# Κεφάλαιο 1

# Αλυσίδες Markov

## 1.1 Εισαγωγή και Ορισμός της Αλυσίδας Markov

Έστω $\mathscr{S}$ ένα σύνολο με πεπερασμένο αριθμό στοιχείων, π.χ. $\mathscr{S} := \{0, 1, 2, \ldots, N\}$. Το σύνολο αυτό το ονομάζουμε χώρο καταστάσεων. Θεωρούμε μια διαδικασία που εξελίσσεται σε διακριτό χρόνο και η οποία, σε κάθε χρονική στιγμή, βρίσκεται σε μια από τις καταστάσεις του $\mathscr{S}$. Η διαδικασία μεταπηδά με τυχαίο τρόπο από κατάσταση σε κατάσταση και, δεδομένης της ιστορίας της διαδικασίας, η πιθανότητα μετάβασης από μια κατάσταση σε μια άλλη εξαρτάται μόνο από την παρούσα θέση της.

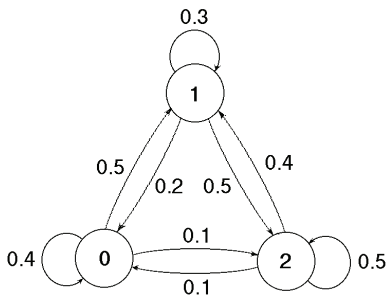
Ας υποθέσουμε για παράδειγμα ότι $\mathscr{S} = \{0, 1, 2\}$. Θεωρούμε ένα σωματίδιο το οποίο μεταπηδά μεταξύ των τριών αυτών καταστάσεων δηλαδή μεταξύ των στοιχείων του συνόλου $\mathscr{S}$ (σχήμα 1). Αν τη χρονική στιγμή $n$ βρίσκεται στην κατάσταση 0 τότε την χρονική στιγμή $n + 1$ θα βρεθεί στην κατάσταση 1 με πιθανότητα 0.5, στην κατάσταση 2 με πιθανότητα 0.1 ή θα παραμείνει στην κατάσταση 0 με πιθανότητα 0.4. Σε κάθε χρονική στιγμή η μεταπήδηση είναι τυχαία και η πιθανότητα μεταπήδησης εξαρτάται από τη θέση του σωματιδίου τη δεδομένη χρονική στιγμή αλλά όχι από την προηγούμενη ιστορία της διαδικασίας.

Για να γίνουμε πιο συγκεκριμένοι, ας υποθέσουμε ότι $X_n$ είναι η θέση του σωματιδίου τη χρονική στιγμή $n$. Οι διαδοχικές θέσεις του σωματιδίου, $\{X_n, n = 0, 1, 2, \ldots, \}$, είναι τυχαίες μεταβλητές. Υποθέτουμε ότι, για κάθε χρονική στιγμή $n$ και κάθε $j, i, i_{n-1}, \ldots, i_0 \in \mathscr{S}$

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i). \qquad (1.1)$$

Η ιδιότητα αυτή ονομάζεται *Μαρκοβιανή ιδιότητα*.

Αν η χρονική στιγμή $n$ συμβολίζει το παρόν, το $n + 1$ το επόμενο, μελλοντικό βήμα, και το ενδεχόμενο $\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-2} = i_{n-2}, X_{n-1} = i_{n-1}\}$ το παρελθόν της διαδικασίας τότε η μαρκοβιανή ιδιότητα μπορεί να εκφραστεί επιγραμματικά ως:




$$P = \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.5 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$$

Σχήμα 1.1: Παράδειγμα αλυσίδας Markov με τρεις καταστάσεις και πίνακα πιθανοτήτων μετάβασης $P$.

**Μαρκοβιανή Ιδιότητα:** *Το μέλλον είναι ανεξάρτητο από το παρελθόν δεδομένου του παρόντος*

Αν επιπλέον υποθέσουμε ότι η πιθανότητες μετάβασης, $P(X_{n+1} = j|X_n = i)$, δεν εξαρτώνται από τη χρονικη στιγμή $n$, δηλαδή αν ισχύει ότι

$$P(X_{n+1} = j|X_n = i) = P(X_1 = j|X_0 = i) := P_{ij}, \qquad \forall n \in \mathbb{N} \tag{1.2}$$

τότε η διαδικασία που προκυπτει είναι μια *χρονικά ομογενής αλυσίδα Markov*. Ο πινακας $P_{ij}$, $i, j, \in \mathscr{S}$, (ένας πίνακας $(N+1) \times (N+1)$ αν το σύνολο καταστάσεων $\mathscr{S}$ έχει $N+1$ στοιχεία) ονομάζεται *πίνακας πιθανοτήτων μετάβασης* και έχει τις εξής ιδιότητες:

1. $P_{ij} \geq 0$ για κάθε $i, j \in \{0, 1, \ldots, N\}$.
2. $\sum_{j=0}^{N} P_{ij} = 1$ για κάθε $i \in \{0, 1, \ldots, N\}$.

Οι πίνακες που έχουν τις δύο αυτές ιδιότητες ονομάζονται στοχαστικοί.

Ας επιστρέψουμε στο παράδειγμα του σχήματος 1. Υποθέτουμε ότι οι μεταπηδήσεις του σωματιδίου *εξαρτώνται από την θέση του αλλά όχι από την προηγούμενη ιστορία του.* Συγκεκριμένα, αν η θέση του σωματιδίου είναι την χρονική στιγμή $n$ είναι $i$ $(i = 0, 1, 2)$ τότε την χρονική στιγμή $n+1$ σωματίδιο θα βρίσκεται στην θέση $j$ με πιθανότητα $P_{ij}$ όπου $P$ είναι ένας τετραγωνικός πίνακας ($3 \times 3$ στο παράδειγμά μας). Εφόσον ο μηχανισμός που διέπει τις μεταπηδήσεις είναι τυχαίος η θέση του σωματιδίου την χρονική στιγμή $n$ είναι τυχαία μεταβλητή που συμβολίζουμε με $X_n$. Από τις παραπάνω υποθέσεις βλέπουμε για παράδειγμα ότι $P(X_1 = 2|X_0 = 0) = P_{02} = 0.1$ και $P(X_1 = 2|X_0 = 1) = P_{12} = 0.5$. Λόγω της υποθέσεώς μας ότι η κάθε μεταπήδηση προσδιορίζεται αποκλειστικά από τον πίνακα πιθανοτήτων μετάβασης $P$ ανεξαρτήτως της ιστορίας της διαδικασίας (δηλαδή των διαδοχικών θέσεων του σωματιδίου) ισχύει επίσης, για παράδειγμα, ότι $P(X_4 = 0|X_3 = 1, X_2 = 0, X_1 = 0, X_0 = 2) = P_{10} = 0.2$. Με άλλα λόγια οι διαδοχικές θέσεις του σωματιδίου ικανοποιούν τον ορισμό της διαδικασίας Markov.

Χάριν ακριβολογίας ας υποθέσουμε ότι η αρχική θέση του σωματιδίου είναι $X_0 = 0$. Βάσει του ορισμού 4.1 η πιθανότητα $P(X_4 = 1, X_3 = 2, X_2 = 2, X_1 = 1|X_0 = 0)$ γράφεται ως

$$\begin{aligned}
&P(X_4 = 1|X_3 = 2, X_2 = 2, X_1 = 1, X_0 = 0)P(X_3 = 2, X_2 = 2, X_1 = 1|X_0 = 0) \\
&\quad = P(X_4 = 1|X_3 = 2)P(X_3 = 2|X_2 = 2)P(X_2 = 2|X_1 = 1)P(X_1 = 1|X_0 = 0) \\
&\quad = P_{21}P_{22}P_{12}P_{01} = 0.4 \times 0.5 \times 0.5 \times 0.5 = 0.05.
\end{aligned}$$

Ο παραπάνω υπολογισμός βασίστηκε αποκλειστικά στον ορισμό της δεσμευμένης πιθανότητας και στην Μαρκοβιανή ιδιότητα.

## 1.2 Εξισώσεις Chapman-Kolmogorov

Ας υπολογίσουμε τώρα την πιθανότητα το σωματίδιο να βρίσκεται στην κατάσταση 1 την χρονική στιγμή 2 δεδομένου ότι ήταν στην κατάσταση 0 την χρονική στιγμή 0, δηλαδή $P(X_2 = 1|X_0 = 0)$. Από το θεώρημα ολικής πιθανότητας, αυτή η πιθανότητα γράφεται ως $\sum_{k=0}^{2} P(X_2 = 1, X_1 = k|X_0 = 0)$. Από τον ορισμό της δεσμευμένης πιθανότητας όμως ο κάθε όρος του αθροίσματος μπορεί να γραφεί σαν ένα γινόμενο $P(X_2 = 1|X_1 = k, X_0 = 0)P(X_1 = k|X_0 = 0)$. Χρησιμοποιώντας τώρα την Μαρκοβιανή ιδιότητα βλέπουμε ότι ο πρώτος όρος του γινομένου παίρνει την μορφή $P(X_2 = 1|X_1 = k)$ και συνεπώς

$$\begin{aligned}
P(X_2 = 1|X_0 = 0) &= \sum_{k=0}^{2} P(X_1 = k|X_0 = 0)P(X_2 = 1|X_1 = k) \\
&= P_{00}P_{01} + P_{01}P_{11} + P_{02}P_{21} \\
&= P_{01}^2 = 0.39
\end{aligned}$$

Βλέπουμε δηλαδή ότι η πιθανότητα να βρισκόμαστε στην κατάσταση $j$ μετά από δύο βήματα, δεδομένου ότι αρχικά ξεκινάμε από την κατάσταση $i$, δίνεται από το στοιχείο $i, j$ ($i$ γραμμή, $j$ στήλη) του πίνακα $P^2$.

Όπως θα δούμε στη συνέχεια η παραπάνω ανάλυση γενικεύεται. Ας συμβολίσουμε με $P_{ij}^{(n)} := P(X_n = j|X_0 = i)$ την πιθανότητα να βρεθούμε στην κατάσταση $j$ την χρονική στιγμή $n$ δεδομένου ότι η διαδικασία ξεκίνησε από την κατάσταση $i$. Θα αποδείξουμε ότι η άγνωστη ποσότητα $P_{ij}^{(n)}$ που εξαρτάται από τρεις δείκτες, τους $i, j$, και $n$, είναι ίση με την $P_{ij}^n$ δηλαδή το $i, j$ στοιχείο του πίνακα $P$ υψωμένου στην $n$–οστή δύναμη. Βάσει των ιδίων επιχειρημάτων που χρησιμοποιήσαμε προηγουμένως

$$P(X_n = j|X_0 = i) = \sum_{k \in \mathscr{S}} P(X_n = j|X_{n-1} = k)P(X_{n-1} = k|X_0 = i)$$

ή

$$P_{ij}^{(n)} = \sum_{k \in \mathscr{S}} P_{ik}^{(n-1)} P_{kj}.$$

Εφαρμόζοντας την παραπάνω σχέση για $n = 2$, $P_{ij}^{(2)} = \sum_{k \in \mathscr{S}} P_{ik}P_{kj}$. Επομένως βλέπουμε ότι ο πίνακας $P^{(2)}$ μπορεί να εκφραστεί σαν το γινόμενο της $P$ με τον εαυτό της και επαγωγικά βλέπουμε ότι η $P^{(n)}$ είναι η $P^n$. Οι διαδοχικές δυνάμεις του πίνακα πιθανοτήτων μετάβασης του παραδείγματός μας είναι

$$P = \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.5 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}, \quad P^2 = \begin{bmatrix} .27 & .39 & .34 \\ .19 & .39 & .42 \\ .17 & .37 & .46 \end{bmatrix}, \quad P^3 = \begin{bmatrix} .22 & .388 & .392 \\ .196 & .38 & .424 \\ .188 & .38 & .432 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} .205 & .383 & .412 \\ .197 & .382 & .422 \\ .194 & .381 & .425 \end{bmatrix}, \quad P^5 = \begin{bmatrix} .200 & .382 & .418 \\ .197 & .382 & .421 \\ .196 & .381 & .422 \end{bmatrix}, \quad P^6 = \begin{bmatrix} .198 & .382 & .420 \\ .197 & .382 & .421 \\ .197 & .382 & .421 \end{bmatrix}$$

Παρατηρείστε ότι καθώς η δύναμη στην οποία υψώνεται ο πίνακας μεγαλώνει, οι γραμμές του φαίνεται να συγκλίνουν σε κάποια οριακή τιμή. Όπως θα δούμε στην συνέχεια αυτό δεν είναι καθόλου συμπτωματικό αλλά αντιθέτως μια γενική ιδιότητα των πινάκων αυτών.

## 1.3 Πιθανότητες μετάβασης και αρχική κατανομή

Έστω $\{X_n\}$, $n = 0, 1, 2, \ldots$, αλυσίδα Markov με χώρο καταστάσεων $\mathscr{S}$, αρχικές πιθανότητες που δίνονται από το διάνυσμα $\nu = \{\nu_i, i \in \mathscr{S}\}$ και πίνακα πιθανοτήτων μετάβασης $P$. Ισχύει δηλαδή $P(X_0 = i) = \nu_i$, $i \in \mathscr{S}$. Τα στοιχεία αυτά είναι αρκετά για τον προσδιορισμό των πιθανοτήτων οποιουδήποτε γεγονότος σχετικά με την εξέλιξη της αλυσίδας Markov. Για παράδειγμα $P(X_3 = l, X_2 = k, X_1 = j, X_0 = i) = P(X_3 = l | X_2 = k, X_1 = j, X_0 = i) P(X_2 = k, X_1 = j, X_0 = i)$. Ο πρώτος όρος στο δεξιό μέλος αυτής της εξίσωσης μπορεί να γραφτεί σαν $P(X_3 = l | X_2 = k)$ λόγω της μαρκοβιανής ιδιότητας. Ο δεύτερος όρος γράφεται σαν $P(X_2 = k | X_1 = j, X_0 = i) P(X_1 = j, X_0 = i) = P(X_2 = k | X_1 = j) P(X_1 = j | X_0 = i) P(X_0 = i)$ όπου χρησιμοποιήσαμε ξανά την μαρκοβιανή ιδιότητα. Επομένως,

$$P(X_3 = l, X_2 = k, X_1 = j, X_0 = i) = \nu_i P_{ij} P_{jk} P_{kl}.$$

Με τον ίδιο ακριβώς τρόπο μπορούμε να αποδείξουμε τη γενική σχέση

$$\begin{aligned} P(X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) & \\ = \nu_{i_0} P_{i_0 i_1} P_{i_1 i_2} \cdots P_{i_{n-2} i_{n-1}} P_{i_{n-1} i_n}. & \end{aligned} \tag{1.3}$$

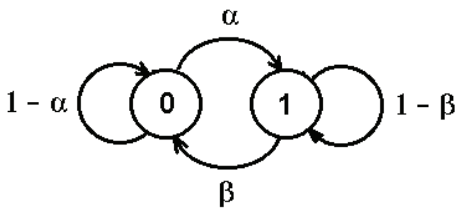
για κάθε $n \in \mathbb{N}$ και $i_0, i_1, \ldots, i_n \in \mathscr{S}$.

Γενικότερα, για κάθε $k_1, k_2, \ldots, k_n \in \mathbb{N}$ με $k_1 < k_2 < \cdots < k_n$ και $i_0, i_1, \ldots, i_n \in \mathscr{S}$ ισχύει ότι

$$\begin{aligned} P(X_{k_n} = i_n, X_{k_{n-1}} = i_{n-1}, \ldots, X_{k_1} = i_1, X_0 = i_0) & \\ = \nu_{i_0} P^{k_1}_{i_0 i_1} P^{k_2 - k_1}_{i_1 i_2} \cdots P^{k_{n-1} - k_{n-2}}_{i_{n-2} i_{n-1}} P^{k_n - k_{n-1}}_{i_{n-1} i_n}. & \end{aligned} \tag{1.4}$$

Για παράδειγμα

$$P(X_6 = l, X_4 = k, X_3 = j, X_0 = i) = \nu_i P^3_{ij} P_{jk} P^2_{kl}.$$




Σχήμα 1.2: Αλυσίδα Markov με δύο καταστάσεις

**Ορισμός 1.1.** Μια αλυσίδα Markov θα ονομάζεται αδιαχώριστη αν από κάθε κατάσταση $i$ του χώρου καταστάσεων μπορούμε να μεταβούμε σε κάθε άλλη, $j$, όχι υποχρεωτικά σε ένα βήμα, με μη μηδενική πιθανότητα.

Η αλυσίδα του σχήματος 1 είναι βεβαίως αδιαχώριστη. Όμως η αλυσίδα με χώρο καταστάσεων $\{0, 1, 2, 3\}$ και πίνακα πιθανοτήτων μετάβασης

$$\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 2/5 & 3/5 \end{bmatrix}$$

δεν είναι αδιαχώριστη. Αν ξεκινήσουμε από την κατάσταση 0 δεν μπορούμε να μεταβούμε στην κατάσταση 3 οσαδήποτε βήματα και αν κάνουμε.

Το αν είναι ή όχι αδιαχώριστη μια αλυσίδα εξαρτάται αποκλειστικά από τον πίνακα πιθανοτήτων μετάβασης $P$. Γι' αυτό το λόγο μπορούμε να μιλάμε ισοδύναμα για πίνακες πιθανοτήτων μετάβασης που είναι αδιαχώριστοι ή όχι.

## 1.4 Αλυσίδες Markov με δύο καταστάσεις

Έστω $X_n$ διαδικασία Markov με δύο καταστάσεις, 0 και 1. Μια δειγματική συνάρτηση αυτής της διαδικασίας θα μπορούσε να είναι λόγου χάριν $01101011100100010\ldots$. Ας υποθέσουμε ότι ο πίνακας πιθανοτήτων μετάβασης είναι

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

Γράφουμε $P = I - S$ όπου $I$ είναι ο ταυτοτικός πίνακας

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{και} \qquad S = \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}.$$

Παρατηρούμε ότι $S^2 = (\alpha+\beta)S$ και επαγωγικά $S^k = (\alpha+\beta)^{k-1}S$, $k = 1, 2, \ldots$. Συνεπώς, από το διωνυμικό θεώρημα έχουμε ότι

$$\begin{aligned} P^n &= (I-S)^n = I + \sum_{k=1}^{n}\binom{n}{k}(-1)^k S^k = I + S\sum_{k=1}^{n}\binom{n}{k}(-1)^k(\alpha+\beta)^{k-1} \\ &= I + \frac{(\alpha+\beta)^n - 1}{\alpha+\beta}S = I - S(\alpha+\beta)^{-1} + S(\alpha+\beta)^{n-1} \end{aligned}$$

Συνεπώς ο πίνακας μετάβασης $n$ βημάτων δίνεται από τον τύπο

$$P^n = \begin{bmatrix} P_{00}^n & P_{01}^n \\ P_{10}^n & P_{11}^n \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix} + (1-\alpha-\beta)^n \begin{bmatrix} \frac{\alpha}{\alpha+\beta} & \frac{-\alpha}{\alpha+\beta} \\ \frac{-\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \end{bmatrix}.$$

Από την παραπάνω σχέση βλέπουμε ότι, όταν $|1-\alpha-\beta| < 1$

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix}.$$

Δεδομένου ότι $0 \le \alpha \le 1$ και $0 \le \beta \le 1$ παρατηρούμε ότι η ποσότητα $|1-\alpha-\beta|$ είναι πάντα μικρότερη της μονάδος εκτός αν

**Περίπτωση 1:** $\alpha = \beta = 1$. Τότε $1-\alpha-\beta = -1$ και $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

**Περίπτωση 2:** $\alpha = \beta = 0$. Τότε $1-\alpha-\beta = 1$ και $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Στην πρώτη περίπτωση παρατηρούμε ότι

$$P^{2n} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{και} \quad P^{2n+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{για } n = 1, 2, \ldots.$$

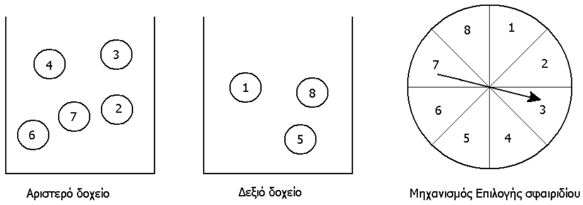
Είναι σαφές ότι το σωματίδιο μεταπηδά από την μία κατάσταση στην άλλη ντετερμινιστικά σε κάθε βήμα. Είναι επίσης προφανές ότι η $P^n$ δεν συγκλίνει όταν $n \to \infty$.

Στην δεύτερη περίπτωση

$$P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Εδώ βλέπουμε ότι το σωματίδιο παραμένει για πάντα στην αρχική του κατάσταση.

Και στις δύο αυτές περιπτώσεις αξίζει να παρατηρήσουμε το γεγονός ότι το σωματίδιο δεν ξεχνάει ποτέ την αρχική του θέση η οποία συνεχίζει να επηρεάζει την συμπεριφορά του ανεξαρτήτως του χρόνου που έχει παρέλθει από την αρχή της διαδικασίας (σε αντίθεση με την γενική περίπτωση $|1-\alpha-\beta| < 1$ όπου βλέπουμε ότι η επίδραση της αρχικής θέσης του σωματιδίου φθίνει γεωμετρικά).




Σχήμα 1.3: Πρότυπο διάχυσης Ehrenfest με 8 σφαιρίδια τα οποία χωρίς βλάβη της γενικότητος θεωρούμε αριθμημένα. Κάθε φορά, με τη βοήθεια του βέλους, διαλέγουμε στην τύχη ένα σφαιρίδιο και το μεταθέτουμε στο αντίθετο δοχείο.

## 1.5 Πρότυπο Διάχυσης των Ehrenfest

Το πρότυπο αυτό είναι μία από τις πρώτες αλυσίδες Markov που μελετήθηκαν (από τους φυσικούς Paul και Tatiana Ehrenfest) και έπαιξε σημαντικό ρόλο στην ανάπτυξη της στατιστικής φυσικής και συγκεκριμένα στη μελέτη των φαινομένων διαχύσεως των αερίων. Έστω δύο δοχεία και $N$ σφαιρίδια, $i$ από τα οποία βρίσκονται αρχικά στο αριστερό δοχείο και $N-i$ στο δεξιό δοχείο.

Τις χρονικές στιγμές $1, 2, \ldots, n, \ldots$ διαλέγουμε κάθε φορά ένα σφαιρίδιο στην τύχη (ανεξάρτητα από τις προηγούμενες επιλογές μας) και το τοποθετούμε στο αντίθετο δοχείο. Ο χώρος καταστάσεων μπορεί να ληφθεί εδώ ως το σύνολο $\mathscr{S} = \{0, 1, 2, \ldots, N-1, N\}$. Όταν η αλυσίδα βρίσκεται στην κατάσταση $i$ αυτό σημαίνει ότι έχουμε $i$ σφαιρίδια στο αριστερό δοχείο (και συνεπώς $N-i$ στο δεξιό). Ο πίνακας πιθανοτήτων μετάβασης δίδεται από τις σχέσεις

$$P_{ij} = \begin{cases} \dfrac{i}{N} & \text{αν } j = i-1 \\[2ex] \dfrac{N-i}{N} & \text{αν } j = i+1 \\[2ex] 0 & \text{διαφορετικά} \end{cases}$$

## 1.6 Στάσιμες κατανομές αλυσίδων Markov

**Ορισμός 1.2.** Αν κάποια κατανομή πιθανότητας $\pi$ στο $\mathscr{S}$ ικανοποιεί τις σχέσεις

$$\pi_j = \sum_{i \in \mathscr{S}} \pi_i P_{ij} \qquad \forall j \in \mathscr{S} \tag{1.5}$$

τότε θα ονομάζεται *στάσιμη κατανομή της αλυσίδας Markov*. Οι εξισώσεις αυτές ονομάζονται εξισώσεις στοχαστικής ισορροπίας ή στάσιμες εξισώσεις.

Αν ξεκινήσουμε την αλυσίδα Markov με την στάσιμη κατανομή, δηλαδή αν $P(X_0 = j) = \pi_j$ για κάθε $j$ τότε

$$P(X_n = j) = \pi_j \qquad \text{για κάθε } n, \tag{1.6}$$

δηλαδή η αλυσίδα Markov συνεχίζει να έχει αυτή την κατανομή για όλες τις επόμενες χρονικές στιγμές. Πράγματι,

$$P(X_1 = j) = \sum_{i \in \mathscr{S}} P(X_1 = j | X_0 = i) P(X_0 = i) = \sum_{i \in \mathscr{S}} \pi_i P_{ij} = \pi_j$$

από την (1.5). Επαναλαμβάνοντας το επιχείρημα αποδεικνύεται η (1.6). Το προηγούμενο επιχείρημα εφαρμόζεται και αντίστροφα, δηλαδή αν η (1.6) ισχύει για κάποια κατανομή $\pi_j$, τότε αυτή θα πρέπει να ικανοποιεί το σύστημα (1.5).

**Σημείωση:** Το γραμμικό σύστημα (1.5) το οποίο μπορούμε να εκφράσουμε συνοπτικά ως $\pi = \pi P$, (όπου $\pi$ είναι το διάνυσμα γραμμής που αντιστοιχεί στην στάσιμη κατανομή) είναι ομογενές και επομένως, αν το διάνυσμα $\pi$ είναι λύση, τότε και το διάνυσμα $c\pi$ είναι λύση για κάθε πραγματικό $c$. Πρέπει επομένως να συμπληρωθεί με τη *εξίσωση κανονικοποίησης*

**Εξίσωση Κανονικοποίησης:** $\sum_{i \in \mathscr{S}} \pi_i = 1, \qquad \pi_i \geq 0$

### 1.6.1 Βασικό Θεώρημα Σύγκλισης για Αλυσίδες Markov

Μια αλυσίδα Markov με πεπερασμένο αριθμό καταστάσεων ονομάζεται απεριοδική αν για κάποιο $m \in \mathbb{N}$ όλα τα στοιχεία του πίνακα $P^m$ είναι αυστηρά θετικά.

Για παράδειγμα, η αλυσίδα του σχήματος 1 είναι απεριοδική. Όμως η διαδικασία Ehrenfest δεν είναι απεριοδική. Αυτό μπορούμε να το καταλάβουμε εύκολα αν συνειδητοποιήσουμε ότι ξεκινόντας από μια κατάσταση με άρτιο αριθμό, (π.χ. την κατάσταση 0 ή την κατάσταση 2) μετά από άρτιο αριθμό βημάτων θα βρισκόμαστε πάλι σε κατάσταση με άρτιο αριθμό. Συνεπώς $P_{00}^{2n+1} = 0$ (αφού ξεκινώντας από το μηδέν είναι αδύνατον να επιστρέψουμε στο μηδέν σε περιττό αριθμό βημάτων) και επομένως η αλυσίδα δεν μπορεί να είναι απεριοδική.

**Θεώρημα 1.1** (Βασικό Θεώρημα Σύγκλισης)**.** *Για μια αδιαχώριστη, απεριοδική αλυσίδα Markov με πεπερασμένο αριθμό καταστάσεων*

$$\lim_{n \to \infty} P_{ij}^n = \pi_j \quad \forall i, j \in \mathscr{S} \tag{1.7}$$

*όπου $\pi_j$ είναι η στάσιμη κατανομή που δίνεται από το σύστημα (1.5). Σάυτή την περίπτωση η* λύση του συστήματος είναι μοναδική.

*Παράδειγμα* 1.1. Έστω $X_n$ αλυσίδα Markov με χώρο καταστάσεων $\{0, 1, 2\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{bmatrix}$$

Να ευρεθεί η στάσιμη κατανομή.

**Λύση:** Το σύστημα $\pi = \pi P$ δίνει τις εξισώσεις

$$\pi_0 = \frac{\pi_0}{3} + \frac{\pi_1}{4} + \frac{\pi_2}{6} \tag{1.8}$$

$$\pi_1 = \frac{\pi_0}{3} + \frac{\pi_1}{2} + \frac{\pi_2}{3} \tag{1.9}$$

$$\pi_2 = \frac{\pi_0}{3} + \frac{\pi_1}{4} + \frac{\pi_2}{2} \tag{1.10}$$

Μία από αυτές τις εξισώσεις είναι περιττή και θα πρέπει να αντικατασταθεί από την εξίσωση κανονικοποίησης

$$\pi_0 + \pi_1 + \pi_2 = 1.$$

Πολλαπλασιάζοντας την (1.8) επί 2 και αφαιρώντας την από την (1.9) απαλείφουμε το $\pi_2$ και παίρνουμε την σχέση $\pi_1 = \frac{5\pi_0}{3}$. Αντικαθιστώντας αυτή τη σχέση στην (1.8) προκύπτει ότι $\pi_2 = \frac{3\pi_0}{2}$. Συνεπώς, από την σχέση κανονικοποίησης έχουμε $\pi_0 + \frac{5\pi_0}{3} + \frac{3\pi_0}{2} = 1$ ή $\pi_0 = \frac{6}{25}$ κι' έτσι $\pi_1 = \frac{2}{5}$ και $\frac{9}{25}$. Κατά την επίλυση του συστήματος δεν χρησιμοποιήσαμε την (1.10).

Αν η αλυσίδα είναι αδιαχώριστη αλλά όχι υποχρεωτικά απεριοδική τότε ισχύει ότι

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P_{ij}^k = \pi_j. \tag{1.11}$$

Παρατηρείστε ότι η (1.7) συνεπάγεται την (1.11) αλλά το αντίστροφο δεν ισχύει.

**Από την παραπάνω σχέση βλέπου με ότι για κάθε αδιαχώριστη αλυσίδα η στάσιμη κατανομή δίνει το ποσοστό του χρόνου που η αλυσίδα Markov βρίσκεται σε κάθε κατάσταση.** Η παραπάνω σχέση μας επιτρέπει να χρησιμοποιούμε την στάσιμη κατανομή για να υπολογίζουμε το ποσοστό του χρόνου που μια αλυσίδα βρίσκεται σε κάθε κατάσταση ασχέτως αν είναι απεριοδική ή όχι. Για το λόγο αυτό το ζήτημα της απεριοδικότητας δεν θα μας απασχολήσει ιδιαίτερα.

**Θεώρημα 1.2.** *Σε μια αδιαχώριστη αλυσίδα η κατανομή ισορροπίας είναι αυστηρά θετική, δηλαδή $\pi_i > 0\ \forall i \in \mathscr{S}$. Αν $\mu_{ii} := E_i T_i$ συμβολίζει τον μέσο χρόνο επιστροφής στην κατάσταση $i$ τότε*

$$\pi_i = \frac{1}{\mu_{ii}} \tag{1.12}$$

## 1.7 Μέσος χρόνος μετάβασης μεταξύ καταστάσεων

Έστω $\mu_{ij}$ ο μέσος χρόνος μετάβασης από μία κατάσταση $i$ σε μια άλλη $j$ (όταν $i = j$ τότε εννοούμε το χρόνο επιστροφής στην κατάσταση $j$). Οι ποσότητες αυτές συνδέονται με τις σχέσεις

$$\mu_{ij} = 1 + \sum_{k\neq j} P_{ik}\mu_{kj} \qquad \text{όταν} \qquad j \neq i.$$

$$\mu_{jj} = 1 + \sum_{k\neq j} P_{jk}\mu_{kj}.$$

Ο μέσος χρόνος επιστροφής στην κατάσταση $j$ μπορεί να υπολογισθεί και κατευθείαν αν είναι γνωστή η στάσιμη κατανομή:

$$\mu_{jj} = \frac{1}{\pi_j}.$$

Θεωρώντας την κατάσταση $j$ σταθερή και την $i$ μεταβλητή έχουμε ένα γραμμικό σύστημα από την επίλυση του οποίου προκύπτουν οι άγνωστοι $\mu_{ij}$.

Για παράδειγμα ας εξετάσουμε την αλυσίδα Ehrenfest με 6 σφαιρίδια. Ο χώρος καταστάσεων σ' αυτή την περίπτωση είναι $\mathscr{S} = \{0, 1, 2, 3, 4, 5, 6\}$ και ο πίνακας πιθανοτήτων μετάβασης είναι

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 0 & 5/6 & 0 & 0 & 0 & 0 \\ 0 & 2/6 & 0 & 4/6 & 0 & 0 & 0 \\ 0 & 0 & 3/6 & 0 & 3/6 & 0 & 0 \\ 0 & 0 & 0 & 4/6 & 0 & 2/6 & 0 \\ 0 & 0 & 0 & 0 & 5/6 & 0 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Η στάσιμη κατανομή είναι

$$\pi_i = \frac{1}{64}\binom{6}{i} \qquad \text{για } \ i = 0, 1, \ldots, 6.$$

Ας υπολογίσουμε τον μέσο χρόνο μετάβασης στην κατάσταση 0 από την κατάσταση 6. Θέτοντας $m_i := \mu_{i0}$ για $i = 1, 2, \ldots, 6$ έχουμε το σύστημα

$$\begin{array}{rcllllll} m_1 & = & 1 & & + \frac{5}{6}m_2 & & & \\ m_2 & = & 1 & + \frac{2}{6}m_1 & & + \frac{4}{6}m_3 & & \\ m_3 & = & 1 & & + \frac{3}{6}m_2 & & + \frac{3}{6}m_4 & \\ m_4 & = & 1 & & & + \frac{4}{6}m_3 & & + \frac{2}{6}m_5 \\ m_5 & = & 1 & & & & + \frac{5}{6}m_4 & \qquad + \frac{1}{6}m_6 \\ m_6 & = & 1 & & & & & + m_5 \end{array} \tag{1.13}$$

Το σύστημα αυτό δίνει

$$m_1 = 63, \qquad m_2 = 74\tfrac{2}{5}, \qquad m_3 = 78\tfrac{3}{5}, \qquad m_4 = 80\tfrac{4}{5}, \qquad m_5 = 82\tfrac{1}{5}, \qquad m_6 = 83\tfrac{1}{5}.$$

Ο χρόνος πρώτης επανόδου στο 0 μπορεί να υπολογισθεί από την σχέση

$$\mu_{00} = 1 + \mu_{10} = 1 + m_1 = 64.$$

Επισημαίνουμε όμως ότι υπάρχει απλούστερος τρόπος που δεν απαιτεί την επίλυση του συστήματος (1.13), υπό την προϋπόθεση ότι έχουμε υπολογίσει την στάσιμη κατανομή. Από το θεώρημα 4.5 έχουμε $\mu_{00} = \frac{1}{\pi_0} = \frac{1}{\frac{1}{64}\binom{6}{0}} = 64$.

## 1.8 Ακολουθιακός έλεγχος ποιότητας

Μια διαδικασία παράγει προϊόντα τα οποία μπορεί να ικανοποιούν τις προδιαγραφές ποιότητας (με πιθανότητα $p$) ή όχι (με πιθανότητα $1-p=q$). Ένας ελεγκτής ελέγχει την ποιότητα των προϊόντων με βάση την ακόλουθη πολιτική. Αρχικά ελέγχει κάθε κομμάτι μέχρις ότου βρει $k$ διαδοχικά κομμάτια καλά. Στο σημείο αυτό ελέγχει δειγματοληπτικά μόνο ένα κομμάτι στα $r$ δηλαδή αφήνει $r-1$ να περάσουν χωρίς έλεγχο και ελέγχει το $r$. Η δειγματοληπτική αυτή διαδικασία σταματάει τη στιγμή που σε έναν από τους δειγματοληπτικούς ελέγχους βρει ένα χαλασμένο κομμάτι. Τότε αρχίζει πάλι να εξετάζει όλα τα κομμάτια μέχρι τη στιγμή που θα βρει $k$ διαδοχικά καλά κομμάτια οπότε η ξεκινά πάλι η δειγματοληπτική διαδικασία. Υποθέτουμε ότι τα ελαττωματικά κομμάτια που βρίσκει ο ελεγκής καταστρέφονται. (Θα μπορούσαμε επίσης να υποθέσουμε ότι επιδιορθώνονται αν αυτό είναι τεχνικά ή οικονομικά εφικτό.) Ο σκοπός αυτής της δειγματοληπτικής διαδικασίας είναι να διατηρείται μια αποδεκτή ποιότητα στην παραγωγή δηλαδή ένα χαμηλό ποσοστό ελαττωματικών κομματιών που διαφεύγουν τον έλεγχο χωρίς το κόστος της διασφάλισης της ποιότητας να γίνεται πολύ μεγάλο.

Στην ανάλυση αυτού του συστήματος θα υποθέσουμε ότι κάθε κομμάτι είναι αποδεκτό ή ελαττωματικό ανεξάρτητα από τα άλλα. Αυτή η υπόθεση απλουστεύει την ανάλυση όμως στην πράξη συνήθως δεν ισχύει. Η παραγωγή ενός ελαττωματικού κομματιού μπορεί να σημαίνει ότι η διαδικασία είναι «εκτός ελέγχου» και κατά συνέπεια συνήθως δεν υπάρχει ανεξαρτησία στην ποιότητα διαδοχικών κομματιών. Αυτός είναι ο λόγος για τον οποίο μόλις βρούμε ένα ελαττωματικό κομμάτι διακόπτουμε τον δειγματοληπτικό έλεγχο μέχρις ότου βεβαιωθούμε (με την εξέταση $k$ διαδοχικών καλών κομματιών) ότι η παραγωγή είναι πάλι εντός ελέγχου.

Προκειμένου να αναλύσουμε την συμπεριφορά αυτής της διαδικασίας (υποθέτοντας πάντα ότι η ποιότητα κάθε κομματιού είναι ανεξάρτητη από την ποιότητα των άλλων) θεωρούμε μια αλυσίδα Markov με χώρο καταστάσεων $\mathscr{S}=\{0,1,\ldots,k,k+1,\ldots,k+r-1\}$. Όταν η αλυσίδα βρίσκεται στην κατάσταση $i=0,1,2,\ldots k$ αυτό σημαίνει ότι ο ελεγκτής βρίσκεται σε φάση πλήρους ελέγχου και μετά από το τελευταίο ελαττωματικό κομμάτι που βρήκε έχει βρεί $i$ διαδοχικά καλά. Όταν βρίσκεται στην κατάσταση $k+j$, $j=1,2,\ldots,r-2$ αυτό σημαίνει ότι ο ελεγκτής είναι σε φάση δειγματοληπτικού ελέγχου και έχουν περάσει $j$ κομμάτια χωρίς να τα ελέγξει. Τέλος όταν είναι στην κατάσταση $k+r-1$ εκτελεί τον δειγματοληπτικό έλεγχο. Για να γίνουν σαφή τα παραπάνω ας δούμε ένα παράδειγμα με $k=3$ και $r=4$. Ο πίνακας πιθανοτήτων μετάβασης θα είναι

$$P=\begin{bmatrix} q & p & & & & & \\ q & & p & & & & \\ q & & & p & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \\ q & & & p & & & \end{bmatrix} \tag{1.14}$$

Παρατηρείστε ότι όταν η αλυσίδα είναι στις καταστάσεις 0,1,2 τότε ο έλεγχος είναι πλήρης. Αν βρεθεί ένα ελαττωματικό κομμάτι (με πιθανότητα $q$) γυρνάμε στο 0. Με καλό κομμάτι (με πιθανότητα $p$) προχωράμε κατά 1 μέχρι να βρεθούμε στην κατάσταση 3. Εκεί αρχίζει ο δειγματοληπτικός έλεγχος και προχωράμε κατά 1 (χωρίς να ελέγχουμε.) Αφού αφήσουμε 3 κομμάτια χωρίς να τα ελέγξουμε, ελέγχουμε το τέταρτο. Αν είναι ελαττωματικό πηγαίνουμε στην κατάσταση 0. Αν είναι καλό πηγαίνουμε στην κατάσταση 3 και αφήνουμε να περάσουν ξανά 3 κομμάτια χωρίς να τα ελέγξουμε.

Η αλυσίσα αυτή είναι αδιαχώριστη και η στάσιμη κατανομή προκύπτει εύκολα ως η λύση του συστήματος

$$\begin{aligned}
\pi_0 &= q(\pi_0 + \pi_1 + \pi_2 + \pi_6) \\
\pi_1 &= p\pi_0 \\
\pi_2 &= p\pi_1 \\
\pi_3 &= p\pi_2 + p\pi_6 \\
\pi_4 &= \pi_3 \\
\pi_5 &= \pi_4 \\
\pi_6 &= \pi_5
\end{aligned}$$

μαζί με την εξίσωση κανονικοποίησης

$$\sum_{i=0}^{6} \pi_i = 1.$$

Επομένως έχουμε, παραλείποντας την πρώτη εξίσωση,

$$\pi_1 = \pi_0 p, \quad \pi_2 = \pi_0 p^2, \quad \pi_3 = \pi_4 = \pi_5 = \pi_6 = \pi_0 q^{-1} p^3.$$

Αντικαθιστώντας στη εξίσωση κανονικοποίησης έχουμε

$$\pi_0(1 + p + p^2 + 4q^{-1}p^3) = 1$$

και χρησιμοποιώντας την γεωμετρική πρόοδο $1 + p + p^2 = q^{-1}(1 - p^3)$ καταλήγουμε στην στάσιμη κατανομή

$$\pi_0 = \frac{q}{1+3p^3}, \qquad \pi_1 = \frac{qp}{1+3p^3}, \qquad \pi_2 = \frac{qp^2}{1+3p^3}, \qquad \pi_3 = \pi_4 = \pi_5 = \pi_6 = \frac{p^3}{1+3p^3}.$$

Από την παραπάνω ανάλυση μπορούμε να υπολογίσουμε τα χαρακτηριστικά αυτής της διαδικασίας. Το ποσοστό των ελαττωματικών κομματιών που ο ελεγκτής βρίσκει και καταστρέφει θα είναι

$$q(\pi_0 + \pi_1 + p_2 + \pi_6).$$

Το ποσοστό των ελαττωματικών κομματιών που θα περάσουν στον καταναλωτή θα είναι

$$\frac{q(\pi_3 + \pi_4 + \pi_5)}{1 - q(\pi_0 + \pi_1 + p_2 + \pi_6)}.$$

(Ο παρονομαστής λαμβάνει υπ΄ όψιν του ότι τα ελαττωματικά κομμάτια που βρίσκονται καταστρέφονται.)

### 1.8.1 Πρόβλημα

Ένας άνθρωπος έχει 4 ομπρέλλες κάποιες από τις οποίες βρίσκονται στο σπίτι ενώ οι άλλες στο γραφείο. Κάθε φορά που βγαίνει από το σπίτι να πάει στο γραφείο ή αντίστροφα βρέχει με πιθανότητα $p$ ή δεν βρέχει με πιθανότητα $q$ *ανεξάρτητητα από τις άλλες φορές*. Όταν βρίσκεται στο σπίτι και βρέχει, παίρνει μια ομπρέλλα και πηγαίνει στο γραφείο. Αν δεν βρέχει στην επιστροφή την αφήνει στο γραφείο. Ποτέ δεν παίρνει ομπρέλλα μαζί του αν δεν βρέχει. Αναπόφευκτα, μερικές φορές θα βρέχει και δεν θα έχει ομπρέλλα γιατί θα βρίσκεται στο σπίτι και όλες οι ομπρέλλες θα είναι στο γραφείο ή αντίστροφα. Τι ποσοστό των μετακινήσεών του βρέχεται;

Θα μπορούσαμε να χρησιμοποιήσουμε μια αλυσίδα Markov με 10 καταστάσεις (αριθμός ομπρελλών που βρίσκεται στο σπίτι, από 0 ως 4 και θέση του ανθρώπου σπίτι ή γραφείο). Είναι όμως πιο εύκολο να χρησιμοποιήσουμε μια αλυσίδα με 5 μόνο καταστάσεις που μας δίνει όλες τις πληροφορίες που χρειαζόμαστε. Έστω $X_n$ ο αριθμός ομπρελλών στο μέρος που βρισκόμαστε (σπίτι ή γραφείο) λίγο πρίν από την $n$-οστή διαδρομή. Αν $X_n = 3$ τότε $X_{n+1} = 1$ με πιθανότητα $q$ (γιατί με πιθανότητα $p$ δεν βρέχει κατά την διαδρομή και δεν παίρνουμε ομπρέλλα επομένως φτάνουμε στον προορισμό μας που έχει 1 ομπρέλλα και λίγο πριν την επόμενη αναχώρηση εκεί θα υπάρχουν 2 ομπρέλλες). Αντίστοιχα $X_{n+1} = 2$ με πιθανότητα $p$. Ο πίνακας πιθανοτήτων μετάβασης είναι συνεπώς

$$P = \begin{bmatrix} & & & & 1 \\ & & & q & p \\ & & q & p & \\ & q & p & & \\ q & p & & & \end{bmatrix}.$$

Ο πίνακας είναι αδιαχώριστος και η στάσιμη κατανομή δίνεται από τις εξισώσεις

$$\begin{aligned} \pi_0 &= q\pi_4 \\ \pi_1 &= q\pi_3 + p\pi_4 \\ \pi_2 &= q\pi_2 + p\pi_3 \\ \pi_3 &= q\pi_1 + p\pi_2 \\ \pi_4 &= \pi_0 + p\pi_1 \end{aligned}$$

και την εξίσωση κανονικοποίησης

$$\sum_{i=0}^{4} \pi_i = 1.$$

Η λύση του παραπάνω συστήματος είναι $\pi_i = 1/5$, $i = 0, 1, 2, 3, 4$. Συνεπώς μακροπρόθεσμα ο αριθμός των ομπρελλών (είτε στο σπίτι είτε στο γραφείο) είναι ομοιόμορφα κατανεμημένος. Το ποσοστό των διαδρομών που βρεχόμαστε είναι $p/5$.

## 1.9 Ασκήσεις

*Άσκηση* 1.1. Έστω αλυσίδα Markov με χώρο καταστάσεων $\mathscr{S} = \{0, 1, 2\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Είναι η αλυσίδα μη διαχωρίσιμη; Να ευρεθεί η στάσιμη κατανομή. Ξεκινώντας από την κατάσταση 0 ποιος είναι ο μέσος χρόνος μέχρι να φθάσουμε στην κατάσταση 2;

*Άσκηση* 1.2. Δίδεται ο πίνακας πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Ποια είναι η στάσιμη κατανομή που αντιστοιχεί στον πίνακα $P$;

*Άσκηση* 1.3 (Έλεγχος αποθεμάτων $s-S$). Η ζήτηση για ένα είδος κάθε μέρα είναι μια τυχαία μεταβλητή με τιμές στους φυσικούς αριθμούς, $f_n$, $n = 0, 1, 2, \ldots$. Αρχικά τα αποθέματα είναι $S$ τεμάχια. Έστω $X_n$ το ύψος των αποθεμάτων στην αρχή της μέρας $n$ και $Y_n$ η ζήτηση κατά τη διάρκεια αυτής της ημέρας. Έτσι

$$X_{n+1} = \begin{cases} X_n - Y_n & \text{όταν } X_n - Y_n > s \\ S & \text{όταν } X_n - Y_n \leq s \end{cases}$$

Όταν στο τέλος της ημέρας τα αποθέματα πέσουν κάτω από το $s$, γίνεται καινούργια παραγελία μεγέθους $S - X_n$ έτσι ώστε στην αρχή της επόμενης ημέρας τα αποθέματα αποκαθίστανται στο αρχικό ύψος $S$. Όταν κάποιο τμήμα της ζήτησης δεν ικανοποιείται άμεσα χάνεται. Έστω ότι η ζήτηση, είναι ισοπίθανα $0, 1, 2, 3, 4$, $s = 2$ και $S = 4$. Να ευρεθεί η κατανομή της στάθμης των αποθεμάτων κάτω από συνθήκες στάσιμης λειτουργίας, καθώς και το ποσοστό της ζήτησης που χάνεται.

*Άσκηση* 1.4. Μια γάτα κυνηγάει ένα ποντίκι σε μια αποθήκη με δύο δωμάτια. Αρχικά η γάτα βρίσκεται στο δωμάτιο 0 και το ποντίκι στο δωμάτιο 1. Το δωμάτιο στο οποίο βρίσκεται η γάτα κάθε χρονική στιγμή περιγράφεται από μια αλυσίδα Markov με πίνακα πιθανοτήτων μετάβασης $P_1$. (Ο χρόνος είναι διακριτός.) Το ποντίκι επίσης κινείται σύμφωνα με μια αλυσίδα Markov με πίνακα πιθανοτήτων μετάβασης $P_2$. Όταν η γάτα και το ποντίκι βρεθούν στο ίδιο δωμάτιο, η γάτα το πιάνει και το κυνήγι τελειώνει. Έστω

$$P_1 = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \qquad \text{και} \qquad P_2 = \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{bmatrix}.$$

Δείξτε ότι η διαδικασία του κυνηγιού μπορεί να περιγραφεί από μια διαδικασία Markov με τρεις καταστάσεις, υπό την προϋπόθεση ότι δεν μας ενδιαφέρει να ξέρουμε σε ποιο δωμάτιο τελειώνει το κυνήγι. Βρείτε τον πίνακα πιθανοτήτων μετάβασης αυτής της διαδικασίας και την μέση διάρκεια του κυνηγιού.

*Άσκηση* 1.5. Ένας παρατηρητής σε κάποιο σημείο της εθνικής οδού παρατηρεί ότι ένα στα πέντε επιβατικά αυτοκίνητα ακολουθείται από φορτηγό ενώ τρία στα πέντε φορτηγά ακολουθούνται από επιβατικά αυτοκίνητα. Ποιό είναι το ποσοστό των φορτηγών που κινούνται στο τμήμα αυτό του δρόμου κατά την διάρκεια των παρατηρήσεων.

*Άσκηση* 1.6. Έστω αλυσίδα Markov με χώρο καταστάσεων $\{0, 1, 2, 3\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}$$

Να δείξετε ότι όλες οι καταστάσεις της αλυσίδας επικοινωνούν μεταξύ τους και να βρείτε την στάσιμη κατανομή.

*Άσκηση* 1.7. Έστω ένα εξάρτημα του οποίου η διάρκεια ζωής (μετρημένη σε μήνες) είναι ομοιόμορφα κατανεμημένη στους ακεραίους $\{1, 2, 3, 4\}$. Στο τέλος κάθε μήνα, το εξάρτημα επιθεωρείται και αν έχει αστοχήσει (υποστεί βλάβη) τότε αντικαθίσταται με ένα παρόμοιο (του οποίου η διάρκεια ζωής έχει την ίδια κατανομή και είναι ανεξάρτητη από εκείνες των προηγουμένων εξαρτημάτων). Κατασκευάστε μία αλυσίδα Markov η οποία να περιγράφει αυτή την διαδικασία.

α) Ποιος είναι ο πίνακας πιθανοτήτων μετάβασης;

β) Ποιες είναι οι στάσιμες πιθανότητες;

γ) Υποθέτουμε ότι το κόστος κάθε αντικατάστασης εξαρτήματος λόγω βλάβης είναι $C_1$ δρχ. Ποιος είναι ο μέσος ρυθμός κόστους (σε δρχ. ανά μήνα) αυτής της διαδικασίας;

Ας υποθέσουμε τώρα ότι θέτουμε σε εφαρμογή την εξής πολιτική προληπτικών αντικαταστάσεων: Το εξάρτημα αντικαθίσταται είτε όταν υφίσταται βλάβη, είτε προληπτικά, *στο τέλος του τρίτου μήνα λειτουργίας*. Ποια είναι η αλυσίδα Markov που περιγράφει αυτή την διαδικασία;

α) Ποιος είναι πίνακας πιθανοτήτων μετάβασης σ΄ αυτή την περίπτωση;

β) Τι ποσοστό των αντικαταστάσεων οφείλεται σε βλάβες και τι ποσοστό σε προληπτικές αντικαταστάσεις;

γ) Υποθέτουμε ότι το κόστος κάθε αντικατάστασης εξαρτήματος λόγω βλάβης είναι $C_1$ δρχ. ενώ το κόστος κάθε προληπτικής αντικατάστασης $C_2 < C_1$. Ποια επιπλέον συνθήκη πρέπει να πληρούν τα $C_1, C_2$, έτσι ώστε η πολιτική προληπτικής αντικατάστασης να είναι οικονομικά συμφέρουσα;

*Άσκηση* 1.8. Παρατηρούμε τις διαδοχικές ρίψεις ενός νομίσματος που έρχεται κορώνα με πιθανότητα $p$ και γράμματα με πιθανότητα $q$.

α) Αν $T_2$ είναι ο χρόνος που απαιτείται μέχρι να εμφανισθούν δύο *διαδοχικές* κορώνες να υπολογίσετε την μέση τιμή $ET_2$.

β) Θεωρούμε την αλυσίδα Markov με χώρο καταστάσεων τον $\{0, 1, 2, \ldots, k\}$ και πίνακα πιθανοτήτων μετάβασης τον $(k+1) \times (k+1)$ πίνακα

$$P = \left[ \begin{array}{ccccccc} q & p & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 & 0 \\ q & 0 & 0 & p & 0 & 0 & 0 \\ & & & \ddots & & & \\ q & 0 & 0 & 0 & 0 & p & 0 \\ q & 0 & 0 & 0 & 0 & 0 & p \\ q & 0 & 0 & 0 & 0 & 0 & p \end{array} \right].$$

Να υπολογίσετε την στάσιμη κατανομή.

γ) Η αλυσίδα του ερωτήματος β) μπορεί να χρησιμοποιηθεί για να υπολογίσουμε τον μέσο χρόνο μέχρι να παρατηρήσουμε $k$ διαδοχικές κορώνες. Να περιγράψετε την διαδικασία και να υπολογίσετε τον $ET_k$, δηλαδή τον μέσο χρόνο που απαιτείται μέχρι να εμφανιστούν $k$ κορώνες.

*Άσκηση* 1.9. Θεωρούμε ένα σύστημα ελέγχου ποιότητας που λειτουργεί ως εξής: Τεμάχια ενός προιόντος έρχονται διαδοχικά σε ένα σταθμό ελέγχου. Κάθε τεμάχιο θεωρούμε ότι είναι καλό με πιθανότητα $p$ και ελαττωματικό με πιθανότητα $q = 1 - p$ *ανεξάρτητα από τα άλλα*. Ο σταθμός ελέγχου έχει δύο διαφορετικές φάσεις λειτουργίας οι οποίες εναλλάσσονται μεταξύ τους. Όταν βρίσκεται στην φάση I εξετάζει *όλα* τα τεμάχια που έρχονται, ένα–ένα, απορρίπτοντας τα ελαττωματικά και αφήνοντας τα καλά να περάσουν. Όταν ο σταθμός ελέγχου βρει $k$ *διαδοχικά* καλά τεμάχια τότε περνάει στην φάση II και αλλάζει τρόπο λειτουργίας. Στην φάση II εξετάζει κάθε τεμάχιο που έρχεται με πιθανότητα $\beta > 0$ ενώ με πιθανότητα $1 - \beta$ το αφήνει να περάσει *χωρίς να το ελέγξει*. Αν, σ' αυτή τη φάση, ο σταθμός ελέγχου τύχει να εξετάσει ένα τεμάχιο και το τεμάχιο τύχει να είναι ελαττωματικό τότε το απορρίπτει και επιστρέφει στην φάση I όπου εξετάζει όλα τα τεμάχια ανεξαιρέτως, μέχρις ότου, ξανά, να βρεί $k$ διαδοχικά καλά τεμάχια.

α) Να βρείτε μια αλυσίδα Markov που περιγράφει αυτό το σύστημα. (Υπόδειξη: Αρκεί μια μικρή τροποποίηση της αλυσίδας του προβλήματος 1.)

β) Να βρείτε την στάσιμη κατανομή της αλυσίδας Markov.

γ) Να υπολογίσετε το ποσοστό των τεμαχίων που το σύστημα ελέγχου απορρίπτει και το ποσοστό των ελαττωματικών τεμαχίων που περνάει χωρίς να απορριφθεί από το σύστημα.

# Κεφάλαιο 2

# Μαρκοβιανές διαδικασίες αποφάσεων

## 2.1 Διαδικασίες Markov με αμοιβές

Έστω μια αδιαχώριστη διαδικασία Markov $(X_n)$ με χώρο καταστάσεων $\mathscr{S}$ και πίνακα πιθανοτήτων μετάβασης $P_{ij}$. Θα υποθέσουμε ότι κάθε φορά που η διαδικασία βρίσκεται στην κατάσταση $j$ παίρνουμε μια αμοιβή $g(j)$ όπου $g : \mathscr{S} \mapsto \mathbb{R}$ μια δεδομένη συνάρτηση. Αν υποθέσουμε ότι την χρονική στιγμή 0 η αλυσίδα βρίσκεται στην κατάσταση $i$ τότε η συνολική αμοιβή που θα πάρουμε τις πρώτες $n-1$ χρονικές στιγμές είναι $\sum_{k=0}^{n-1} g(X_k)$. Η μέση τιμή της συνολικής αμοιβής, δεδομένου ότι ξεκινάμε από την κατάσταση $i$ είναι

$$\begin{aligned} E[\sum_{k=0}^{n-1} g(X_k)|X_0 = i] &= \sum_{k=0}^{n-1} E[g(X_k)|X_0 = i] = \sum_{k=0}^{n-1} \sum_{j\in\mathscr{S}} g(j)P(X_k = j|X_0 = i) \\ &= \sum_{k=0}^{n-1} \sum_{j\in\mathscr{S}} g(j)P_{ij}^k. \end{aligned} \tag{2.1}$$

Μας ενδιαφέρει η μέση αμοιβή ανά μονάδα χρόνου μακροπρόθεσμα, δηλαδή το όριο

$$V := \lim_{n\to\infty} \frac{1}{n} E[\sum_{k=0}^{n-1} g(X_k)|X_0 = i]. \tag{2.2}$$

Χρησιμοποιώντας την (2.1) η παραπάνω σχέση γράφεται και ως

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{j\in\mathscr{S}} g(j)P(X_k = j|X_0 = i) = \sum_{j\in\mathscr{S}} g(j) \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P(X_k = j|X_0 = i) \tag{2.3}$$

Για μια αδιαχώριστη αλυσίδα Markov ισχύει πάντα ότι

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P(X_k = j|X_0 = i) = \pi_j \tag{2.4}$$

όπου $\pi_j$, $j \in \mathscr{S}$ είναι η στάσιμη κατανομή. Καταλήγουμε επομένως στην ακόλουθη έκφραση για την μέση αμοιβή ανά μονάδα χρόνου όταν ο χρονικός ορίζοντας τείνει στο άπειρο:

$$V = \sum_{j \in \mathscr{S}} g(j)\pi_j. \tag{2.5}$$

Παρατηρείστε ότι η παραπάνω έκφραση δεν εξαρτάται από την αρχική κατάσταση εκκίνησης, $i$.

**Παράδειγμα 1.** Οι μηνιαίες πωλήσεις ενός προϊόντος μεταβάλλονται τυχαία ακολουθώντας μια διαδικασία Markov με χώρο καταστάσεων $\mathbb{S} = \{0, 1, 2, 3\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

Η κατάσταση 0 συμβολίζει χαμηλές πωλήσεις, η 1 μέτριες, η 2 ικανοποιητικές και η 3 υψηλές. (Στην πράξη τα στοιχεία του πίνακα πιθανοτήτων μετάβασης θα έπρεπε να εκτιμηθούν από στατιστικά δεδομένα για τις πωλήσεις του προϊόντος.) Η συνάρτηση $g$ δίνει το καθαρό κέρδος στο τέλος ενός μήνα με το αντίστοιχο ύψος πωλήσεων. Ας υποθέσουμε ότι $g(0) = -3$, $g(1) = 1$, $g(2) = 4$ και $g(3) = 6$ (σε δεκάδες χιλιάδες ευρώ. Η αρνητική τιμή για την περίπτωση 0 οφείλεται στο ότι τα έσοδα δεν καλύπτουν τα έξοδα παραγωγής και διάθεσης.) Προκειμένου να βρούμε το μέσο μηνιαίο κέρδος από την διάθεση του προϊόντος βρίσκουμε πρώτα τη στάσιμη κατανομή λύνοντας το σύστημα

$$\pi_0 = \frac{1}{2}\pi_0 + \frac{1}{3}\pi_1 \tag{2.6}$$

$$\pi_1 = \frac{1}{2}\pi_0 + \frac{1}{3}\pi_1 + \frac{1}{3}\pi_2 \tag{2.7}$$

$$\pi_2 = \frac{1}{3}\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{3}\pi_3 \tag{2.8}$$

$$\pi_3 = \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3 \tag{2.9}$$

μαζί με την εξίσωση κανονικοποίησης

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1. \tag{2.10}$$

Γνωρίζουμε ότι μία από τις εξισώσεις (2.6)–(2.9) είναι περιττή και συνεπώς μπορούμε να χρησιμοποιήσουμε τις (2.6), (2.7), (2.9) μαζί με την (2.10) για να βρούμε την στάσιμη κατανομή η οποία προκύπει να είναι

$$\pi_0 = \frac{2}{10}, \quad \pi_1 = \frac{3}{10}, \quad \pi_2 = \frac{3}{10}, \quad \pi_3 = \frac{2}{10}.$$

## 2.2 Μαρκοβιανές Διαδικασίες Αποφάσεων

Θεωρούμε δύο πεπερασμένα σύνολα, το σύνολο καταστάσεων, $\mathscr{S}$ το οποίο συνήθως θα ταυτίζουμε με κάποιο υποσύνολο των φυσικών $\{0, 1, 2, \ldots, N\}$ και το σύνολο αποφάσεων $A = \{a, b, c, \ldots q\}$. Θεωρούμε επίσης μια *οικογένεια* από πίνακες πιθανοτήτων μετάβασης, $\{P_{ij}(a), a \in A\}$ καθώς και μια συνάρτηση κόστους $C : \mathscr{S} \times A \mapsto \mathbb{R}$. Συγκεκριμένα, $C(i, a)$ είναι το κόστος που πληρώνουμε όταν βρισκόμαστε στην κατάσταση $i$ και επιλέξουμε την απόφαση $a$. Αν $X_k = i$ είναι η κατάσταση την χρονική στιγμή $k$ και $Y_n = a$ η απόφαση που παίρνουμε την ίδια χρονική στιγμή, τότε η επόμενη κατάσταση είναι $X_{k+1} = j$ με πιθανότητα $P(X_{k+1} = j | X_k = i, Y_k = a) = P_{ij}(a)$. Το συνολικό κόστος μετά αυτής της διαδικασίας μετά από $n$ βήματα, θα είναι

$$\sum_{k=0}^{n-1} C(X_k, Y_k).$$

Σκοπός μας στο προκείμενο πρόβλημα είναι να βρούμε μια ακολουθία αποφάσεων $Y_k$, $k = 0, 1, 2, \ldots$ τέτοια ώστε να ελαχιστοποιεί το μέσο ρυθμό κόστους σε μακροπρόθεσμο χρονικό ορίζοντα (ή να μεγιστοποιεί την μέση συνολική αμοιβή, στην περίπτωση που τα $C(i, a)$ αντιπροσωπεύουν αμοιβές). Δεν είμαστε όμως έτοιμοι να γράψουμε το πρόβλημα υπό τη μορφή προβλήματος βελτιστοποίησης κάποιας μέσης τιμής πριν ξεκαθαρίσουμε ένα σημαντικό ζήτημα. Η εξέλιξη της διαδικασίας $\{X_n\}$ δεν έχει προσδιορισμένα στατιστικά χαρακτηριστικά εφ' όσον δεν προσδιορίσουμε την ακολουθία αποφάσεων. Αν θεωρήσουμε ότι ζητάμε να βρούμε την ακολουθία $\{Y_k\}$ που βελτιστοποιεί κάποιο κριτήριο τότε το πρόβλημα γίνεται πολύ δύσκολο μια και στον μακροχρόνιο ορίζοντα βελτιστοποιούμε ως προς ένα άπειρο (ή πολύ μεγάλο) πλήθος παραμέτρων. Ευτυχώς υπάρχει μια έννοια που απλοποιεί σημαντικά το πρόβλημα και αυτή είναι η έννοια της πολιτικής (policy).

Θα ονομάζουμε *πολιτική* κάθε συνάρτηση $f : \mathscr{S} \mapsto A$ από τον χώρο καταστάσεων στον χώρο αποφάσεων. Μια πολιτική με άλλα λόγια είναι ένας κανόνας που υπαγορεύει μια απόφαση, έστω $a$, κάθε φορά που η αλυσίδα Markov βρίσκεται στην κατάσταση $i$. Έτσι έχουμε $f(i) = a$. Η υιοθέτηση μιας πολιτικής περιορίζει την ελευθερία που έχουμε να διαλέγουμε τις αποφάσεις μας. Μας αναγκάζει κάθε φορά που βρισκόμαστε σε μια συγκεκριμένη κατάσταση να διαλέγουμε την ίδια πάντα απόφαση. Αποδεικνύεται όμως ότι, λόγω της μαρκοβιανής φύσης της διαδικασίας, υπάρχει *βελτιστη πολιτική* η οποία εξασφαλίζει συνολικό μέσο κόστος εξίσου χαμηλό ή χαμηλότερο από οποιαδήποτε άλλη ακολουθία αποφάσεων που δεν προκύπτει από κάποια πολιτική. Συνεπώς είναι αρκετό να αναζητήσουμε την βέλτιστη πολιτική χωρίς να χρειάζεται να ασχοληθούμε όλες τις δυνατές ακολουθίες αποφάσεων.

Ένα σημαντικότατο κέρδος που έχουμε περιοριζόμενοι σε αποφάσεις που προκύπτουν από πολιτικές είναι η ακόλουθη: Αν χρησιμοποιήσουμε μια οποιαδήποτε πολιτική, $f$, η διαδικασία που προκύπτει ειναι μαρκοβιανή με πίνακα πιθανοτήτων μετάβασης $P_{ij}(f(i))$. Θα συμβολίζουμε τον πίνακα αυτό ως $P^f$ και το $i, j$ στοιχείο του ως $P^f_{ij} := P_{ij}(f(i))$.

[0]Παραπέμπουμε τους ενδιαφερόμενους στον Sheldon Ross, (1970). *Applied Probability Models with Optimization Applications,* Εκδόσεις Dover.

Ας ορίσουμε τώρα τον ρυθμό κέρδους (ή αντίστοιχα κόστους) σε μακροχρόνιο ορίζοντα ως συνάρτηση της πολιτικής $f$. Αφού επιλέγοντας μια πολιτική $f$ προσδιορίζουμε μια αλυσίδα Markov θα ισχύει ότι

$$\begin{aligned}
V_f(i) &= \lim_{n\to\infty} \frac{1}{n} E[\sum_{k=0}^{n-1} C(X_k, f(X_k))|X_0 = i] \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{j\in\mathscr{S}} C(j, f(j))(P^f)^k_{ij} \\
&= \sum_{j\in\mathscr{S}} C(j, f(j)) \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (P^f)^k_{ij} \\
&= \sum_{j\in\mathscr{S}} C(j, f(j))\pi^f_j
\end{aligned}$$

Το νόημα της παραπάνω σχέσης είναι το εξής: Από τη στιγμή που υιοθετούμε μια πολιτική $f$ η διαδικασία που προκύπτει είναι μαρκοβιανή με πίνακα πιθανοτήτων μετάβασης $P^f$. Συνεπώς η δεσμευμένη πιθανότητα $P(X_k = j|X_0 = i)$ δίνεται από το στοιχείο $i, j$ του πίνακα $P^f$ υψωμένου στην δύναμη $k$, $(P^f)^k_{ij}$. Αν ο πίνακας αυτός είναι αδιαχώριστος, πράγμα που θα υποθέσουμε, τότε έχει στάσιμη κατανομή $\pi^f$ (που ασφαλώς εξαρτάται από την πολιτική $f$) που προκύπτει από την μοναδική λύση του συστήματος

$$\pi^f_j = \sum_{i\in\mathscr{S}} \pi^f_i (P^f)_{ij} \qquad j \in \mathscr{S}, \tag{2.11}$$

μαζί με την εξίσωση κανονικοποίησης

$$\sum_{i\in\mathscr{S}} \pi^f_i = 1. \tag{2.12}$$

Παρατηρείστε τέλος ότι ο ρυθμός κέρδους που δίνεται από την (2.11) δεν εξαρτάται από την αρχική κατάσταση $i$.

## 2.3 Εύρεση της βέλτιστης πολιτικής με τη βοήθεια του γραμμικού προγραμματισμού

Το πρόβλημα της βέλτιστης πολιτικής μπορεί βεβαίως να λυθεί εξετάζοντας όλες τις πολιτικές, υπολογίζοντας το μέσο ρυθμό κόστους για την κάθε μία από αυτές και επιλέγοντας την καλύτερη. Αυτό όμως μπορεί να είναι δύσκολο. Ας αριθμήσουμε τις καταστάσεις και τις αποφάσεις, δηλαδή ας θέσουμε $\mathscr{S} = \{1, 2, \ldots, N\}$ και $A = \{1, 2, \ldots, M\}$. Τότε υπάρχουν $M^N$ διαφορετικές πολιτικές από τις οποίες θα πρέπει να επιλέξουμε την καλύτερη και για μεγάλα $M$ και $N$ το υπολογιστικό κόστος μπορεί να είναι απαγορευτικό. Υπάρχει όμως μια επαναδιατύπωση του προβλήματος η οποία μπορεί να λυθεί εύκολα με τη βοήθεια του γραμμικού προγραμματισμού. Η επαναδιατύπωση αυτή χρησιμοποιεί την έννοια των τυχαιοποιημένων πολιτικών, αποδεικνύεται όμως ότι τελικά η βέλτιστη λύση που προκύπτει είναι πάντα μη τυχαιοποιημένη και συνεπώς η χρήση των τυχαιοποιημένων πολιτικών δεν επηρρεάζει την τελική λύση του προβλήματος.

Το αρχικό πρόβλημα έχει ως εξής: Δεδομένων δύο συνόλων $\mathscr{S} = \{1, 2, \ldots, N\}$ και $A = \{1, 2, \ldots, M\}$, μιας οικογένειας πινάκων πιθανοτήτων μετάβασης $P_{ij}(k)$ (όπου βεβαίως $i, j \in \mathscr{S}$ και $k \in A$) και μιας συνάρτησης κόστους $C : \mathscr{S} \times A \to \mathbb{R}$ να βρεθεί μια συνάρτηση $f : \mathscr{S} \to A$ τέτοια ώστε να ελαχιστοποιεί το κριτήριο

$$V_f = \sum_{j=1}^{N} C(j, f(j))\pi_j^f$$

υπό τους περιορισμούς

$$\begin{aligned}
&P_{ij}^f = P(f(i))_{ij} \\
&\pi_j^f = \sum_{i=1}^{N} \pi_i^f (P^f)_{ij} \qquad j = 1, 2, \ldots, N, \\
&\sum_{i=1}^{N} \pi_i^f = 1 \\
&\pi_i^f \geq 0, \qquad i = 1, 2, \ldots, N.
\end{aligned}$$

Η επαναδιατύπωση του προβλήματος που θα δώσουμε θα το μετατρέψει σε ένα πρόβλημα γραμμικού προγραμματισμού. Ας εξετάσουμε τυχαιοποιημένες στρατηγικές: Όταν βρισκόμαστε στην κατάσταση $i$ δεν επιλέγουμε πάντα την ίδια δράση. Αντίθετα, με πιθανότητα $d_{ik}$ επιλέγουμε την δράση $k$ ανεξάρτητα από την προηγούμενη ιστορία της διαδικασίας. Ισχύχει βεβαίως ότι

$$\sum_{k=1}^{M} d_{ik} = 1 \qquad i = 1, 2, \ldots, M. \tag{2.13}$$

Επομένως, το σύνολο των τυχαιοποιημένων πολιτικών με την παραπάνω έννοια είναι το σύνολο όλων των πινάκων $D$ $N \times M$ με μη αρνητικά στοιχεία των οποίων κάθε γραμμή αθροίζεται στη μονάδα. Το σύνολο αυτό περιέχει ασφαλώς και τις μη τυχαιοποιημένες πολιτικές που αντιστοιχούν σε πίνακες $D$ που έχουν ακριβώς μια μονάδα σε κάθε γραμμή και όλα τα υπόλοιπα στοιχεία 0. Έστω τώρα ότι έχουμε υιοθετήσει μια τυχαιοποιημένη πολιτική με πίνακα $D$ και ότι βρισκόμαστε στην κατάσταση $i$. Η πιθανότητα να μεταπηδήσουμε στο επόμενο βήμα στην κατάσταση $j$ είναι

$$P_{ij}^d := \sum_{k=1}^{M} d_{ik} P_{ij}(k). \tag{2.14}$$

Η παραπάνω σχέση ορίζει ένα πίνακα πιθανοτήτων μετάβασης. Πράγματι, $P_{ij}^d \geq 0$ και

$$\sum_{j=1}^{N} P_{ij}^d = \sum_{j=1}^{N} \sum_{k=1}^{M} d_{ik} P_{ij}(k) = \sum_{k=1}^{M} d_{ik} \sum_{j=1}^{N} P_{ij}(k) = \sum_{k=1}^{M} d_{ik} = 1 \qquad \text{για κάθε } i = 1, \ldots, N.$$

Η στάσιμη κατανομή που αντιστοιχεί στον πίνακα πιθανοτήτων μετάβασης $P^d$ δίνεται από τις σχέσεις

$$\pi_j^d = \sum_{i=1}^{N} \pi_i^d P_{ij}^d \qquad j = 1, \ldots, N$$

και

$$\sum_{j=1}^{N} \pi_j^d = 1.$$

Όταν βρισκόμαστε στην κατάσταση $i$ το μέσο κόστος που έχουμε για ένα βήμα είναι $\sum_{k=1}^{M} d_{ik} C(i,k)$ αφού επιλέγουμε την δράση $k$ με πιθανότητα $d_{ik}$ ακολουθώντας την τυχαιοποιημένη πολιτική $D$. Συνεπώς ο ρυθμός κόστους που προκύπτει από την υιοθέτηση της τυχαιοποιημένης πολιτικής $D$ είναι

$$V = \sum_{i=1}^{N} \pi_i^d \sum_{k=1}^{M} d_{ik} C(i,k) \tag{2.15}$$

Το πρόβλημα βελτιστοποίησης επομένως είναι η επιλογή των $d_{ik}$ έτσι ώστε να ελαχιστοποιήσουμε ή να μεγιστοποιήσουμε την (2.15) ανάλογα με το αν πρόκειται για κόστος ή κέρδος κάτω από τους περιορισμούς

$$d_{ik} \geq 0, \quad i = 1, \ldots, N, \;\; k = 1, \ldots, M,$$

$$\sum_{k=1}^{M} d_{ik} = 1, \quad i = 1, \ldots, N, \tag{2.16}$$

$$\pi_j^d = \sum_{i=1}^{N} \pi_i^d \sum_{k=1}^{M} d_{ik} P_{ij}(k) \qquad j = 1, \ldots, N \tag{2.17}$$

$$\sum_{j=1}^{N} \pi_j^d = 1, \tag{2.18}$$

$$\pi_j^d \geq 0, \qquad j = 1, \ldots, N. \tag{2.19}$$

Το πρόβλημα αυτό δεν είναι στη μορφή ενός γραμμικού προγράμματος δεδομένου ότι, τόσο στο κριτήριο όσο και σε έναν από τους περιορισμούς εμφανίζονται γινόμενα μεταβλητών της μορφής $\pi_i^d d_{ik}$. Μπορούμε όμως να το μετατρέψουμε εύκολα σε ένα πρόβλημα γραμμικού προγραμματισμού ορίζοντας νέες μεταβλητές, τις

$$y_{ik} := \pi_i^d d_{ik}. \tag{2.20}$$

Δεδομένου ότι η στάσιμη κατανομή $\pi_i^d$ εκφράζει, μεταξύ άλλων, και το ποσοστό του χρόνου που η διαδικασία βρίσκεται στην κατάσταση $i$ (όταν χρησιμοποιείται η τυχαιοποιημένη πολιτική $D$) οι μεταβλητές $y_{ik}$ εκφράζουν το ποσοστό του χρόνου που είμαστε στην κατάσταση $i$ και επιλέγουμε την δράση $k$. (Οι μεταβλητές $d_{ik}$ μπορούν να θεωρηθούν δεσμευμένες πιθανότητες να επιλέξουμε την δράση $k$ δεδομένου ότι είμαστε στην κατάσταση $i$.) Αφού με τον παραπάνω ορισμό ισχύει ότι

$$\pi_i^d = \sum_{k=1}^{M} y_{ik} \tag{2.21}$$

το πρόβλημα βελτιστοποίησης μπορεί να διατυπωθεί ως

$$\max V = \sum_{i=1}^{N}\sum_{k=1}^{M} y_{ik}C(i,k) \tag{2.22}$$

υπό τους περιορισμούς

$$\sum_{k=1}^{M} y_{jk} = \sum_{i=1}^{N}\sum_{k=1}^{M} y_{ik}P_{ij}(k) \qquad j=1,\ldots,N, \tag{2.23}$$

$$\sum_{j=1}^{N}\sum_{k=1}^{M} y_{jk} = 1, \tag{2.24}$$

$$y_{ik} \geq 0, \quad i=1,\ldots,N, \;\; k=1,\ldots,M. \tag{2.25}$$

Το πρόβλημα τώρα έχει διατυπωθεί υπό τη μορφή γραμμικού προγράμματος και επομένως μπορεί να επιλυθεί με τη μέθοδο Simplex.

Παρατηρείστε ότι με την εισαγωγή των μεταβλητών $y_{ik}$ εξαφανίστηκε ο περιορισμός

$$\sum_{k=1}^{M} d_{jk} = 1, \qquad j=1,\ldots,N. \tag{2.26}$$

Αυτό οφείλεται στο γεγονός ότι ο περιορισμός αυτός είναι ισοδύναμος με την (2.21) όπως μπορούμε να διαπιστώσουμε αν πολλαπλασιάσουμε και τα δύο σκέλη της (2.26) με $\pi_i^d$. Συνεπώς αυτός ο περιορισμός έχει απορροφηθεί από τους (2.23) και (2.24).

## 2.4 Ένα πρόβλημα διαφήμισης

Οι πωλήσεις ενός προϊόντος μπορεί να βρίσκονται σε μία από δύο καταστάσεις, 0=χαμηλές ή 1=υψηλές. (Η χονδρική αυτή διακριτοποίηση αποσκοπεί απλά στην παρουσίαση των ιδεών χωρίς τις λεπτομέριες που θα τις συσκότιζαν αν το παράδειγμα ήταν πιο ρεαλιστικό). Χωρίς διαφήμηση οι πωλήσεις από μήνα σε μήνα ακολουθούν μια διαδικασία Markov με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{bmatrix} 4/5 & 1/5 \\ 1/2 & 1/2 \end{bmatrix}.$$

Με διαφήμιση οι πωλήσεις από μήνα σε μήνα ακολουθούν μια άλλη διαδικασία Markov με πίνακα πιθανοτήτων μετάβασης

$$P(1) = \begin{bmatrix} 2/3 & 1/3 \\ 1/4 & 3/4 \end{bmatrix}.$$

Ο χώρος αποφάσεων κωδικοποιείται ως $\{0,1\}$. Η συνάρτηση κέρδους δίδεται ως $f_{00} = 1$ (χαμηλές πωλήσεις, όχι διαφήμιση), $f_{01} = -1$ (χαμηλές πωλήσεις, διαφήμιση), $f_{10} = 4$ (υψηλές πωλήσεις, όχι διαφήμιση), $f_{11} = 2$ (υψηλές πωλήσεις, διαφήμηση). Το πρόβλημα βελτιστοποίσης γίνεται τότε

$$\begin{aligned}
&\max \; f_{00}y_{00} + f_{01}y_{01} + f_{10}y_{10} + f_{11}y_{11}\\
&\text{ υπό τους περιορισμούς}\\
&y_{00} + y_{01} = y_{00}P(0)_{00} + y_{01}P(1)_{00} + y_{10}P(0)_{10} + y_{11}P(1)_{10}\\
&y_{10} + y_{11} = y_{00}P(0)_{01} + y_{01}P(1)_{01} + y_{10}P(0)_{11} + y_{11}P(1)_{11}\\
&y_{00} + y_{01} + y_{10} + y_{11} = 1\\
&y_{00},\; y_{01},\; y_{10},\; y_{11} \geq 0
\end{aligned}$$

Αντικαθιστώντας τις τιμές έχουμε

$$\begin{aligned}
&\max \; y_{00} - y_{01} + 4y_{10} + 2y_{11}\\
&\text{ υπό τους περιορισμούς}\\
&y_{00} + y_{01} = \frac{4}{5}y_{00} + \frac{2}{3}y_{01} + \frac{1}{2}y_{10} + \frac{1}{4}y_{11}\\
&y_{10} + y_{11} = \frac{1}{5}y_{00} + \frac{1}{3}y_{01} + \frac{1}{2}y_{10} + \frac{3}{4}y_{11}\\
&y_{00} + y_{01} + y_{10} + y_{11} = 1\\
&y_{00},\; y_{01},\; y_{10},\; y_{11} \geq 0.
\end{aligned}$$

Λύνοντας το γραμμικό πρόγραμμα με την μέθοδο Simplex παίρνουμε την λύση $y_{00} = 0,71$, $y_{01} = 0$, $y_{10} = 0,29$, $y_{11} = 0$. Η βέλτιστη τιμή του κριτηρίου είναι 1,86. Προκειμένου να βρούμε την βέλτιστη πολιτική παρατηρούμε ότι

$$d_{00} = \frac{y_{00}}{y_{00} + y_{01}} = 1, \quad d_{01} = \frac{y_{01}}{y_{00} + y_{01}} = 0, \quad d_{10} = \frac{y_{10}}{y_{10} + y_{11}} = 1, \quad d_{01} = \frac{y_{11}}{y_{10} + y_{11}} = 0.$$

Συνεπώς η βέλτιστη στρατηγική, η οποία προκύπτει (όπως πρέπει) μη τυχαιοποιημένη είναι να διαφημίζουμε όταν είμαστε σε κατάσταση χαμηλών πωλήσεων και να μη διαφημίζουμε όταν είμαστε σε κατάσταση υψηλών πωλήσεων. Το ποσοστό του χρόνου στις χαμηλές πωλήσεις είναι $y_{00} + y_{01} = 0,71$ ενώ σε κατάσταση υψηλών πωλήσεων είναι $y_{10} + y_{11} = 0,29$. Το κέρδος ανά μήνα είναι 1,86 μονάδες.