

Στοχαστικά Μοντέλα
στην Επιχειρησιακή Έρευνα
Μέρος 2
Στοχαστικός Δυναμικός
Προγραμματισμός

Αντώνης Οικονόμου

Μεταπτυχιακό Πρόγραμμα Σπουδών
Στατιστική και Επιχειρησιακή Έρευνα

6/10/2022-20/10/2022

Εισαγωγή στον Δυναμικό Προγραμματισμό σε πεπερασμένο χρονικό ορίζοντα

Βασικό πλαίσιο Δυναμικού Προγραμματισμού

- Διακριτός χρόνος.
- Πεπερασμένος χρονικός ορίζοντας.
- Χώρος καταστάσεων αριθμήσιμος (θεωρία) (ή κάποιο διάστημα του $\mathbb{R}$ σε μερικές εφαρμογές).
- Δέσμη αποφάσεων σε κάθε κατάσταση πεπερασμένη (ή κάποιο συμπαγές υποσύνολο του $\mathbb{R}$).
- Ακολουθιακή λήψη αποφάσεων:

$$X_t \rightarrow A_t \rightarrow X_{t-1} \rightarrow A_{t-1} \rightarrow \cdots \rightarrow X_1 \rightarrow A_1 \rightarrow X_0.$$

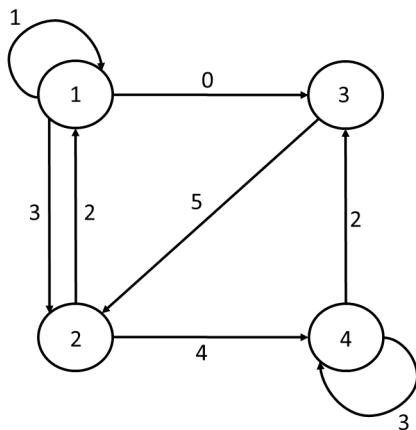
- Απόφαση $\rightarrow$ Κόστος.
- Στόχος: Ελαχιστοποίηση συνολικού κόστους.

Βασικές υποθέσεις Δυναμικού Προγραμματισμού

- Το κόστος σε κάθε στάδιο/βήμα εξαρτάται μόνο από τα παρόντα στάδιο, κατάσταση και απόφαση.
- Η επόμενη κατάσταση επιλέγεται μόνο με βάση τα παρόντα στάδιο, κατάσταση και απόφαση.

Παράδειγμα: Κίνηση οχήματος

- Όχημα κινείται πάνω στο οδικό δίκτυο.



Σχήμα: Δίκτυο κίνησης οχήματος.

Παράδειγμα: Κίνηση οχήματος (συνέχεια)

- Κόμβοι = Πόλεις.
- Ακμές = Απευθείας οδικές συνδέσεις.
- Το όχημα σε κάθε περίοδο επιλέγει μια από τις ακμές που έχουν αφετηρία τον κόμβο στον οποίον βρίσκεται και μετακινείται πάνω σε αυτή.
- Αριθμός ακμής = Κόστος μετακίνησης σε επί αυτής.
- Το όχημα την περίοδο 0 βρίσκεται στον κόμβο 1.
- Πρόκειται να κινηθεί για 4 περιόδους.
- Διαδρομή ελάχιστου κόστους ;

Αρχή Bellman

- Αν μια διαδρομή είναι βέλτιστη, τότε κάθε υποδιαδρομή της είναι βέλτιστη για τη σύνδεση των αντίστοιχων αρχικών τελικών κόμβων.
- Π.χ., στην κίνηση οχήματος, αν είναι βέλτιστη η διαδρομή $4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 3$, για μετακίνηση για 4 περιόδους, ξεκινώντας από την 4, τότε η διαδρομή $3 \rightarrow 2 \rightarrow 1 \rightarrow 3$ είναι βέλτιστη για μετακίνηση για 3 περιόδους, ξεκινώντας από την 3.

Κίνηση οχήματος (συνέχεια)

- Συνάρτηση βέλτιστης τιμής:
 $v_n(x)$: Ελάχιστο συνολικό κόστος μετακίνησης, ξεκινώντας από την κατάσταση x , μετά τη συμπλήρωση n μετακινήσεων.
- Αναδρομικός υπολογισμός:

$$v_0^*(x_0) = 0, \quad x_0 \in \{1, 2, 3, 4\}$$

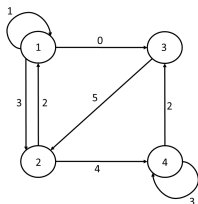
$$v_n^*(x_n) = \min_{a_n} (c(x_n, a_n) + v_{n-1}^*(g(x_n, a_n))), \quad x_n \in \{1, 2, 3, 4\}$$

x_n η κατάσταση στο στάδιο n ,

a_n η απόφαση στο στάδιο n ,

$g(x_n, a_n)$ η επόμενη κατάσταση αν στο στάδιο n , στην κατάσταση x_n , ληφθεί η απόφαση a_n .

Κίνηση οχήματος (συνέχεια)



Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 0$: $v_0(x_0) = 0$, $x_0 \in \{1, 2, 3, 4\}$.
- Για $n = 1$: $v_1(x) = \min_a c(x, a)$. Οπότε:

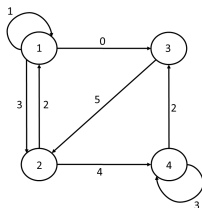
$$v_1(1) = 0, \quad a_1^*(1) = 3$$

$$v_1(2) = 2, \quad a_1^*(2) = 1$$

$$v_1(3) = 5, \quad a_1^*(3) = 2$$

$$v_1(4) = 2, \quad a_1^*(4) = 3$$

Κίνηση οχήματος (συνέχεια)



Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 2$: $v_2(x) = \min_a (c(x, a) + v_1(g(x, a)))$. Οπότε:

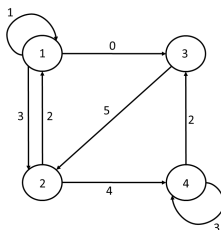
$$\begin{aligned} v_2(1) &= \min\{c(1, 1) + v_1(1), c(1, 2) + v_1(2), c(1, 3) + v_1(3)\} \\ &= \min\{1 + 0, 3 + 2, 0 + 5\} = 1, \quad a_2^*(1) = 1. \end{aligned}$$

$$v_2(2) = 2, \quad a_2^*(2) = 1$$

$$v_2(3) = 7, \quad a_2^*(3) = 2$$

$$v_2(4) = 5, \quad a_2^*(4) = 4$$

Κίνηση οχήματος (συνέχεια)



Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 3$:

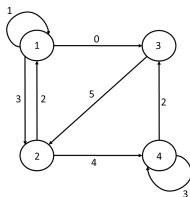
$$v_3(1) = 2, \quad a_3^*(1) = 1$$

$$v_3(2) = 3, \quad a_3^*(2) = 1$$

$$v_3(3) = 7, \quad a_3^*(3) = 2$$

$$v_1(4) = 8, \quad a_3^*(4) = 4$$

Κίνηση οχήματος (συνέχεια)



Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 4$:

$$v_4(1) = 3, \quad a_4^*(1) = 1$$

$$v_4(2) = 4, \quad a_4^*(2) = 1$$

$$v_4(3) = 8, \quad a_4^*(3) = 2$$

$$v_4(4) = 9, \quad a_4^*(4) = 3.$$

- Ελάχιστο δυνατό κόστος σε 4 βήματα ξεκινώντας από τον κόμβο 1: $v_4(1) = 3$: $1 - 1 - 1 - 1 - 1$.
- Από κόμβο 4: $v_4(4) = 9$: $4 - 3 - 2 - 1 - 3$.

Γενικό πλαίσιο Μαρκοβιανών Διαδικασιών Αποφάσεων πεπερασμένου ορίζοντα

- N : πλήθος σταδίων (συνολικό πλήθος αποφάσεων που θα ληφθούν).
- n : στάδιο (πλήθος αποφάσεων που απομένουν να ληφθούν), $n = N, N - 1, N - 2, \dots, 0$.
- x_n : κατάσταση στο στάδιο n .
- a_n : απόφαση στο στάδιο n (μετά την παρατήρηση της x_n).
- $\mathcal{X}$: αριθμήσιμος χώρος καταστάσεων συστήματος.
- $\mathcal{A}(x_n; n)$: πεπερασμένη δέσμη αποφάσεων στην κατάσταση x_n στο στάδιο n .

Γενικό πλαίσιο Μαρκοβιανών Διαδικασιών

Αποφάσεων πεπερασμένου ορίζοντα (Συνέχεια)

- Δομή κόστους:

$c(x, a; n)$: Άμεσο κόστος, αν όντας στο στάδιο n , στην κατάσταση x , ληφθεί η απόφαση a .

$\tilde{c}(x)$: Τερματικό κόστος που επάγεται στο τελικό στάδιο (στάδιο 0), όταν δεν μένουν πια αποφάσεις και η κατάσταση είναι x .

- Δυναμική συστήματος:

$p_{xy}(a; n)$: Πιθανότητα η επόμενη κατάσταση να είναι η y , αν όντας στο στάδιο n , στην κατάσταση x , ληφθεί η απόφαση a .

Χρονικά ομογενή προβλήματα

- Σε αρκετά προβλήματα τα άμεσα κόστη και οι πιθανότητες μετάβασης δεν εξαρτώνται από το στάδιο στο οποίο βρισκόμαστε, οπότε μιλάμε για χρονικά ομογενές πρόβλημα.
- Στην περίπτωση αυτή τα συμβολίζουμε με $c(x, a)$ και $p_{xy}(a)$ αντίστοιχα.

Κριτήριο Βελτιστοποίησης

- Ζητούμενο: Εύρεση βέλτιστης πολιτικής π .
- Κριτήριο βελτιστότητας: Ελαχιστοποίηση συνολικού κόστους.
- Αντικειμενική συνάρτηση:

$$E_{\pi}[c(X_N, A_N; N) + c(X_{N-1}, A_{N-1}; N-1) + \dots + c(X_1, A_1; 1) + \tilde{c}(X_0) | X_N = x].$$

Βασικές υποθέσεις μοντέλου Μαρκοβιανών Διαδικασιών Αποφάσεων

- 1 Μαρκοβιανή δυναμική: Η επόμενη κατάσταση εξαρτάται μόνο από την τρέχουσα κατάσταση και την τρέχουσα απόφαση.
- 2 Μαρκοβιανή δομή κόστους: Το άμεσο κόστος εξαρτάται μόνο από την τρέχουσα κατάσταση και την τρέχουσα απόφαση.
- 3 Προσθετική συνάρτηση κόστους: Τα κόστη των διαφόρων περιόδων αθροίζονται.
- 4 Τέλεια παρατηρησιμότητα: Ο ελεγκτής γνωρίζει ακριβώς την κατάσταση πριν τη λήψη κάθε απόφασης.

Στοιχεία ορισμού Μαρκοβιανής διαδικασίας αποφάσεων

- Μαρκοβιανή Διαδικασία Αποφάσεων Πεπερασμένου Ορίζοντα:

$$(\mathcal{X}, \mathcal{A}(x) : x \in \mathcal{X}, c, p, \tilde{c}, N, x_0).$$

- $\mathcal{X}$: Χώρος καταστάσεων.
- $\mathcal{A}(x) : x \in \mathcal{X}$: Δέσμες αποφάσεων.
- c : Συνάρτηση άμεσου κόστους.
- p : Συνάρτηση πιθανοτήτων μετάβασης.
- $\tilde{c}$: Συνάρτηση τερματικού κόστους.
- N : Διάρκεια χρονικού ορίζοντα.
- x_0 : Αρχική κατάσταση.

Έννοια ιστορίας

- Ιστορία σε ένα πρόβλημα N σταδίων στο στάδιο n :

$$h_N = x_N a_N x_{N-1} a_{N-1} \cdots x_{n+1} a_{n+1} x_n$$

ακολουθία καταστάσεων-αποφάσεων μέχρι το στάδιο n .

Έννοια πολιτικής

- Πολιτική:

$$\pi = (\pi(N), \pi(N-1), \dots, \pi(1))$$

σύνολο κανόνων, ένας για κάθε στάδιο, για το ποια απόφαση θα ληφθεί.

- $\pi(n)$: κανόνας λήψης απόφασης για το στάδιο n .
- $\pi_{h_n}(a; n)$: πιθανότητα επιλογής της απόφασης a δεδομένης της ιστορίας h_n στο στάδιο n .
- Πεδίο ορισμού: το σύνολο όλων των δυνατών ιστοριών στο στάδιο n .
- Πεδίο τιμών: το σύνολο των κατανομών πιθανότητας στο σύνολο των δυνατών αποφάσεων.

Κατηγοριοποίηση πολιτικών ως προς την εξάρτηση από την ιστορία

- π ιστοριοεξαρτώμενη (history-dependent - H) στη γενική περίπτωση.
- π Μαρκοβιανή (Markovian - M), αν για κάθε στάδιο n και ιστορία $h_n = x_N a_N x_{N-1} a_{N-1} \cdots x_{n+1} a_{n+1} x_n$, είναι

$$\pi_{h_n}(a; n) = \pi_{x_n}(a; n).$$

- π στάσιμη (stationary - S), αν είναι Μαρκοβιανή και για κάθε στάδιο n και κατάσταση x_n , είναι

$$\pi_{x_n}(a; n) = \pi_{x_n}(a).$$

Κατηγοριοποίηση πολιτικών ως προς την επιλογή

- π τυχαιοποιημένη (randomized -R) στη γενική περίπτωση.
- π προσδιοριστική, αν για κάθε στάδιο και ιστορία h_n , υπάρχει μια απόφαση $a_n(h_n)$ που επιλέγεται με βεβαιότητα.
Δηλαδή:

$$\begin{aligned}\pi_{h_n}(a_n(h_n); n) &= 1 \\ \pi_{h_n}(a; n) &= 0, \quad a \neq a_n(h_n).\end{aligned}$$

Κλάσεις πολιτικών

- H , M και S : ιστοριοεξαρτώμενες, Μαρκοβιανές και στάσιμες πολιτικές.
- R και D : τυχαιοποιημένες ντετερμινιστικές πολιτικές.
- Κλάσεις πολιτικών: Π^{HR} , Π^{HD} , Π^{MR} , Π^{MD} , Π^{SR} , Π^{SD} .
- Σχέσεις:

$$\begin{aligned}\Pi^{SD} &\subseteq \Pi^{SR} \subseteq \Pi^{MR} \subseteq \Pi^{HR}, \\ \Pi^{SD} &\subseteq \Pi^{MD} \subseteq \Pi^{MR} \subseteq \Pi^{HR}, \\ \Pi^{SD} &\subseteq \Pi^{MD} \subseteq \Pi^{HD} \subseteq \Pi^{HR}.\end{aligned}$$