

Wissen1_Preisalarm

Problemstellung

Preisvergleichsportale im Internet ermöglichen es, den aktuell verfügbaren günstigsten Preis für ein gewünschtes Produkt genau im Auge zu behalten. Ein lernfähiger Agent soll auf Basis des beobachteten günstigsten Preises ein Produkt online erwerben. Folgende Rahmenbedingungen charakterisieren die Aufgabenstellung:

- Der Agent kann jeden Tag die Entscheidung treffen, ob er zum tagesaktuellen Preis das Produkt kaufen möchte, oder ob er weiter warten will.
- Da das gewünschte Produkt zu einem bestimmten Zeitpunkt garantiert vorliegen soll, gibt es einen spätestmöglichen Zeitpunkt t_{final} , zu dem der Agent das Produkt kaufen muss. D.h. ein weiteres Warten (auf einen günstigeren Preis) ist dann nicht mehr möglich.
- Der Preis am ersten ($t = 1$) Tag sei $\mathcal{P}_0 \in \mathbb{N}$.
- Aufgrund der auf Angebot und Nachfrage basierenden Gesetze des Marktes nehmen wir an, dass der Preis des Produkts nicht negativ werden kann und sich von Tag zu Tag wie folgt ändert:

$$\mathcal{P}_{t+1} = \max(0, \mathcal{P}_t + k - \delta)$$

wobei sich der Wert $k \in \{0, \dots, 2\delta\}$ stochastisch gemäß einer Binomialverteilung $B(k|n, q)$ mit $n = 2\delta$ und $q \in [0, 1]$ ergibt. Hierbei steht δ für die maximale absolute Preisschwankung an einem Tag.

- Ziel des Agenten ist es, das Produkt zum bestmöglichen (geringsten) Preis zu kaufen. Kosten entstehen aus Sicht des Agenten also ausschließlich im Moment des Kaufes.

Problemspezifikation

MDP = $[T, S, A, p, c]$

T

N diskrete Entscheidungszeitpunkte: $T = \{1, 2, \dots, N\}$ (Problem mit endlichem Horizont)

A

Aktionsmenge: $A = \{a_0, a_1\}$

- a_0 = warten ; a_1 = kaufen
- A ist endlich.

S

Zustandsmenge:

$$S = \{(t, g) \mid t \in T, g \in \mathbb{W}(\mathcal{P}_t)\} \cup \{''0''\}$$

$$\mathbb{W}(\mathcal{P}_1) = \{\mathcal{P}_0\}, \mathcal{P}_0 \in \mathbb{N}$$

$$\forall t \in T : \mathbb{W}(\mathcal{P}_{t+1}) = \{\max(0, \mathcal{P}_t + k - \delta) \mid k \in \{0, \dots, 2\delta\}\}$$

- \mathbb{W} = Wertebereich
- "0" ist ein Terminalzustand mit folgenden Eigenschaften:

- "0" ist absorbierend (wird nicht mehr verlassen).
- Es fallen in ihm keine Kosten mehr an.
- Umgangssprachlich: Mit Ausnahme vom Terminalzustand, jeder Zustand stellt einen Tag im Beobachtungszeitraum und einen möglichen Wert des Produktpreises an diesem Tag dar.
- S ist endlich, denn T ist endlich und k ist beschränkt.

c

Direkte Kosten $c : (S \times A) \setminus \{((N, g), a_0) \mid g \in \mathbb{W}(\mathcal{P}_N)\} \rightarrow \mathbb{R}$

$$c(s, a) = \begin{cases} g_s, & \text{falls } s \neq "0" \text{ und } a = a_1 \\ 0, & \text{sonst} \end{cases}$$

Umgangssprachlich: Kosten entstehen ausschließlich im Moment des Kaufes.

p

Übergangswahrscheinlichkeiten $p : S \times A \times S \rightarrow [0, 1]$

- "0" ist absorbierend (wird nicht mehr verlassen):

$$\forall a \in A : \quad p("0", a, j) = \begin{cases} 1, & \text{falls } j = "0" \\ 0, & \text{sonst} \end{cases}$$

- Jeder Kauf (a_1) bewirkt den sofortigen Übergang in den Terminalzustand.:

$$\forall i, j \in S, i \neq "0" : \quad p(i, a_1, j) = \begin{cases} 1, & \text{falls } j = "0" \\ 0, & \text{sonst} \end{cases}$$

- Für jedes Warten (a_0):

$$\forall i, j \in S, i \neq "0" : \quad p(i, a_0, j) = \begin{cases} \sum_{m=0}^{\delta-g_i} \binom{2\delta}{m} q^m (1-q)^{2\delta-m}, & \begin{aligned} &\text{falls } t_j - t_i = 1 \\ &\text{und } \exists k \in \{0, \dots, 2\delta\} : g_j = \max(0, g_i + k - \delta) \\ &\text{und } g_j = 0 \\ &\text{und } \delta - g_i \geq 0 \end{aligned} \\ \binom{2\delta}{h-g_i+\delta} q^{h-g_i+\delta} (1-q)^{\delta-h+g_i}, & \begin{aligned} &\text{falls } t_j - t_i = 1 \\ &\text{und } \exists k \in \{0, \dots, 2\delta\} : g_j = \max(0, g_i + k - \delta) \\ &\text{und } g_j = h > 0 \end{aligned} \\ 0, & \text{sonst} \end{cases}$$

Begründung:

- j ist ein möglicher Folgezustand von i nach Ausführung der Aktion a_0 , gdw. $t_j - t_i = 1$ und $\exists k \in \{0, \dots, 2\delta\} : g_j = \max(0, g_i + k - \delta)$.
- Wenn die oberen Bedingungen erfüllt sind:
 - $g_j = 0 \Leftrightarrow g_i + k - \delta \leq 0 \Leftrightarrow k \leq \delta - g_i \Rightarrow P(g_j = 0) = P(k \leq \delta - g_i)$
Da $k \in \{0, \dots, 2\delta\}, k \sim B(k \mid 2\delta, q), q \in [0, 1]$, es folgt:
Wenn $\delta - g_i \geq 0$: $P(k \leq \delta - g_i) = \sum_{m=0}^{\delta-g_i} \binom{2\delta}{m} q^m (1-q)^{2\delta-m}$
 - Aus $g_j = h > 0$ folgt $h = g_i + k - \delta$ oder äquivalent $k = h - g_i + \delta \Rightarrow P(g_j = h) = P(k = h - g_i + \delta)$
Da $k \in \{0, \dots, 2\delta\}, k \sim B(k \mid 2\delta, q), q \in [0, 1]$, es folgt:
 $P(k = h - g_i + \delta) = \binom{2\delta}{h-g_i+\delta} q^{h-g_i+\delta} (1-q)^{\delta-h+g_i}$