική διαδικασία είναι ότι ο χώρος των δ.π ,Π, είναι συνεχής. Κατά το στάδιο policy evaluation στο οποίο υπολογίζουμε την συνάρτηση τιμών μιας πολιτικής, το πρόβλημα απλουστεύεται σημαντικά αν η πολιτική επάγει μία Μαρκοβιανή διαμέριση του χώρου Π. Στην περίπτωση αυτή η συνάρτηση τιμών είναι κατά τμήματα γραμμική και ο υπολογισμός της ανάγεται στην επίλυση ενός γραμμικού συστήματος εξισώσεων. Στις επόμενες ενότητες του κεφαλαίου αυτού παρουσιάζουμε ικανές συνθήκες , που εξασφαλίζουν ότι μια πολιτική επάγει Μαρκοβιανή διαμέριση του χώρου Π.

Στην ενότητα 5.2, παρουσιάζουμε μια ικανή συνθήκη που προτάθηκε από τον Sondik [120]:η πολιτική να είναι πεπερασμένα μεταβατική (finitely transient).

Στην ενότητα 5.3,προτείνουμε μια διαφορετική ικανή συνθήκη:η πολιτική να είναι περιοδική. Επιπλέον το βασικό αποτέλεσμα αυτής της ενότητας είναι η διατύπωση μιας γενικότερης συνθήκης από την πεπερασμένη μεταβατικότητα, και περιοδικότητα ,που όπως αποδεικνύουμε επάγει Μαρκοβιανή διαμέριση.

5.1. Εισαγωγή

Μία άλλη μέθοδος για επίλυση του προβλήματος μιας μερικά παρατηρήσιμης Μαρκοβιανής διαδικασίας POMDP σε άπειρο χρονικό ορίζοντα (εκτός από την μέθοδο των επαναληπτικών τιμών), είναι η μέθοδος επαναληπτικής βελτίωσης μιας πολιτικής *(policy-iteration)*.Η μέθοδος αυτή παρουσιάσθηκε για την περίπτωση μιας Μαρκοβιανής διαδικασίας αποφάσεων MDP με πεπερασμένο πλήθος καταστάσεων και αποφάσεων (βλέπε παράγραφο 1.1). Στην Μαρκοβιανή διαδικασία αποφάσεων MDP, η παραπάνω επαναληπτική μέθοδος στα πλαίσια του βήματος βελτίωσης της πολιτικής *(policy-improvement)*,βελτιώνει την πολιτική σε κάθε επανάληψη και τερματίζεται μετά από πεπερασμένο αριθμό επαναλήψεων με μια βέλτιστη πολιτική. Επέκταση αυτής της επαναληπτικής μεθόδου έχει γίνει από τον Blackwell [15],αλλά μονάχα από θεωρητική σκοπιά, ο βασικός ωστόσο αλγόριθμος οφείλεται στον Sondik [120]. Αυτός χρησιμοποίησε μια ανάλογη προσέγγιση για την εύρεση μιας βέλτιστης πολιτικής σε μια μερικά παρατηρήσιμη Μαρκοβιανή διαδικασία αποφάσεων POMDP σε άπειρο χρονικό ορίζοντα.

Η επαναληπτική μέθοδος πολιτικής *(policy iteration)* περιλαμβάνει δύο στάδια Στο πρώτο στάδιο *(policy evaluation)* υπολογίζεται η συνάρτηση τιμών μίας στάσιμης πολιτικής δ^{∞}.Στο δεύτερο στάδιο *(policy-improvement)* αυτή η συνάρτηση τιμών χρησιμοποιείται ως εφαλτήριο για να βρούμε μια πολιτική με μεγαλύτερη συνάρτηση τιμών (βελτιωμένη πολιτική). Η επαναληπτική διαδικασία συνεχίζεται, μέχρις ότου μια διαφορετική πολιτική με μεγαλύτερη συνάρτηση τιμών να μην μπορεί πλέον να βρεθεί. *Η τελευταία αυτή πολιτική είναι τώρα βέλτιστη.*

Θα εξετάσουμε πρώτα το στάδιο *(policy evaluation).*Έστω δ^{∞} μία (μη τυχαιοποιημένη) στάσιμη πολιτική και $V(\pi/\delta),\ \pi\in\Pi$ η αντίστοιχη συνάρτηση του αναμενόμενου ολικού εκπίπτοντος κέρδους για άπειρο χρονικό ορίζοντα (συνάρτηση τιμών της δ^{∞}). Επειδή ο τελεστής H_{δ} (βλέπε κεφάλαιο 1) είναι συστολή modulus β (όπου β είναι ο συντελεστής έκπτωσης, $0 \le \beta < 1$), με μοναδικό σταθερό σημείο τη συνάρτηση $V(./\delta)$, η προσέγγιση της $V(./\delta)$ μπορεί να γίνει με επαναληπτική εφαρμογή του τελεστή H_{δ}:

$$V_n(\pi/\delta) = H_{\delta}V_{n-1}(\pi/\delta), \pi \in \Pi,\ \ n \ge 1,$$

όπου $V_0(\pi/\delta), \pi \in \Pi$ είναι τυχούσα φραγμένη συνάρτηση. Η ακολουθία $\{V_n(./\delta)\}$ συγκλίνει όταν $n \to \infty$ στην συνάρτηση $V(./\delta)$ ανεξάρτητα από την επιλογή της αρχικής συνάρτησης $V_0(./\delta)$ και η σύγκλιση είναι ομαλή στο χώρο Π. Όμως η παραπάνω επαναληπτική μέθοδος δύσκολα εφαρμόζεται επειδή δεν υπάρχουν αλγόριθμοι του ενός βήματος για τον τελεστή H_{δ} ανάλογοι με εκείνους για τον τελεστή H που περιγράψαμε στα κεφάλαια 2 και 3. Αυτό οφείλεται στο γεγονός ότι ο τελεστής H_δ δεν διατηρεί την κατά τμήματα γραμμικότητα και κυρτότητα. Η δυσκολία αυτή παρακάμπτεται με μία διαδικασία που προτάθηκε από τον Sondik [120]. Η *V(./δ)*προσεγγίζεται μέσω κατά τμήματα γραμμικών συναρτήσεων με οποιαδήποτε επιθυμητή ακρίβεια και βασίζεται σε θεμελιώδεις δομικές ιδιότητες της POMDP μάλλον παρά σε ιδιότητες σύγκλισης που σχετίζονται με τη συστολή του τελεστή H_δ.Στην περίπτωση που η πολιτική δ^∞ επάγει Μαρκοβιανή διαμέριση του χώρου Π τότε η *V(./δ)* υπολογίζεται επακριβώς μέσω της επίλυσης ενός γραμμικού συστήματος εξισώσεων (ενότητα 5.2). Επιπλέον στην ενότητα 5.3 γενικεύουμε την συνθήκη του Sondik που εξασφαλίζει Μαρκοβιανή διαμέριση.Το στάδιο ***policy-improvement*** στηρίζεται στο ακόλουθο θεώρημα

***Θεώρημα 5.1.1*: (*Howard-Blackwell* *policy improvement*)** [16].

Έστω *V(./δ) η συνάρτηση τιμών για μια πολιτική* δ^∞ *και* V^* *η άριστη συνάρτηση τιμών. Αν* δ^1 *είναι η συνάρτηση ελέγχου, που* για κάθε $\pi \in \Pi$ *επιλέγει απόφαση που μεγιστοποιεί την*

$$H_\alpha\,[V(./\delta)]=\pi.q^\alpha+\beta.\sum_\theta\{\theta/\pi,a\}.V\,[T(\pi,\theta,\alpha)/\delta]\ ,\alpha\in A$$

δηλαδή : $$\delta^1(\pi)=\arg\max_a H_\alpha\,[V(./\delta)],\ \pi\in \Pi.$$

Τότε : $$V(\pi/\delta^1)\geq V(\pi/\delta),\ \forall\,\pi\in\Pi.$$

Επιπλέον, αν $V(./\delta)\neq V^*$*, τότε υπάρχει κάποιο δ.π ,π , ώστε:*

$$V(\pi/\delta^1)>V(\pi/\delta).$$

Επισημαίνουμε ότι το παραπάνω θεώρημα δεν εγγυάται την εύρεση άριστης πολιτικής σε πεπερασμένο αριθμό επαναλήψεων, που οφείλεται στο γεγονός ότι ο χώρος Π των δ.π είναι συνεχής,*οπότε* το σύνολο των (μη τυχαιοποιημένων) στάσιμων πολιτικών *είναι μη αριθμήσιμο*.Αφού λοιπόν, δεν υπάρχει εγγύηση σύγκλισης σε πεπερασμένο αριθμό επαναλήψεων,μπορούμε να περιοριστούμε στην αναζήτηση μιας $\varepsilon - άριστης$ πολιτικής ($\varepsilon > 0$).Το κριτήριο τερματισμού της *policy iteration* στηρίζεται τότε στο κατάλοιπο Bellman *μιας πολιτικής* δ^{∞} *(Bellman –residual), που ορίζεται ως η μέγιστη απόλυτη διαφορά των συναρτήσεων* $V(./\delta)$ *και* $HV(./\delta)$, *δηλαδή η ποσότητα:*

$$\left\| HV(./\delta) - V(./\delta) \right\|$$

Αν το κατάλοιπο Bellman *για την πολιτική* δ^{∞} είναι αρκούντως μικρό:

$$\left\| HV(./\delta) - V(./\delta) \right\| \le \varepsilon.(1-\beta),$$

τότε η πολιτική δ^{∞} *είναι* $\varepsilon - άριστη$ και η υπολογιστική διαδικασία τερματίζεται. Πράγματι από το λήμμα 4.1.1 (iii) παίρνουμε:

$$\left\| V^{*} - V(./\delta) \right\| \le \frac{1}{1-\beta} \left\| \mathrm{H}V(./\delta) - V(./\delta) \right\| \le \varepsilon .$$

Επισημαίνουμε μία δυσκολία που εμφανίζεται στο στάδιο ***policy improvement*** και σχετίζεται με την εφαρμογή του τελεστού H στην συνάρτηση $V(./\delta)$, που γενικά δεν είναι κατά τμήματα γραμμική και κυρτή. Οι αλγόριθμοι όμως του ενός βήματος των κεφαλαίων 2 και 3 εφαρμόζονται μόνο σε συναρτήσεις που είναι κατά τμήματα γραμμικές και κυρτές (p.w.l.c.). Η δυσκολία αυτή αντιμετωπίσθηκε από τον Sondik [120] με την αντικατάσταση της $V(./\delta)$ ή μιας κατά τμήματα γραμμικής προσέγγισής της από την κυρτή θήκη της (convex hull).

Στο σημείο αυτό θέλουμε να επισημάνουμε, ότι αρκετοί συγγραφείς όπως οι Kakalic και Eckles [56] παρακάμπτουν τις παραπάνω δυσκολίες προσεγγίζοντας τον χώρο Π με ένα σύνολο σημείων (grid of points).Ωστόσο οι υπολογιστικές δυσκολίες μιας τέτοιας προσέγγισης είναι τεράστιες.

5.2.Μαρκοβιανή διαμέριση και πεπερασμένα μεταβατικές πολιτικές

Από το γεγονός ότι η συνάρτηση τιμών μιας πολιτικής δ^{∞},V(./δ),είναι το μοναδικό σταθερό σημείο του τελεστού H_{δ}, αποδεικνύεται εύκολα το ακόλουθο λήμμα

Λήμμα 5.2.1: *Η V(π/δ) μπορεί να γραφεί σαν*

$$V(\pi/\delta)=\pi.\gamma(\pi/\delta),\ \pi\in\Pi,$$

όπου γ(π/δ) είναι ένα διάνυσμα-στήλη $N\times 1$ *(όπου N το πλήθος καταστάσεων του συστήματος) και μάλιστα η μοναδική φραγμένη λύση της διανυσματικής εξίσωσης*

$$\gamma(\pi/\delta)=q^{\delta(\pi)}+\beta.P^{\delta(\pi)}\sum_{\theta}R_{\theta}^{\delta(\pi)}.\gamma[T(\pi,\theta,\delta)/\delta],\ \pi\in\Pi. \qquad \underline{\textbf{5.2.1}}$$

(απόδειξη Sondik)[120]. □

Βέβαια το παραπάνω λήμμα, δεν εκφράζει ότι η V(./δ) είναι κατά τμήματα γραμμική συνάρτηση. Στη συνέχεια θα μελετήσουμε συνθήκες κάτω από τις οποίες η συνάρτηση τιμών V(./δ) είναι κατά τμήματα γραμμική συνάρτηση (p.w.l.).

Ορισμός 5.2.1:*Έστω* δ^{∞} *μια στάσιμη πολιτική. Μια διαμέριση* **V**=[V_1,V_2,…] *του χώρου Π, λέγεται* ***Μαρκοβιανή διαμέριση*** *επαγόμενη από την* δ^{∞} *αν ικανοποιούνται οι παρακάτω ιδιότητες**:***

α)*Σε όλα τα δ.π (information-vectors) που ανήκουν στο ίδιο κελλί αντιστοιχεί η ίδια απόφαση μέσω της δ.*

β)*Μέσω της συνάρτησης μεταφοράς Τ (.,θ,δ) όλα τα δ.π που ανήκουν σε κάποιο κελλί της διαμέρισης έστω* V_j*, απεικονίζονται στο ίδιο κελλί* $V_{j'}$ *, όπου ο δείκτης j΄, εξαρτάται αποκλειστικά από τα j και θ, j′ =* ν*(j,θ).* Η συνάρτηση ν*(j,θ) καλείται Μαρκοβιανή απεικόνιση (Markov mapping) επαγόμενη από την* δ^{∞} .

Σχετικά με την ιδιότητα (α), η απόφαση που αντιστοιχεί στο κελλί V_j μέσω της δ συμβολίζεται με δ_j :

$$\delta(\pi) \equiv \delta_j \quad \forall\, \pi \in V_j .$$

Σχετικά με την ιδιότητα (β), η σχέση ανάμεσα στα σύνολα V_j της διαμέρισης **V** μέσω της συνάρτησης μεταφοράς *Τ (.,θ,δ) παρέχεται από την Μαρκοβιανή απεικόνιση* ν *(j,θ)*:

$$\pi \in V_j \Rightarrow T\,(\pi,\theta,\delta) \in V_{\nu(j,\theta)} .$$

Η Μαρκοβιανή απεικόνιση ν *και ή Μαρκοβιανή διαμέριση* **V***, πού επάγεται από την πολιτική* δ^{∞}*, λέμε ότι είναι **«ισοδύναμες μέσω της δ».***

Σημειώνουμε ότι στην περίπτωση που η πολιτική δ^{∞} ***επάγει Μαρκοβιανή διαμέριση*** **V***=[$V_1, V_2, \ldots V_m$]* ***η συνάρτηση τιμών*** $V(./\delta)$ ***είναι κατά τμήματα γραμμική (p.w.l) και υπολογίζεται μέσω επίλυσης ενός γραμμικού συστήματος εξισώσεων, που έχει πάντοτε λύση.***

Πράγματι από το λήμμα 5.2.1 έχουμε:

$$V(\pi/\delta) = \pi.\gamma(\pi/\delta),$$

όπου το διάνυσμα-στήλη γ(π/δ) είναι η μοναδική φραγμένη λύση της διανυσματικής εξίσωσης (5.2.1) .

Σε κάθε κελλί της διαμέρισης αντιστοιχούμε σταθερό διάνυσμα γ(./δ) και θέτουμε:

$$\gamma(\pi/\delta) \equiv \gamma_j \quad \forall\, \pi \in V_j \,.$$

Επειδή για κάθε $\pi \in V_j$ έχουμε $\delta(\pi)=\delta_j$ και $\gamma(T(\pi,\theta,\delta)/\delta)=\gamma_{\nu(j,\theta)}$,$\theta \in \Theta$, η εξίσωση (5.2.1) ανάγεται στο ακόλουθο σύστημα διανυσματικών εξισώσεων

$$\gamma_j = q^{\delta_j} + \beta . \sum_{\theta} P^{\delta_j} . R_{\theta}^{\delta_j} . \gamma_{\nu(j,\theta)} \,,\; 1 \le j \le m \,. \qquad \textbf{\underline{5.2.2}}$$

Οι λύσεις για τα γ_j μέσω των σχέσεων (5.2.2) είναι μοναδικές (πρόκειται για vectors-contraction), και ισχύει:

$$V(\pi/\delta) = \pi . \gamma_j \qquad \forall\, \pi \in V_j . \qquad \textbf{\underline{5.2.3}}$$

Τίθεται τώρα το ερώτημα,πότε μια στάσιμη πολιτική δ^{∞} ***επάγει μια Μαρκοβιανή διαμέριση.*** Ο Sondik [120] έδωσε ικανή συνθήκη, ώστε μια στάσιμη πολιτική δ^{∞},να επάγει μια πεπερασμένη Μαρκοβιανή διαμέριση του χώρου Π. *Μια πολιτική, που ικανοποιεί αυτή τη συνθήκη λέμε ότι είναι, πεπερασμένα μεταβατική, σύντομα f.t (finitely transient).*Χονδρικά ως *πεπερασμένα μεταβατική* πολιτική, ορίζεται μια στάσιμη πολιτική ,δ^{∞},για την οποία οποιοδήποτε δ.π δεν αποτελεί σημείο ασυνέχειας της συνάρτηση ελέγχου δ μετά από πεπερασμένο αριθμό χρονικών περιόδων. Έτσι τα δ.π στα οποία η συνάρτηση ελέγχου είναι ασυνεχής είναι ουσιαστικά «μεταβατικά» (*transient).*Για να ορίσουμε την έννοια αυτή αναλυτικότερα εισάγουμε τους ακόλουθους συμβολισμούς.

Αν δ^{∞} είναι μια στάσιμη πολιτική και $B \subseteq \Pi$,ορίζουμε:

$$T_\delta(B) := \text{κλειστή θήκη}\{T(\pi, \theta, \delta): \pi \in B,\ \theta \in \Theta\}.$$

Με άλλα λόγια το $T_\delta(B)$ είναι το ελάχιστο κλειστό σύνολο που περιέχει τις δυνατές μεταβάσεις των δ.π. του συνόλου Β την επόμενη χρονική περίοδο εφαρμόζοντας την πολιτική δ^∞.

Ορίζουμε επίσης : $S_\delta^0 := \Pi$

και

$$S_\delta^n := T_\delta(S_\delta^{n-1}),\ n \geq 1,$$

Το S_δ^n περιέχει όλα τα δυνατά δ.π, στην n$^{\text{στη}}$ χρονική περίοδο μετά την έναρξη της λειτουργίας του συστήματος εφαρμόζοντας την πολιτική δ^∞.

Αποδεικνύεται επαγωγικά ότι :

$$S_\delta^n \subset S_\delta^{n-1},\ n \geq 1. \qquad \textbf{\underline{5.2.4}}$$

Θεωρούμε

$$D_\delta := \text{κλειστή θήκη}\ \{\pi: \delta(\pi)\ \text{είναι ασυνεχής στο}\ \pi\}.$$

Δηλαδή ,D_δ, είναι το ελάχιστο κλειστό σύνολο που περιέχει τις ασυνέχειες της συνάρτησης ελέγχου δ.

Ορίζουμε την ακολουθία των συνόλων D^n *, n=0,1,2…ως εξής:*

$$D^0 := D_\delta$$

$$D^{n+1} := \{\pi: T(\pi,\theta,\delta) \in D^n\ \text{για κάποιο}\ \theta\}\ \ n \geq 0\,.$$

Το σύνολο D^n *αναφέρεται σαν σύνολο ασυνεχειών n-τάξης της* δ^∞.

Ορισμός5.2.2:*Μια στάσιμη πολιτική* δ^∞ *είναι πεπερασμένα μεταβατική(finitely-transient) ,σύντομα f.t. αν υπάρχει ένας ακέραιος n , ώστε :*

$$D_\delta \cap S_\delta^n = \varnothing. \qquad \textbf{\underline{5.2.5}}$$

Ο μικρότερος ακέραιος, που πληροί την (5.2.5) σημειώνεται με n_δ και ονομάζεται δείκτης (index) της παραπάνω πολιτικής.

Αν η στάσιμη πολιτική, δ^∞,είναι πεπερασμένα μεταβατική με index n_δ τότε από την (5.2.4) προκύπτει άμεσα ότι:

$$D_\delta \cap S_\delta^n = \varnothing \qquad \forall\ n \geq n_\delta.$$

Λήμμα 5.2.2: *Η πολιτική δ^∞ είναι πεπερασμένα μεταβατική f.t. με δείκτη ,n_δ, αν και μόνον αν D^{n_δ} είναι το πρώτο κενό σύνολο στην ακολουθία $D^0, D^1, D^2, \ldots$,*

Απόδειξη Sondik [120].

Από τα παραπάνω συνάγεται ότι:αν μια πολιτική δ^∞ είναι *f.t.* με δείκτη $n_\delta \geq 1$ τότε **:** $D^n \neq \varnothing$, $n < n_\delta$ και $D^n = \varnothing$, $n \geq n_\delta$.

Θεώρημα5.2.1:*Αν δ^∞ είναι f.t.,τότε αυτή επάγει* πεπερασμένη Μαρκοβιανή διαμέριση και η συνάρτηση,*$V(\pi/\delta)$, είναι κατά τμήματα γραμμική συνάρτηση (p.w.l).*

(απόδειξη Sondik) [120].

Παρατηρήσεις

1)Αν η συνάρτηση ελέγχου δ είναι σταθερή τότε προφανώς η δ δεν έχει σημεία ασυνέχειας και $D^\delta = \varnothing$. Η πολιτική δ^∞ είναι τετριμμένα *f.t.* με δείκτη $n_\delta=0$. Η Μαρκοβιανή διαμέριση που επάγεται από την δ^∞ είναι τετριμμένη:αποτελείται από ένα μόνο σύνολο, τον χώρο Π.

2)Τα σύνορα ανάμεσα στα σύνολα (κελλιά) της Μαρκοβιανής διαμέρισης του χώρου Π που επάγεται από μία f.t . πολιτική δ^{∞} με δείκτη $n_{\delta} \geq 1$ σχηματίζονται από το σύνολο $\bigcup_{k=0}^{n_{\delta}-1} D^{k}$.

3)Το αντίστροφο του παραπάνω θεωρήματος 5.2.1, δεν αληθεύει, όπως αποδεικνύεται με αντιπαράδειγμα Sondik)[117].Επομένως αν έχουμε μια συνάρτηση τιμών,V(π/δ), που είναι κατά τμήματα γραμμική, δεν σημαίνει ότι η πολιτική που την επάγει είναι πεπερασμένα μεταβατική.

Εφαρμογή 5.2.1: (πηγή Sondik [117]). Θεωρούμε ένα πρόβλημα POMDP με δύο καταστάσεις, δύο μηνύματα και δύο αποφάσεις . S={1,2}, Θ={1,2}, A={1,2} με πίνακες μετάβασης και μηνυμάτων και διανύσματα άμεσου κέρδους:

$$
\begin{array}{cccc}
 & P^{\alpha} & R^{\alpha} & q^{\alpha} \\
\alpha=1 & \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{pmatrix} & \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} & \begin{pmatrix} 4 \\ -4 \end{pmatrix} \\
 & & & \\
\alpha=2 & \begin{pmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{pmatrix} & \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix} & \begin{pmatrix} 0 \\ -3 \end{pmatrix}
\end{array}
$$

$$
\delta(\pi) = \begin{cases} 1 & 0 \leq \pi_1 \leq 0.7 \\ 2 & 0.7 < \pi_1 \leq 1 \end{cases} \qquad \beta=0.9
$$

Πηγή Sondik [117]

Τα δ.π $\pi = (\pi_1, \pi_2 = 1-\pi_1)$ καθώς και η συνάρτηση μεταφοράς

$T(\pi,\theta,\alpha) = (T_1(\pi,\theta,\alpha), 1-T_1(\pi,\theta,\alpha))$ θα εξετάζονται ως προς την πρώτη συνιστώσα. Από τον τύπο Bayes (1.4.4) παίρνουμε:

i) $T_1(\pi,\theta=1,\delta(\pi)=1)=\dfrac{0.4+0.24.\pi_1}{0.7+0.06.\pi_1},\quad 0\le\pi_1\le 0.7,$

γνήσια αύξουσα, με πεδίο τιμών $W_1=\left[0.5714, 0.7655\right]$.

ii) $T_1(\pi,\theta=2,\delta(\pi)=1)=\dfrac{0.1+0.06.\pi_1}{0.3-0.06.\pi_1},\quad 0\le\pi_1\le 0.7,$

γνήσια αύξουσα, με πεδίο τιμών $W_2=\left[0.3333, 0.5504\right]$.

iii) $T_1(\pi,\theta=1,\delta(\pi)=2)=\dfrac{0.56+0.42.\pi_1}{0.64-0.18.\pi_1},\quad 0.7<\pi_1\le 1,$

γνήσια φθίνουσα, με πεδίο τιμών $W_3=\left[0.3043, 0.5175\right)$.

iv) $T_1(\pi,\theta=2,\delta(\pi)=2)=\dfrac{0.24-0.18.\pi_1}{0.36+0.18.\pi_1},\quad 0.7<\pi_1\le 1,$

γνήσια φθίνουσα, με πεδίο τιμών $W_4=\left[0.1111, 0.2346\right)$.

Προφανώς $D^0 = D_\delta = \{0.7\}$.

Για να βρούμε το σύνολο ασυνεχειών $1^{\eta s}$ τάξης D^1 της δ ,εξετάζουμε τις εξισώσεις

$$T_1(\pi,\theta,\delta)= 0.7.$$

Επειδή $0.7\in W_1$ ενώ $0.7\notin W_2$, $0.7\notin W_3$, $0.7\notin W_4$, μόνο η εξίσωση

$$T_1(\pi,\theta=1,\delta(\pi)=1)=\frac{0.4+0.24.\pi_1}{0.7+0.06.\pi_1}=0.7$$ έχει αποδεκτή λύση $\pi_1=0.4545$.

Αρα $D^1=\{0.4545\}$.

Συνεχίζοντας παρόμοια, βρίσκουμε τα σύνολα ασυνεχειών $2^{\eta s}$, $3^{\eta s}$ κλπ. τάξης της δ:

$$D^2=\{0.4167, 0.7957\}$$

$$D^3=\{0.2941, 0.8502\}$$

$$D^4=\varnothing.$$

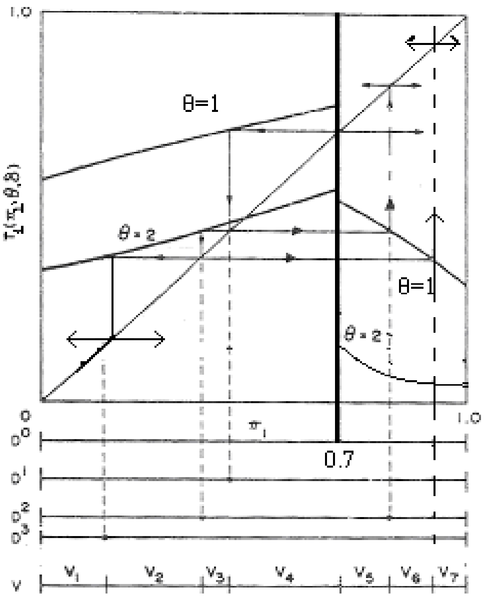
Άρα η στάσιμη πολιτική δ^{∞} είναι πεπερασμένα μεταβατική με δείκτη $n_{\delta} = 4$.Τα σύνολα D^n απεικονίζονται στο σχήμα 5.1. Το σύνολο D^1 βρίσκεται με αντανάκλαση του $D^0=\{0.7\}$,όπως αποδεικνύουν τα βέλη.Παρόμοια με αντανάκλαση του D^1 βρίσκεται το D^2 κ.ο.κ.

Η Μαρκοβιανή διαμέριση του διαστήματος [0,1] που επάγεται από την πολιτική δ^{∞} αποτελείται από 7 σύνολα (κελλιά) με συνοριακά σημεία τα στοιχεία του συνόλου $D^0 \cup D^1 \cup D^2 \cup D^3$ (βλέπε παρατήρηση 2 και σχήμα 5.1):

$$V_1=\left[0, 0.2941\right], V_2=(0.2941, 0.4167\right], V_3=(0.4167, 0.4545\right], V_4=(0.4545, 0.7\right],$$

$$V_5=(0.7, 0.7957\right], V_6=\left[0.7957, 0.8502), V_7=\left[0.8502, 1\right].$$

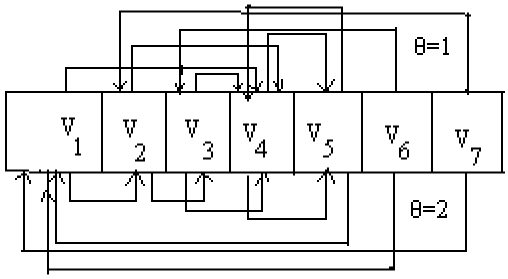
Μαρκοβιανή απεικόνιση



Σύνολο V_j	$v(j,\theta)$	
	θ=1	θ=2
1	4	2
2	4	3
3	4	4
4	5	4
5	4	1
6	3	1
7	2	1

Σχήμα 5.1: Διάγραμμα μεταφοράς στην εφαρμογή 5.2.1.

(πηγή Sondik [117])



Σχήμα 5.2: Διάγραμμα ροής της πολιτικής δ^{∞}.

Αφού η παραπάνω πολιτική δ^{∞} επάγει πεπερασμένη Μαρκοβιανή διαμέριση ***V****=[$V_1, V_2, \ldots V_7$] η συνάρτηση τιμών $V(./\delta)$ είναι κατά τμήματα γραμμική (p.w.l) και υπολογίζεται με επίλυση του παρακάτω πεπερασμένου γραμμικού συστήματος εξισώσεων μέσω του προγράμματος Mathematica 5.*

$$\gamma_1 = q^1 + \beta.P^1.R_1^1.\gamma_4 + \beta.P^1.R_2^1.\gamma_2$$

$$\gamma_2 = q^1 + \beta.P^1.R_1^1.\gamma_4 + \beta.P^1.R_2^1.\gamma_3$$

$$\gamma_3 = q^1 + \beta.P^1.R_1^1.\gamma_4 + \beta.P^1.R_2^1.\gamma_4$$

$$\gamma_4 = q^1 + \beta.P^1.R_1^1.\gamma_5 + \beta.P^1.R_2^1.\gamma_4$$

$$\gamma_5 = q^2 + \beta.P^2.R_1^2.\gamma_4 + \beta.P^2.R_2^2.\gamma_1$$

$$\gamma_6 = q^2 + \beta.P^2.R_1^2.\gamma_3 + \beta.P^2.R_2^2.\gamma_1$$

$$\gamma_7 = q^2 + \beta.P^2.R_1^2.\gamma_2 + \beta.P^2.R_2^2.\gamma_1$$

Βρίσκουμε λοιπόν τα gradients:

$$\gamma_1 = (11.2925, 0.6228)^{\mathrm{T}}, \gamma_2 = (11.2483, 0.5694)^{\mathrm{T}}, \gamma_3 = (11.0372, 0.3783)^{\mathrm{T}}$$

$$\gamma_4 = (9.7884, -0.0588)^{\mathrm{T}}, \gamma_5 = (7.3027, 1.6364)^{\mathrm{T}}, \gamma_6 = (8.0691, 2.2040)^{\mathrm{T}},$$

$$\gamma_7 = (8.2115, 2.3315)^{\mathrm{T}}.$$

$$V(\pi/\delta) = \pi.\gamma_j \ , \quad \pi \in V_j \ , \ j=1,2,3,\ldots,7$$

5.3. Περιοδικές πολιτικές

Στην ενότητα αυτή, παρουσιάζουμε τη συνθήκη της περιοδικότητας η οποία πέραν από τη συνθήκη της πεπερασμένης μεταβατικότητας του Sondik εξασφαλίζει επίσης ότι μία στάσιμη πολιτική επάγει Μαρκοβιανή διαμέριση του χώρου Π.Το πιο σημαντικό όμως είναι ότι υπάρχει γενικότερη συνθήκη από τις δύο που αναφέραμε η οποία όπως θα διαπιστώσουμε στη συνέχεια εξασφαλίζει Μαρκοβιανή διαμέριση. Πρώτα όμως θα δώσουμε την έννοια της διαμέρισης τάξης k που επάγεται από μια στάσιμη πολιτική (βλέπε Sondik [120]).

Έστω δ^{∞} μια στάσιμη πολιτική . Ορίζουμε

$$\bar{D}^n := \bigcup_{i=0}^{n} D^i, n = 0,1,2,...$$

Το $\bar{D}^n$ δηλώνει το σύνολο των δ.π που οδηγούνται σε ασυνέχειες της συνάρτησης ελέγχου δ μέσω της συνάρτησης T (.,.,δ) σε n το πολύ βήματα.

Σημειώνουμε ότι $\bar{D}^0 = D^0 \equiv D_{\delta}$ και $\bar{D}^n \subseteq \bar{D}^m \quad \forall m > n$.

Θεωρούμε την ακολουθία διαμερίσεων του χώρου Π,$\mathbf{V}^0$,$\mathbf{V}^1$,…η οποία κατασκευάζεται ως εξής:Τα σύνορα ανάμεσα στα σύνολα (κελλιά) της αρχικής διαμέρισης $\mathbf{V}^0$ του χώρου Π σχηματίζονται από το σύνολο των ασυνεχειών της συνάρτησης ελέγχου δ, $D^0 = D_{\delta}$.Έτσι η συνάρτηση ελέγχου δ είναι σταθερή (παίρνει ως τιμή την ίδια απόφαση) σε κάθε κελλί της $\mathbf{V}^0$.Λέμε ότι η διαμέριση $\mathbf{V}^0$ επάγεται από την δ^{∞}.

Για $k \geq 1$, η διαμέριση $\mathbf{V}^k$ του χώρου Π κατασκευάζεται έτσι ώστε τα σύνορα ανάμεσα στα κελλιά αυτής να σχηματίζονται από το σύνολο

$$\bar{D}^k := \bigcup_{n=0}^{k} D^n.$$

Παρατηρούμε ότι η διαμέριση **V^k** αποτελεί λέπτυνση της διαμέρισης **V^{k-1}**,δηλαδή για κάθε κελλί της **V^k** υπάρχει κελλί της **V^{k-1}** που το περιέχει. Επομένως συμπεραίνουμε επαγωγικά ότι η συνάρτηση ελέγχου δ είναι σταθερή σε κάθε κελλί της **V^k**. Η **V^k** καλείται διαμέριση τάξης k που επάγεται από την δ^{∞}.

Σημειώνουμε ότι αν $\bar{D}^k \neq \bar{D}^{k-1}$, η διαμέριση **V^k** προκύπτει από την **V^{k-1}** μέσω του συνόλου των ασυνεχειών k τάξης της δ, D^k: τα κελλιά της **V^{k-1}** που τέμνουν το σύνολο D^k διαμερίζονται σε νέα σύνολα (κελλιά) τα οποία διαχωρίζονται με σύνορα από το D^k, ενώ τα κελλιά της **V^{k-1}** που δεν τέμνουν το D^k παραμένουν τα ίδια στη διαμέριση **V^k**. Αν $\bar{D}^k = \bar{D}^{k-1}$, τότε η διαμέριση **V^k** συμπίπτει με την **V^{k-1}**.

Πρόταση 5.3.1: *Έστω **V^k** διαμέριση τάξης k ($k \geq 1$) του χώρου Π που επάγεται από μια στάσιμη πολιτική δ^{∞} και $\theta \in \Theta$. Τότε κάθε κελλί της **V^k** απεικονίζεται εντός κάποιου κελλιού της διαμέρισης **V^{k-1}** μέσω της T (.,θ,δ). Με άλλα λόγια, αν V_i^k ,είναι κελλί της **V^k** ,τότε υπάρχει κελλί της **V^{k-1}**, έστω V_j^{k-1} έτσι ώστε $T(V_i^k, \theta, \delta) \subseteq V_j^{k-1}$.*

Απόδειξη

Συμβολίζουμε με δ^i την απόφαση που παίρνει η δ στο κελλί V_i^k: $\delta^i \equiv \delta(\pi), \pi \in V_i^k$.

Θεωρούμε δύο τυχόντα δ.π. π΄, π΄΄ στο εσωτερικό της περιοχής V_i^k και υποθέτουμε ότι το π΄ απεικονίζεται στο κελλί V_j^{k-1} *της διαμέρισης* ***V^{k-1}*** μέσω της συνάρτησης

$T(.,\theta,\delta^i)$, δηλαδή $T(\pi',\theta,\delta^i)\in V_j^{k-1}$.Θα δείξουμε ότι το δ.π. π΄΄ απεικονίζεται στο ίδιο κελλί, δηλαδή $T(\pi'',\theta,\delta^i)\in V_j^{k-1}$.Διακρίνουμε τις ακόλουθες περιπτώσεις:

ι) Υποθέτουμε ότι το ευθύγραμμο τμήμα

$$l : \lambda.\pi' + (1-\lambda).\pi'' \ ,0 \le \lambda \le 1$$

που συνδέει τα π',π'' περιέχεται εξ ολοκλήρου στο εσωτερικό της περιοχής V_l^k.Επειδή η συνάρτηση μεταφοράς T(.,θ,α), όπου $a\in A$,απεικονίζει ευθύγραμμα τμήματα σε ευθύγραμμα τμήματα (βλέπε πρόταση 1.4.1 ,Κεφ 1), συνάγεται ότι το σύνολο

$$T(l,\theta,\delta^i)=\{T(\pi,\theta,\delta^i) : \pi\in l\}$$

είναι ευθύγραμμο τμήμα με άκρα $T(\pi',\theta,\delta^i)$ και $T(\pi'',\theta,\delta^i)$.Ας θεωρήσουμε ότι το π'' απεικονίζεται σε διαφορετικό κελλί της διαμέρισης $\boldsymbol{V^{k-1}}$ από αυτό που απεικονίζεται το π' μέσω της $T(.,\theta,\delta^i)$, δηλαδή $T(\pi'',\theta,\delta^i)\in V_s^{k-1}$*,όπου* $s\neq j$.Τότε το ευθύγραμμο τμήμα $T(l,\theta,\delta^i)$ διασχίζει κάποιο σύνορο ανάμεσα στις περιοχές V_j^{k-1},V_s^{k-1}. Επομένως υπάρχει $\pi^*\in l$, έτσι ώστε $T(\pi^*,\theta,\delta^i)\in \bar{D}^{k-1}$. Τότε $\pi^*\in \bar{D}^k$ το οποίο σημαίνει ότι το π^* είναι συνοριακό σημείο της περιοχής V_l^k. *Αυτό* όμως αντίκειται στην υπόθεσή μας ότι το ευθύγραμμο τμήμα l περιέχεται στο εσωτερικό της περιοχής V_l^k. Συνεπώς τα δ.π. π',π'' καθώς και το ευθύγραμμο τμήμα l που τα ενώνει, απεικονίζονται μέσω της $T(.,\theta,\delta^i)$ στο ίδιο κελλί της διαμέρισης $\boldsymbol{V^{k-1}}$:

$$T(l,\theta,\delta^i)\subseteq V_j^{k-1}, \quad T(\pi'',\theta,\delta^i)\in V_j^{k-1}.$$

ii) Θεωρούμε ότι το ευθύγραμμο τμήμα l που συνδέει τα π',π'' δεν περιέχεται εξ ολοκλήρου στο εσωτερικό της περιοχής V_l^k. Τότε υπάρχει διαδρομή από ευθ. τμήματα

$l_1, l_2, ..., l_m$ η οποία συνδέει τα π', π'', έτσι ώστε τα $l_1, l_2, ..., l_m$ να περιέχονται εξ ολοκλήρου στο εσωτερικό της περιοχής V_i^k. Συμβολίζουμε με π_t ,π_{t+1} τα άκρα του ευθυγράμμου τμήματος l_t (t=1,2,…,m), όπου $\pi_1 \equiv \pi'$, $\pi_{m+1} \equiv \pi''$.

Επειδή $T(\pi_1, \theta, \delta^i) \in V_j^{k-1}$, από την περίπτωση (i) συνάγεται ότι $T(l_1, \theta, \delta^i) \subseteq V_j^{k-1}$, $T(\pi_2, \theta, \delta^i) \subseteq V_j^{k-1}$.Εφαρμόζοντας διαδοχικά το ίδιο επιχείρημα για τα ευθύγραμμα τμήματα $l_2, ..., l_m$, παίρνουμε:

$$T(l_t, \theta, \delta^i) \subseteq V_j^{k-1},\ T(\pi_{t+1}, \theta, \delta^i) \subseteq V_j^{k-1}, t=2,...,m$$

Άρα $T(\pi'', \theta, \delta^i) \in V_j^{k-1}$ και συμπεραίνουμε ότι το κελλί V_i^k *της διαμέρισης* $\mathbf{V^k}$ απεικονίζεται μέσω της $T(.,\theta, \delta^i)$ εντός του κελλιού V_j^{k-1} *της διαμέρισης* $\mathbf{V^{k-1}}$.□

Πρόταση 5.3.2:

Οι ακόλουθες προτάσεις είναι ισοδύναμες

i) $\bar{D}^{n+1} = \bar{D}^n$ για κάποιο $n \in \mathbb{N}_0$.

*ii)*Υπάρχει $n \in \mathbb{N}_0$ έτσι ώστε $\bar{D}^m = \bar{D}^n \quad \forall m \geq n$.

Απόδειξη

$i \Rightarrow ii)$.As υποθέσουμε ότι το *ιι)* δεν αληθεύει. Θα καταλήξουμε σε άτοπο. Έστω $\bar{D}^m$ το πρώτο σύνολο στην ακολουθία $\bar{D}^{n+2}, \bar{D}^{n+3}, ...,$ το οποίο διαφέρει από το $\bar{D}^n$, δηλαδή $\bar{D}^{m-1} = \bar{D}^n$ και $\bar{D}^n \underset{\neq}{\subset} \bar{D}^m$.

Επομένως υπάρχει $\pi \in \bar{D}^m$ έτσι ώστε $\pi \notin \bar{D}^n$. Τότε υπάρχει $\theta \in \Theta$ έτσι ώστε $T(\pi, \theta, \delta(\pi)) \in \bar{D}^{m-1}$.Συνεπώς $T(\pi, \theta, \delta(\pi)) \in \bar{D}^n$ και $\pi \in \bar{D}^{n+1}$.Αυτό όμως οδηγεί σε αντίφαση επειδή υποθέσαμε $\bar{D}^{n+1} = \bar{D}^n$ και $\pi \notin \bar{D}^n$. Άρα το (ii) αληθεύει.

ιι) $\Rightarrow i)$. Είναι προφανές. □

Θα δώσουμε τώρα τον ορισμό της περιοδικής πολιτικής.

Ορισμός 5.3.1: Μία στάσιμη πολιτική δ^{∞} λέγεται περιοδική αν η ακολουθία των συνόλων $D^0, D^1, D^2 \ldots$ δεν περιέχει κενό σύνολο και υπάρχουν φυσικοί αριθμοί $l \geq 0, s \geq 1$ έτσι ώστε :

$$D^{n+s} = D^n \quad \forall n \geq l .$$

Από τον παραπάνω ορισμό και το λήμμα 5.2.2 συνάγεται ότι μία περιοδική πολιτική αποκλείεται να είναι πεπερασμένα μεταβατική (f.t.) και αντιστρόφως. Με άλλα λόγια οι συνθήκες της περιοδικότητας και της πεπερασμένης μεταβατικότητας είναι αμοιβαία αποκλειόμενες.

Μία γενίκευση των συνθηκών περιοδικότητας και πεπερασμένης μεταβατικότητας είναι η ακόλουθη συνθήκη **(Α)**: *Λέμε ότι μία στάσιμη πολιτική* δ^{∞} *ικανοποιεί τη συνθήκη (Α) αν:*

$$\bar{D}^{n+1} = \bar{D}^n \quad \textit{για κάποιο } n \in \mathbb{N}_0 . \qquad \textbf{\textit{(A)}}$$

Η γενίκευση αυτή προκύπτει από την ακόλουθη πρόταση.

Πρόταση 5.3.3: Αν η πολιτική δ^{∞} είναι πεπερασμένα μεταβατική ή περιοδική τότε η δ^{∞} ικανοποιεί τη συνθήκη **(Α)**.

Απόδειξη

Αν η πολιτική δ^{∞} είναι πεπερασμένα μεταβατική με δείκτη n_δ, τότε το συμπέρασμα έπεται προφανώς από το λήμμα 5.2.2 με $n = n_\delta$. Αν η πολιτική δ^{∞} είναι περιοδική, τότε το συμπέρασμα έπεται από τον ορισμό 5.3.1 με $n = l + s - 1$.

Πράγματι,

$$\bar{D}^{n+1} = \bar{D}^{l+s} = \bigcup_{i=0}^{l+s} D^i = (\bigcup_{i=0}^{l+s-1} D^i) \cup D^{l+s}$$

$$= (\bigcup_{i=0}^{l+s-1} D^i) \cup D^l = \bigcup_{i=0}^{l+s-1} D^i = \bar{D}^{l+s-1} = \bar{D}^n .$$

□

Ορισμός 5.3.2: Εστω δ^∞ μία στάσιμη πολιτική *που ικανοποιεί τη συνθήκη (Α).* Καλούμε δείκτη της δ^∞ τον ελάχιστο $n \in \mathbb{N}_0$ για τον οποίο $\bar{D}^{n+1} = \bar{D}^n$. Ο δείκτης συμβολίζεται με n_δ.

Από την πρόταση 5.3.2 έπεται ότι αν η πολιτική δ^∞ ικανοποιεί την συνθήκη (Α) και έχει δείκτη n_δ, τότε

$$\bar{D}^n = \bar{D}^{n_\delta} \qquad \forall n \geq n_\delta$$

Επομένως η διαμέριση του χώρου Π τάξης n_δ που επάγεται από την δ^∞, $\mathbf{V}^{n_\delta}$, είναι μια τελική διαμέριση, η οποία δεν μπορεί να λεπτυνθεί περαιτέρω. Με άλλα λόγια, όλες οι διαμερίσεις τάξης $k \geq n_\delta$ που επάγονται από την δ^∞ ταυτίζονται με την διαμέριση $\mathbf{V}^{n_\delta}$: $\mathbf{V}^k = \mathbf{V}^{n_\delta}$.

Σημειώνουμε ότι ο δείκτης μιας πεπερασμένης μεταβατικής πολιτικής όπως ορίσθηκε στην ενότητα 5.2 (βλέπε ορισμό 5.2.2 και λήμμα 5.2.2) δεν συμπίπτει με τον δείκτη σύμφωνα με τον ορισμό 5.3.2. Έτσι στην εφαρμογή 5.2.1 ο δείκτης της δ^∞ σύμφωνα με τον ορισμό 5.2.2 είναι $n_\delta = 4$ ενώ σύμφωνα με τον ορισμό 5.3.1 είναι $n_\delta = 3$, επειδή $\bar{D}^2 \neq \bar{D}^3$ και $\bar{D}^3 = \bar{D}^4$.

Το βασικό αποτέλεσμα αυτής της ενότητας δίνεται από την ακόλουθη πρόταση.

Πρόταση 5.3.4: Αν η πολιτική δ^{∞} ικανοποιεί την συνθήκη (Α),τότε αυτή επάγει Μαρκοβιανή διαμέριση του χώρου Π και η συνάρτηση V(π/δ), $\pi \in \Pi$ είναι κατά τμήματα γραμμική.

Απόδειξη

Έστω n_δ ο δείκτης της πολιτικής δ^{∞}. Θα δείξουμε ότι η διαμέριση του χώρου Π τάξεως n_δ που επάγεται από την δ^{∞}, $\mathbf{V}^{n_\delta}$,είναι Μαρκοβιανή. Πράγματι, η διαμέριση $\mathbf{V}^{n_\delta}$ (όπως και όλες οι διαμερίσεις $\mathbf{V}^{k}$ τάξης k=0,1,2… που επάγονται από την δ^{∞}) ικανοποιεί την ιδιότητα (α) του ορισμού 5.2.1: η συνάρτηση ελέγχου δ είναι σταθερή σε κάθε κελλί της διαμέρισης. Θεωρούμε την διαμέριση τάξεως $n_\delta + 1$, $\mathbf{V}^{n_\delta+1}$. Σύμφωνα με την πρόταση 5.3.1 κάθε κελλί της διαμέρισης $\mathbf{V}^{n_\delta+1}$ απεικονίζεται εντός κάποιου κελλιού της διαμέρισης $\mathbf{V}^{n_\delta}$ μέσω της συνάρτησης μεταφοράς $T(.,\theta,\delta)$,όπου $\theta \in \Theta$.Λαμβάνοντας υπόψη την ταύτιση των διαμερίσεων $V^{n_\delta}, V^{n_\delta+1}$ συνάγεται ότι η διαμέριση $\mathbf{V}^{n_\delta}$ ικανοποιεί την ιδιότητα (β) του ορισμού 5.2.1:κάθε κελλί της $\mathbf{V}^{n_\delta}$ απεικονίζεται εντός κάποιου κελλιού της ίδιας διαμέρισης μέσω της $T(.,\theta,\delta)$,όπου $\theta \in \Theta$.Επομένως η διαμέριση $\mathbf{V}^{n_\delta}$ είναι Μαρκοβιανή. Επίσης η συνάρτηση τιμών της δ^{∞}, V(π/δ), $\pi \in \Pi$ είναι κατά τμήματα γραμμική (βλέπε σχέσεις (5.2.2),(5.2.3)). $\square$

Από τις προτάσεις 5.3.3 και 5.3.4 συμπεραίνουμε ότι αν η πολιτική δ^{∞} είναι πεπερασμένα μεταβατική ή περιοδική τότε αυτή επάγει Μαρκοβιανή διαμέριση στον χώρο Π και η συνάρτηση τιμών V(./δ) είναι κατά τμήματα γραμμική. Αριθμητικές εφαρμογές περιοδικών πολιτικών δίνονται σε ένα πρόβλημα αντικατάστασης στο κεφάλαιο 7.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στο κεφάλαιο αυτό, δείξαμε ότι η συνθήκη της πεπερασμένης μεταβατικότητας του Sondik είναι πιο ισχυρή από αυτή που στην πραγματικότητα χρειάζεται για να επιβεβαιώσουμε ότι μία στάσιμη πολιτική επάγει Μαρκοβιανή διαμέριση και παράλληλα η αντίστοιχη συνάρτηση τιμών είναι κατά τμήματα γραμμική .Εναλλακτικά παρουσιάσαμε τη συνθήκη περιοδικότητας, η οποία είναι αμοιβαία αποκλειόμενη από τη συνθήκη πεπερασμένης μεταβατικότητας.

Το πιο βασικό αποτέλεσμα του κεφαλαίου αυτού είναι μία γενίκευση των παραπάνω συνθηκών (**η συνθήκη (Α)**, ενότητα 5.3),η οποία επάγει Μαρκοβιανή διαμέριση και παράλληλα συνάρτηση τιμών κατά τμήματα γραμμική (πρόταση 5.3.4).Αριθμητικές εφαρμογές περιοδικών πολιτικών δίνονται σε ένα πρόβλημα αντικατάστασης στο κεφάλαιο 7.