

Game Theory

Γιάννης Γεωργίου - Μουσσές

April 2025

1 Υπολογισμός μέσου αναμενόμενου payoff με χρήση μαρκοβιανών αλυσίδων

Σε ένα γύρο του παιχνιδιού, τα αποτελέσματα μπορεί να είναι CC ή CD ή DC ή DD , όπου C = cooperation και D = defection και σε κάθε ζευγάρι το πρώτο γράμμα δηλώνει την επιλογή του πρώτου παίκτη και το δεύτερο την επιλογή του δεύτερου παίκτη. Σε παιχνίδια όπου οι δύο παίκτες υιοθετούν στρατηγικές που βασίζονται στη μνήμη του τι παίχτηκε στον προηγούμενο γύρο (memory-one strategies) ή και γενικότερα σε στρατηγικές που βασίζονται σε πεπερασμένη μνήμη, αποδεικνύεται ότι το παιχνίδι μπορεί να περιγραφεί από μια αλυσίδα Markov πεπερασμένων καταστάσεων, όπου το σύστημα μπορεί να μεταβαίνει μεταξύ των τεσσάρων καταστάσεων CC , CD , DC , DD και όπου οι δεσμευμένες πιθανότητες οποιουδήποτε μελλοντικού γεγονότος, δεδομένου οποιουδήποτε παρελθοντικού γεγονότος εξαρτώνται μόνο από την παρούσα κατάσταση και είναι ανεξάρτητες του παρελθοντικού γεγονότος. Σε αυτή τη βάση, είναι σημαντικός ο πίνακας μεταβάσεων καταστάσεων M σε ένα βήμα (one step transition matrix) με την παρακάτω μορφή

$$M = \begin{matrix} & \begin{matrix} CC & CD & DC & DD \end{matrix} \\ \begin{matrix} CC \\ CD \\ DC \\ DD \end{matrix} & \begin{bmatrix} p(CC|CC) & p(CC|CD) & p(CC|DC) & p(CC|DD) \\ p(CD|CC) & p(CD|CD) & p(CD|DC) & p(CD|DD) \\ p(DC|CC) & p(DC|CD) & p(DC|DC) & p(DC|DD) \\ p(DD|CC) & p(DD|CD) & p(DD|DC) & p(DD|DD) \end{bmatrix} \end{matrix}$$

όπου για παράδειγμα, $p(CD|DC) = p(X_{k+1} = CD | X_k = DC)$ δηλώνει την πιθανότητα η επόμενη κατάσταση του παιχνιδιού να είναι η CD , δεδομένου ότι η τρέχουσα κατάσταση είναι η DC (ο πρώτος παίκτης επιλέγει C και ο δεύτερος D ως επόμενη κίνηση, ενώ οι προηγούμενες κινήσεις τους ήταν D και C , αντίστοιχα).

Επειδή όμως για παράδειγμα $p(CD|DC) = p_X(C|DC)p_Y(D|DC) = p_X(C|DC)(1 - p_Y(C|DC))$ λόγω ανεξαρτησίας των κινήσεων των δύο παικτών, όπου $p_X(C|DC)$ η πιθανότητα ο πρώτος παίκτης να επιλέξει C για την επόμενη κίνηση, όταν η προηγούμενη κατάσταση του συστήματος ήταν DC και $p_Y(D|DC)$ η πιθανότητα ο δεύτερος παίκτης να επιλέξει D για την επόμενη κίνηση όταν η προηγούμενη κατάσταση του συστήματος ήταν DC , είναι σημαντικό για κάθε στρατηγική να μπορούμε να υπολογίσουμε τις δεσμευμένες πιθανότητες $p_X(C|CC)$, $p_X(C|CD)$, $p_X(C|DC)$ και $p_X(C|DD)$, οι οποίες εξαρτώνται από την εκάστοτε στρατηγική, δηλαδή τις πιθανότητες που έχει ο παίκτης, ο οποίος ακολουθεί τη στρατηγική, να επιλέξει C (cooperation) για την επόμενη κίνησή του, δεδομένου ότι η παρούσα κατάσταση του συστήματος είναι η CC , ή η CD , ή η DC , ή η DD . Έστω ότι συμβολίζουμε με

$$\mathbf{p} = (p_1, p_2, p_3, p_4) = (p_X(C|CC) \ p_X(C|CD) \ p_X(C|DC) \ p_X(C|DD))$$

και

$$\mathbf{q} = (q_1, q_2, q_3, q_4) = (p_Y(C|CC) \ p_Y(C|CD) \ p_Y(C|DC) \ p_Y(C|DD))$$

τα διανύσματα των στρατηγικών του πρώτου παίκτη και του δεύτερου παίκτη, αντίστοιχα. Το διάνυσμα \mathbf{p} μπορεί να υπολογιστεί για κάθε διαφορετική στρατηγική. Για παράδειγμα, αν κάποιος παίκτης ακολουθεί τη στρατηγική random, το διάνυσμα $\mathbf{p} = (0.5, 0.5, 0.5, 0.5)$ γιατί η πιθανότητα επιλογής του C ως επόμενης κίνησης είναι 50%, ανεξάρτητα από την τρέχουσα κατάσταση CC , CD , DC , DD του συστήματος. Για την All- C το διάνυσμα είναι το $(1, 1, 1, 1)$ αφού η στρατηγική επιλέγει πάντα C ανεξάρτητα της τρέχουσας κατάστασης. Για την All- D το διάνυσμα είναι το $(0, 0, 0, 0)$ αφού η στρατηγική επιλέγει πάντα D ανεξάρτητα της τρέχουσας κατάστασης. Για την Tit-For-Tat, έχουμε $\mathbf{p} = (1, 0, 1, 0)$, αφού όταν η προηγούμενη κατάσταση είναι CC ή DC , ο αντίπαλος έχει επιλέξει C και η στρατηγική Tit-For-Tat θα

επιλέξει C , ενώ όταν η προηγούμενη κατάσταση είναι CD ή DD , ο αντίπαλος έχει επιλέξει D και η στρατηγική Tit-For-Tat θα απαντήσει επίσης με D . Για την Generous Tit-For-Tat (GTFT), όπου όταν ο αντίπαλος έχει επιλέξει D , ο GTFT επιλέγει C με πιθανότητα q και D με πιθανότητα $1 - q$, το διάνυσμα γίνεται $(1, q, 1, q)$. Για την Pavlov (Win Stay Lose Shift) έχουμε το διάνυσμα $(1, 0, 0, 1)$ αφού σε περίπτωση διαφωνίας στην προηγούμενη κίνηση (CD , ή DC) η Pavlov επιλέγει D ενώ για συμφωνία (CC , ή DD) η Pavlov επιλέγει C . Με χρήση των διανυσμάτων $\mathbf{p} = (p_1, p_2, p_3, p_4)$ της στρατηγικής του πρώτου παίκτη X και $\mathbf{q} = (q_1, q_2, q_3, q_4)$ της στρατηγικής του δεύτερου παίκτη Y , ο πίνακας μετάβασης καταστάσεων σε ένα βήμα M , υπολογίζεται όταν παίζουν μεταξύ τους οι στρατηγικές \mathbf{p} και \mathbf{q} , ως

$$M = \begin{bmatrix} p_1 q_1 & p_1(1 - q_1) & (1 - p_1)q_1 & (1 - p_1)(1 - q_1) \\ p_2 q_3 & p_2(1 - q_3) & (1 - p_2)q_3 & (1 - p_2)(1 - q_3) \\ p_3 q_2 & p_3(1 - q_2) & (1 - p_3)q_2 & (1 - p_3)(1 - q_2) \\ p_4 q_4 & p_4(1 - q_4) & (1 - p_4)q_4 & (1 - p_4)(1 - q_4) \end{bmatrix}$$

Το άθροισμα στοιχείων σε κάθε γραμμή του πίνακα M πρέπει να ισούται με 1. Σημειώνουμε ότι στη δεύτερη γραμμή του M έχουμε μετάβαση στην κατάσταση CD για το σύστημα (δεδομένης της τρέχουσας κατάστασης CC , CD , DC , DD) και επομένως για το δεύτερο παίκτη έχουμε μετάβαση στην κατάσταση D (δεδομένης της τρέχουσας κατάστασης CC , CD , DC , DD) και για το λόγο αυτό χρησιμοποιείται το $q_3 = p_Y(C|DC)$ που δηλώνει την πιθανότητα ο δεύτερος παίκτης Y να επιλέξει C , όταν η τρέχουσα είναι η DC , δηλαδή όταν ο δεύτερος παίκτης είναι στην τρέχουσα κατάσταση D και ο πρώτος στην C , και δε χρησιμοποιείται το q_2 . Ανάλογη σημείωση ισχύει και για την τρίτη γραμμή του πίνακα M .

Σημαντικό είναι να μπορούμε να υπολογίσουμε τις μη δεσμευμένες πιθανότητες το σύστημα να βρεθεί σε μια από τις καταστάσεις CC , CD , DC ή DD και το πως αυτές εξελίσσονται από γενιά σε γενιά του παιχνιδιού. Έστω ότι ορίζουμε το διάνυσμα

$$v_k = [p(X_k = CC) \quad p(X_k = CD) \quad p(X_k = DC) \quad p(X_k = DD)]$$

που περιλαμβάνει τις πιθανότητες (μη δεσμευμένες) το σύστημα να βρεθεί σε κάποια από τις καταστάσεις CC , CD , DC , ή DD , στο βήμα k . Τότε, χρησιμοποιώντας το θεώρημα της ολικής πιθανότητας, έχουμε για παράδειγμα

$$\begin{aligned} p(X_{k+1} = CD) &= p(X_{k+1} = CD|X_k = CC)p(X_k = CC) + p(X_{k+1} = CD|X_k = CD)p(X_k = CD) \\ &+ p(X_{k+1} = CD|X_k = DC)p(X_k = DC) + p(X_{k+1} = CD|X_k = DD)p(X_k = DD) \\ &= [p(CD|CC) \quad p(CD|CD) \quad p(CD|DC) \quad p(CD|DD)] \cdot [p(X_k = CC) \quad p(X_k = CD) \quad p(X_k = DC) \quad p(X_k = DD)]^T \end{aligned}$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε τον πιο απλό συμβολισμό που είχαμε υιοθετήσει στον πίνακα M . Δηλαδή, το $p(X_{k+1} = CD)$, το οποίο είναι στοιχείο του v_{k+1} , ισούται με το εσωτερικό γινόμενο της δεύτερης γραμμής του πίνακα μεταβάσεων καταστάσεων σε ένα βήμα, M , και του διανύσματος v_k . Επομένως συνολικά, για τη χρονική εξέλιξη του v_k ισχύει $v_{k+1} = Mv_k$ και τελικά $v_{k+1} = M^{k+1}v_0$, με v_0 το διάνυσμα με τις μη δεσμευμένες πιθανότητες το σύστημα να βρεθεί σε μία από τις καταστάσεις CC , CD , DC ή DD , στο βήμα όπου οι αντίπαλοι επιλέγουν τις αρχικές τους κινήσεις. Επομένως, ο πίνακας M^m είναι ο πίνακας των δεσμευμένων πιθανοτήτων σε m βήματα (πίνακας μεταβάσεων καταστάσεων σε m βήματα - m -step transition matrix) και περιλαμβάνει όλες τις δεσμευμένες πιθανότητες για να μεταβούμε από οποιαδήποτε από τις τρέχουσες καταστάσεις CC , CD , DC ή DD σε οποιαδήποτε από τις καταστάσεις CC , CD , DC ή DD , αλλά σε m βήματα του παιχνιδιού (εξισώσεις Chapman - Kolmogorov).

Αποδεικνύεται ότι όταν η αλυσίδα Markov είναι μη υποβιβάσιμη (irreducible) και εργοδική (ergodic) το διάνυσμα v_k των μη δεσμευμένων πιθανοτήτων συγκλίνει ($\lim_{k \rightarrow \infty} v_k = \pi$) στη στατική κατανομή (stationary distribution) π , η οποία δίνεται και από το συγκλίνον όριο $\pi = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} v_0 M^k$. Όταν υπάρχει το τελευταίο όριο, αποδεικνύεται ότι είναι ανεξάρτητο του v_0 . Προφανώς, για το π ισχύει $\pi M = \pi$ (από τη σχέση $v_{k+1} = Mv_k$, στο όριο $k \rightarrow \infty$) και το άθροισμα των στοιχείων του π ισούται με το 1, αφού μας δίνει τις μη δεσμευμένες πιθανότητες, στη μόνιμη κατάσταση, το σύστημα να βρεθεί σε κάποια από τις καταστάσεις CC , CD , DC ή DD . Τότε και τα στοιχεία κάθε στήλης του πίνακα M^m , με m μεγάλο, λαμβάνουν την ίδια τιμή, αφού η δεσμευμένη πιθανότητα μετάβασης σε κάποια κατάσταση σε m βήματα, γίνεται ανεξάρτητη από την τρέχουσα κατάσταση στην οποία βρίσκεται το σύστημα. Γνωρίζοντας για ένα παιχνίδι δύο αντιπάλων X και Y το διάνυσμα π με τις μακροπρόθεσμες πιθανότητες το σύστημα να βρεθεί σε κάποια από τις καταστάσεις CC , CD , DC ή DD , είναι προφανές ότι μπορούμε να υπολογίσουμε το μέσο αναμενόμενο pay-off σε ένα γύρο για καθένα από τους δύο αντιπάλους, από τις σχέσεις

$$s_X = \pi S_X = \pi \begin{bmatrix} R \\ S \\ T \\ P \end{bmatrix} \quad \text{και} \quad s_Y = \pi S_Y = \pi \begin{bmatrix} R \\ T \\ S \\ P \end{bmatrix} \quad \text{αντίστοιχα.}$$