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:

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

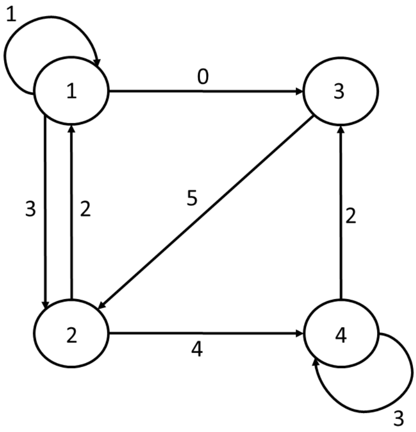
- Απόφαση $\to$ Κόστος.
- Στόχος: Ελαχιστοποίηση συνολικού κόστους.

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 $4 \to 3 \to 2 \to 1 \to 3$, για μετακίνηση για 4 περιόδους, ξεκινώντας από την 4,
  τότε η διαδρομή $3 \to 2 \to 1 \to 3$ είναι βέλτιστη για μετακίνηση για 3 περιόδους, ξεκινώντας από την 3.

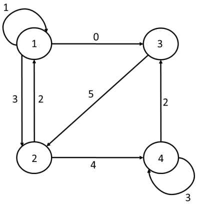
Κίνηση οχήματος (συνέχεια)

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

$$
\begin{aligned}
v_0^*(x_0) &= 0,\ \ x_0 \in \{1,2,3,4\} \\
v_n^*(x_n) &= \min_{a_n} \left(c(x_n,a_n) + v_{n-1}^*(g(x_n,a_n)) \right),\ \ x_n \in \{1,2,3,4\}
\end{aligned}
$$

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)

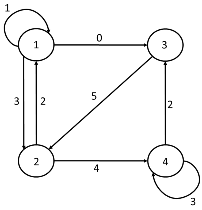


Σχήμα: Δίκτυο κίνησης οχήματος.

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

$$\begin{aligned} v_1(1) &= 0, \quad a_1^*(1) = 3 \\ v_1(2) &= 2, \quad a_1^*(2) = 1 \\ v_1(3) &= 5, \quad a_1^*(3) = 2 \\ v_1(4) &= 2, \quad a_1^*(4) = 3 \end{aligned}$$

Κίνηση οχήματος (συνέχεια)



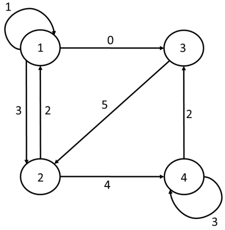
Σχήμα: Δίκτυο κίνησης οχήματος.

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

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

$$\begin{aligned} v_2(2) &= 2, \;\; a_2^*(2) = 1 \\ v_2(3) &= 7, \;\; a_2^*(3) = 2 \\ v_2(4) &= 5, \;\; a_2^*(4) = 4 \end{aligned}$$

Κίνηση οχήματος (συνέχεια)

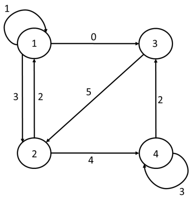


Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 3$:

$$
\begin{aligned}
v_3(1) &= 2, \quad a_3^*(1) = 1 \\
v_3(2) &= 3, \quad a_3^*(2) = 1 \\
v_3(3) &= 7, \quad a_3^*(3) = 2 \\
v_1(4) &= 8, \quad a_3^*(4) = 4
\end{aligned}
$$

Κίνηση οχήματος (συνέχεια)



Σχήμα: Δίκτυο κίνησης οχήματος.

- Για $n = 4$:

$$\begin{aligned} v_4(1) &= 3, \quad a_4^*(1) = 1 \\ v_4(2) &= 4, \quad a_4^*(2) = 1 \\ v_4(3) &= 8, \quad a_4^*(3) = 2 \\ v_4(4) &= 9, \quad a_4^*(4) = 3. \end{aligned}$$

- 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), $n = N, N-1, N-2, \ldots, 0$.
- 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:

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

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:

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

  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:

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

Θεωρία Δυναμικού Προγραμματισμού σε πεπερασμένο χρονικό ορίζοντα

Υπενθύμιση πλαισίου I

- N: πλήθος σταδίων.
- $\mathcal{X}$: χώρος καταστάσεων.
- n: στάδιο.
- x_n: κατάσταση στο στάδιο n.
- a_n: απόφαση στο στάδιο n.
- $\mathcal{A}(x_n; n)$: δέσμη αποφάσεων στην κατάσταση x_n στο στάδιο n.
- $p_{x_n, x_{n-1}}(a_n; n) = \Pr[X_{n-1} = x_{n-1} | X_n = x_n, A_n = a_n]$: Πιθανότητες μετάβασης.
- $c(x_n, a_n; n)$: Άμεσα κόστη.
- $\tilde{c}(x)$: Τερματικά κόστη.

Υπενθύμιση πλαισίου ΙΙ

- Πολιτική:

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

- $\pi_{h_n}(a; n)$: 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.

Συνάρτηση ολικής αξίας πολιτικής

- Δοθείσης μιας πολιτικής $\pi = (\pi(N), \pi(N-1), \ldots, \pi(1))$ και αρχικής κατάστασης x_N:

$$u_N^\pi(x_N) \;=\; E_\pi\left[\sum_{n=1}^{N} c(X_n, A_n; n) + \tilde{c}(X_0)\,|X_N = x_N\right].$$

η συνάρτηση της (ολικής) αξίας της.

Συνάρτηση μερικής αξίας πολιτικής

- Δοθείσης μιας πολιτικής $\pi = (\pi(N), \pi(N-1), \ldots, \pi(1))$ και ιστορίας h_n:

$$u_n^\pi(h_n) = E_\pi\left[\sum_{k=1}^{n} c(X_k, A_k; k) + \tilde{c}(X_0)\,|H_n = h_n\right].$$

η συνάρτηση μερικής αξίας της, από το στάδιο n και μετά δεδομένου ότι έχει παρατηρηθεί η ιστορία h_n.

Αναδρομικός υπολογισμός αξίας πολιτικής

- Αναδρομικό σχήμα υπολογισμού αξίας πολιτικής:

$$
\begin{aligned}
u_0^\pi(h_0) &= \tilde{c}(x_0), \\
u_n^\pi(h_n) &= \sum_{a_n} \pi_{h_n}(a_n; n) \cdot \Bigg[c(x_n, a_n; n) \\
&\quad + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) u_{n-1}^\pi(h_n a_n x_{n-1}) \Bigg],
\end{aligned}
$$

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

- Απόδειξη: Δέσμευση στην απόφαση A_n και στην κατάσταση X_{n-1}.

Συνάρτηση βέλτιστης τιμής

- Έστω πρόβλημα N σταδίων.
- Συνάρτηση βέλτιστης τιμής στο στάδιο n, ως προς την τρέχουσα κατάσταση x_n:

$$v_n(x_n) \quad = \quad \inf_{\pi, x_N, a_N, \ldots, x_{n+1}, a_{n+1}} u_n^\pi(x_N a_N \ldots x_{n+1} a_{n+1} x_n).$$

- $v_n(x_n)$: μέγιστο κάτω φράγμα (infimum) για το κόστος από το στάδιο n και μετά, δεδομένου ότι η τρέχουσα κατάσταση είναι x_n.

Βασικό αποτέλεσμα Δυναμικού Προγραμματισμού σε πεπερασμένο ορίζοντα - προς απόδειξη

- Υπάρχει βέλτιστη πολιτική, δηλαδή π τέτοια ώστε:

$$u_n^\pi(\cdots x_n) = v_n(x_n)$$

  που είναι Μαρκοβιανή Ντετερμινιστική MD.
- Η συνάρτηση βέλτιστης τιμής υπολογίζεται αναδρομικά

$$\begin{aligned} v_0(x_0) &= \tilde{c}(x_0), \\ v_n(x_n) &= \min_{a_n} \left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1}) \right] \end{aligned}$$

  και η βέλτιστη πολιτική επιλέγει σε κάθε κατάσταση την απόφαση που υλοποιεί το min.

Ανισότητα συνάρτησης βέλτιστης τιμής

- Η συνάρτηση βέλτιστης τιμής $v_n(x_n)$ ικανοποιεί την ανισότητα

$$v_n(x_n) \geq \min_{a_n} \left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1}) \right].$$

Ανισότητα συνάρτησης βέλτιστης τιμής - απόδειξη

π αυθαίρετη πολιτική, στάδιο n και ιστορία

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

που οδηγεί στην κατάσταση x_n στο στάδιο n.Τότε:

$$\begin{aligned}
u_n^\pi(h_n) &= \sum_{a_n} \pi_{h_n}(a_n;n)\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) u_{n-1}^\pi(h_n a_n x_{n-1})\right] \\
&\geq \sum_{a_n} \pi_{h_n}(a_n;n)\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) v_{n-1}(x_{n-1})\right] \\
&\geq \sum_{a_n} \pi_{h_n}(a_n;n) \underbrace{\min_{a_n}\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) v_{n-1}(x_{n-1})\right]}_{\text{Ανεξάρτητο των } a_n, h_n, n} \\
&= \min_{a_n}\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) v_{n-1}(x_{n-1})\right],
\end{aligned}$$

και παίρνουμε inf για όλα τα π, h_{n+1} και a_{n+1}.

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα

- Ορίζουμε τις συναρτήσεις $u_n^*(x_n)$, $n = 0, 1, 2, \dots, N$:

$$
\begin{aligned}
u_0^*(x_0) &= \tilde{c}(x_0),\ x_0 \in \mathcal{X}, \\
u_n^*(x_n) &= \min_{a_n} \left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) u_{n-1}^*(x_{n-1}) \right],\ x_n \in \mathcal{X}.
\end{aligned}
$$

- $a_n^*(x_n)$: η απόφαση που πετυχαίνει το ελάχιστο στο δεξιό μέλος, $n = 1, \dots, N, x_n \in \mathcal{X}$.
- $\pi^* = (\pi^*(N), \pi^*(N-1), \dots, \pi^*(1))$ η MD πολιτική που στην κατάσταση x_n στο στάδιο n επιλέγει την $a_n^*(x_n)$.
- Τότε, για κάθε ιστορία $h_n = x_N a_N \cdots x_{N+1} a_{N+1} x_n$:

$$
u_n^{\pi^*}(h_n) = u_n^*(x_n) = v_n(x_n).
$$

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα (συνέχεια)

- Επομένως, η π^* είναι βέλτιστη.
- Επιπλέον, η συνάρτηση βέλτιστης τιμής ικανοποιεί τις εξισώσεις βελτιστότητας

$$\begin{aligned} v_0(x_0) &= \tilde{c}(x_0),\ x_0 \in \mathcal{X}, \\ v_n(x_n) &= \min_{a_n} \left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1}) \right] \end{aligned}$$

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα - Απόδειξη

- Αποδεικνύουμε με επαγωγή την

$$u_n^{\pi^*}(h_n) = u_n^*(x_n) = v_n(x_n).$$

- Για $n = 0$ έχουμε ότι

$$u_0^{\pi^*}(h_0) = \tilde{c}(x_0) = u_0^*(x_0) = v_0(x_0), \; x_0 \in \mathcal{X},$$

  οπότε η πρόταση ισχύει.
- Έστω ότι ισχύει για $n-1$. Θα αποδείξουμε ότι ισχύει και για n.

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα - Απόδειξη (συνέχεια)

- Από την αναδρομική σχέση υπολογισμού αξίας πολιτικής έχουμε:

$$u_n^{\pi^*}(h_n) = \sum_{a_n} \pi^*_{h_n}(a_n;n)\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) u_{n-1}^{\pi^*}(h_n a_n x_{n-1})\right].$$

- Όμως από τον ορισμό της π^* έχουμε

$$\pi^*_{h_n}(a_n;n) = \begin{cases} 1, & \text{αν } a_n = a_n^* \\ 0, & \text{διαφορετικά,} \end{cases}$$

- Επίσης, από την επαγωγική υπόθεση:

$$u_{n-1}^{\pi^*}(h_n a_n x_{n-1}) = u_{n-1}^*(x_{n-1}) = v_{n-1}(x_{n-1}).$$

- Άρα ο τύπος για την $u_n^{\pi^*}(h_n)$ γίνεται:

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα - Απόδειξη (συνέχεια)

- $$\begin{aligned} u_n^{\pi^*}(h_n) &= \min_{a_n}\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) u_{n-1}^*(x_{n-1})\right] \\ &= \min_{a_n}\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) v_{n-1}(x_{n-1})\right]. \end{aligned}$$

- Από την πρώτη ισότητα και τον αναδρομικό ορισμό της $u_n^*(x_n)$

$$u_n^*(x_n) = \min_{a_n}\left[c(x_n,a_n;n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n;n) u_{n-1}^*(x_{n-1})\right],$$

παίρνουμε $u_n^{\pi^*}(h_n) = u_n^*(x_n)$.

Βασικό θεώρημα Δυναμικού Προγραμματισμού πεπερασμένου ορίζοντα - Απόδειξη (συνέχεια)

$$\begin{aligned} u_n^{\pi^*}(h_n) &= \min_{a_n}\left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) u_{n-1}^*(x_{n-1})\right] \\ &= \min_{a_n}\left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1})\right]. \end{aligned}$$

- Από τη δεύτερη ισότητα και τον ορισμό της $v_n(x_n)$:

$$v_n(x_n) \leq \min_{a_n}\left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1})\right],$$

  που μαζί με την ανισότητα βέλτιστης τιμής δίνει

$$v_n(x_n) = \min_{a_n}\left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1})\right].$$

  Συγκρίνοντας δεξιά μέλη: $u_n^{\pi^*}(h_n) = v_n(x_n) = u_n^*(x_n)$.

Γενική μεθοδολογία επίλυσης

- Γράφουμε τις εξισώσεις βελτιστοποίησης:

$$
\begin{aligned}
v_0(x_0) &= \tilde{c}(x_0),\ x_0 \in \mathcal{X}, \\
v_n(x_n) &= \min_{a_n} \left[c(x_n, a_n; n) + \sum_{x_{n-1}} p_{x_n x_{n-1}}(a_n; n) v_{n-1}(x_{n-1}) \right]
\end{aligned}
$$

- Υπολογίζουμε για μικρά $n = 1, 2, \ldots$ και εικάζουμε τη μορφή της $v_n(x_x)$ και την $a_n^*(x_n)$.
- Αποδεικνύουμε επαγωγικά την εικασία.

Μοντέλο κατανάλωσης-επένδυσης - Περιγραφή

- Συσσωρευμένο κεφάλαιο → Ετήσιο εισόδημα.
- Αν το ετήσιο εισόδημα είναι x, τότε πρέπει να αποφασιστεί η ποσότητα προς κατανάλωση a, $0 \leq a \leq x$.
- Αν καταναλώθεί a, τότε το επόμενο έτος το εισόδημα θα είναι $x + \theta(x - a)$,
  (ενσωμάτωση αδιάθετου εισοδήματος στο κεφάλαιο με επιτόκιο θ).
- Σκοπός: Μεγιστοποίηση συνολικού ποσού κατανάλωσης για τα επόμενα N χρόνια, δοθέντος ότι το φετινό ετήσιο εισόδημα είναι x.

Μοντέλο κατανάλωσης-επένδυσης - Μοντελοποίηση

- Χρονικός ορίζοντας: ο αριθμό N των ετών προγραμματισμού.
- Στάδιο: ο αριθμός n των υπόλοιπων ετών.
- Κατάσταση: Το ετήσιο εισόδημα x στην αρχή ενός έτους.
- Απόφαση: Το ποσό a που θα καταναλωθεί.
- Δυναμική: $x \xrightarrow{a} x + \theta(x - a)$.
- Συνάρτηση βέλτιστης τιμής: $v_n(x)$ η μέγιστη ωφέλεια (συνολική κατανάλωση), ξεκινώντας από την κατάσταση x, όταν απομένουν n έτη.

Μοντέλο κατανάλωσης-επένδυσης - Εξισώσεις βελτιστοποίησης

- Η συνάρτηση βέλτιστης τιμής υπολογίζεται αναδρομικά από την

$$
\begin{aligned}
v_0(x) &= 0,\ x \geq 0,\\
v_n(x) &= \max_{a\in[0,x]} \left[a + v_{n-1}(x+\theta(x-a))\right],\\
&\quad x \geq 0,\ n = 1, 2, \ldots .
\end{aligned}
$$

Μοντέλο κατανάλωσης-επένδυσης - Υπολογισμός συνάρτησης βέλτιστης τιμής, $n = 0, 1$

- Για $n = 0$, από την αρχική συνθήκη

$$v_0(x) = 0, \;\; x \geq 0,$$

- Αντικαθιστώντας στην εξίσωση βελτιστοποίησης για $n = 1$, έχουμε

$$v_1(x) = \max_{a \in [0,x]} [a + 0] = x, \;\; x \geq 0.$$

Μοντέλο κατανάλωσης-επένδυσης - Υπολογισμός συνάρτησης βέλτιστης τιμής, $n = 2$

- Αντικαθιστώντας στην εξίσωση για $n = 2$, έχουμε

$$\begin{aligned} v_2(x) &= \max_{a \in [0,x]} \left[a + x + \theta(x-a)\right] \\ &= (1+\theta)x + \max_{a \in [0,x]} \left[(1-\theta)a\right]. \end{aligned}$$

- $(1-\theta)a$ γραμμική με μέγιστο στο 0 ή στο x:

$$\begin{aligned} v_2(x) &= \begin{cases} (1+\theta)x + (1-\theta)0 & \text{αν } \theta \geq 1, \\ (1+\theta)x + (1-\theta)x & \text{αν } \theta \leq 1, \end{cases} \\ &= \begin{cases} (1+\theta)x & \text{αν } \theta \geq 1, \\ 2x & \text{αν } \theta \leq 1, \end{cases} \\ &= \underbrace{\max[1+\theta, 2]}_{\rho_2} x = \rho_2 x, x \geq 0. \end{aligned}$$

Μοντέλο κατανάλωσης-επένδυσης - Εικασία για συνάρτηση βέλτιστης τιμής

- Εικάζουμε ότι $v_n(x) = \rho_n x$, $x \geq 0$.
- Απόδειξη με επαγωγή στο n.
- Για $n = 0$ ισχύει προφανώς με $\rho_0 = 0$.
- Έστω ότι ισχύει για $n - 1$. Τότε από την εξίσωση βελτιστοποίησης έχουμε

$$\begin{aligned} v_n(x) &= \max_{a \in [0,x]} \left[a + \rho_{n-1}(x + \theta(x - a)) \right] \\ &= (1+\theta)\rho_{n-1} x + \max_{a \in [0,x]} \left[(1 - \theta\rho_{n-1}) a \right], \ x \geq 0. \end{aligned}$$

- $(1 - \theta\rho_{n-1})a$ γραμμική συνάρτηση του a $\rightarrow$ μέγιστη τιμή στο $[0, x]$ λαμβάνεται στο 0 ή στο x.

Μοντέλο κατανάλωσης-επένδυσης - Εικασία για συνάρτηση βέλτιστης τιμής (συνέχεια)

- Επομένως:

$$\begin{aligned} v_n(x) &= (1+\theta)\rho_{n-1}x + \max_{a\in\{0,x\}}[(1-\theta\rho_{n-1})a] \\ &= \underbrace{\max[(1+\theta)\rho_{n-1}, 1+\rho_{n-1}]}_{\rho_n} x \\ &= \rho_n x,\ \ x \geq 0. \end{aligned}$$

- Η εικασία αποδείχθηκε και, επιπλέον, βρέθηκε ένας αναδρομικός τύπος για τον υπολογισμό των σταθερών ρ_n.

Μοντέλο κατανάλωσης-επένδυσης - Συνάρτηση βέλτιστης τιμής

- Η συνάρτηση βέλτιστης τιμής στο μοντέλο κατανάλωσης - επένδυσης δίνεται από τον τύπο

$$v_n(x) = \rho_n x,\ x \geq 0,\ n = 0, 1, 2, \ldots,$$

όπου

$$\begin{aligned} \rho_0 &= 0, \\ \rho_n &= \max[(1+\theta)\rho_{n-1}, 1+\rho_{n-1}],\ n = 1, 2, \ldots. \end{aligned}$$

Μοντέλο κατανάλωσης-επένδυσης - Βέλτιστη πολιτική

- Έχουμε δει ότι

$$v_n(x) = (1+\theta)\rho_{n-1}x + \max_{a\in\{0,x\}} \left[(1-\theta\rho_{n-1})a\right], \ x \geq 0.$$

- Στο στάδιο n, η βέλτιστη πολιτική είναι $a_n^*(x) = 0$ (αποταμίευση όλου του ποσού), αν

$$1 - \theta\rho_{n-1} \leq 0 \Leftrightarrow \rho_{n-1} \geq \frac{1}{\theta}.$$

- Άρα:

$$a_n^*(x) = \begin{cases} 0 & \text{αν } \rho_{n-1} \geq \frac{1}{\theta}, \\ x & \text{αν } \rho_{n-1} \leq \frac{1}{\theta}. \end{cases}$$

Μοντέλο κατανάλωσης-επένδυσης - Βέλτιστη πολιτική (συνέχεια)

- Από αναδρομικό τύπο $\rho_n = \max[(1+\theta)\rho_{n-1}, 1+\rho_{n-1}]$, έχουμε ρ_n αύξουσα.
- Έστω n^* ο ελάχιστος n για τον οποίο $\rho_n \geq \frac{1}{\theta}$.
- Τότε, έχουμε $\rho_n < \frac{1}{\theta}$ για $n \leq n^* - 1$, ενώ $\rho_n \geq \frac{1}{\theta}$ για $n \geq n^*$.
- Επίσης,

$$a_n^*(x) = \begin{cases} 0 & \text{αν } n \geq n^* + 1, \\ x & \text{αν } n \leq n^*. \end{cases}$$

Μοντέλο κατανάλωσης-επένδυσης - Βέλτιστη πολιτική (συνέχεια)

- Υπολογισμός των ρ_n:

$$
\begin{aligned}
\rho_0 &= 0, \\
\rho_n &= \rho_{n-1} + \max(1, \theta\rho_{n-1}) \\
&= \begin{cases} \rho_{n-1} + 1 & n \leq n^*, \\ \rho_{n-1}(1+\theta) & n \geq n^* + 1. \end{cases}
\end{aligned}
$$

- Άρα

$$
\rho_n = \begin{cases} n & n \leq n^*, \\ n^*(1+\theta)^{n-n^*} & n \geq n^* + 1. \end{cases}
$$

- Άρα ο n^*, ο ελάχιστος n με $\rho_n \geq \frac{1}{\theta}$, είναι ο

$$
n^* = \left\lceil \frac{1}{\theta} \right\rceil .
$$

Μοντέλο κατανάλωσης-επένδυσης - Βέλτιστη πολιτική (τελική μορφή)

- Έστω
$$n^* = \left\lceil \frac{1}{\theta} \right\rceil .$$
- Η βέλτιστη πολιτική κατανάλωσης είναι
$$a_n^*(x) = \begin{cases} 0 & \text{αν } n \geq n^* + 1, \\ x & \text{αν } n \leq n^*, \end{cases}$$
- Η συνάρτηση βέλτιστης τιμής είναι
$$v_n(x) = \rho_n x,\ x \geq 0,\ n = 0, 1, 2, \dots ,$$
όπου
$$\rho_n = \begin{cases} n & \text{αν } n \leq n^*, \\ n^*(1+\theta)^{n-n^*} & \text{αν } n \geq n^*. \end{cases}$$