12. Řízené markovské řetězce

12.1. Definice: Definice strategie a optimální strategie.

12.2. Věta: Věta o hledání optimální strategie.

12.3. Příklad: Je sledována výrobní linka, která se může nacházet buď v provozu (stav 0) nebo v opravě (stav 1). Ve stavu 0 je možný provoz „bez kontroly agregátů“ (strategie 1) nebo „s kontrolou agregátů“ (strategie 2). Ve stavu 1 je možno rozlišit opravu „bez výměny agregátů“ (strategie 1) nebo „s výměnou agregátů“ (strategie 2). Matice přechodu a matice výnosů jsou následující:

$$^{1}P=\begin{pmatrix}0{,}5 & 0{,}5\\ 0{,}2 & 0{,}8\end{pmatrix}, {}^{1}R=\begin{pmatrix}2 & 0\\ 1 & -5\end{pmatrix}, {}^{2}P=\begin{pmatrix}0{,}4 & 0{,}6\\ 0{,}4 & 0{,}6\end{pmatrix}, {}^{2}R=\begin{pmatrix}3 & 1\\ 2 & -2\end{pmatrix}.$$

Pro první tři kroky najděte maximální střední hodnotu celkového výnosu a optimální strategii.

Řešení:

$$^{1}q_0={}^{1}p_{00}\,{}^{1}r_{00}+{}^{1}p_{01}\,{}^{1}r_{01}=0{,}5\cdot 2+0{,}5\cdot 0=1, {}^{1}q_1={}^{1}p_{10}\,{}^{1}r_{10}+{}^{1}p_{11}\,{}^{1}r_{11}=0{,}2\cdot 1+0{,}8\cdot(-5)=-0{,}2$$

$$^{2}q_0={}^{2}p_{00}\,{}^{2}r_{00}+{}^{2}p_{01}\,{}^{2}r_{01}=0{,}4\cdot 3+0{,}6\cdot 1=1{,}8, {}^{2}q_1={}^{2}p_{10}\,{}^{2}r_{10}+{}^{2}p_{11}\,{}^{2}r_{11}=0{,}4\cdot 2+0{,}6\cdot(-2)=-0{,}4$$

$$^{1}q=\begin{pmatrix}1\\ -0{,}2\end{pmatrix}, {}^{2}q=\begin{pmatrix}1{,}8\\ -0{,}4\end{pmatrix}$$

$$v_0(1)=\max\{1;1{,}8\}=1{,}8\Rightarrow d_0^{*}(1)=2, v_1(1)=\max\{-0{,}2;-0{,}4\}=-0{,}2\Rightarrow d_1^{*}(1)=1$$

$$v_0(2)=\max\left\{{}^{1}q_0+{}^{1}p_{00}v_0(1)+{}^{1}p_{01}v_1(1);{}^{2}q_0+{}^{2}p_{00}v_0(1)+{}^{2}p_{01}v_1(1)\right\}=$$
$$=\max\{1+0{,}5\cdot 1{,}8+0{,}5\cdot(-0{,}2);1{,}8+0.4\cdot 1{,}8+0{,}6\cdot(-0{,}2)\}=\max\{1{,}8;2{,}4\}=2{,}4\Rightarrow d_0^{*}(2)=2$$

$$v_1(2)=\max\left\{{}^{1}q_1+{}^{1}p_{10}v_0(1)+{}^{1}p_{11}v_1(1);{}^{2}q_1+{}^{2}p_{10}v_0(1)+{}^{2}p_{11}v_1(1)\right\}=$$
$$=\max\{-0{,}2+0{,}2\cdot 1{,}8+0{,}8\cdot(-0{,}2);-0{,}4+0.4\cdot 1{,}8+0{,}6\cdot(-0{,}2)\}=\max\{0;0{,}2\}=0{,}2\Rightarrow d_1^{*}(2)=2$$

$$v_0(3)=\max\left\{{}^{1}q_0+{}^{1}p_{00}v_0(2)+{}^{1}p_{01}v_1(2);{}^{2}q_0+{}^{2}p_{00}v_0(2)+{}^{2}p_{01}v_1(2)\right\}=$$
$$=\max\{1+0{,}5\cdot 2{,}4+0{,}5\cdot 0{,}2;1{,}8+0.4\cdot 2{,}4+0{,}6\cdot 0{,}2\}=\max\{2{,}3;2{,}88\}=2{,}88\Rightarrow d_0^{*}(3)=2$$

$$v_1(3)=\max\left\{{}^{1}q_1+{}^{1}p_{10}v_0(2)+{}^{1}p_{11}v_1(2);{}^{2}q_1+{}^{2}p_{10}v_0(2)+{}^{2}p_{11}v_1(2)\right\}=$$
$$=\max\{-0{,}2+0{,}2\cdot 2{,}4+0{,}8\cdot 0{,}2;-0{,}4+0.4\cdot 2{,}4+0{,}6\cdot 0{,}2\}=\max\{0{,}44;0{,}68\}=0{,}68\Rightarrow d_1^{*}(3)=2$$

Závěr: V prvním kroku je vektor optimálních strategií $\begin{pmatrix}2\\ 1\end{pmatrix}$, ve druhém a třetím kroku $\begin{pmatrix}2\\ 2\end{pmatrix}$.

12.4. Poznámka: Uvedená metoda hledání optimálních strategií se nazývá rekurentní metoda. Hodí se jen pro malý počet kroků. Pro větší počet kroků je vhodnější použít iterační metodu.

12.5. Poznámka: Popis iterační metody hledání optimální strategie

Vytvořující funkce posloupnosti vektorů $\{v(n)\}_{n=1}^{\infty}$ má tvar: $G_v(z) = \frac{z}{1-z}(\mathbf{I} - z\mathbf{P})^{-1}\mathbf{q}$. V teorii matic se dokazuje, že matice $(\mathbf{I} - z\mathbf{P})^{-1}$ obsahuje stacionární složku $\mathbf{S} = \begin{pmatrix} \mathbf{a} \\ \vdots \\ \mathbf{a} \end{pmatrix}$, kde **a** je stacionární vektor matice **P** a tranzientní složku **T**. Lze tedy psát

$$G_v(z) = \frac{z}{(1-z)^2}\mathbf{Sq} + \frac{z}{(1-z)\left(1-\frac{z}{c_1}\right)\cdot\cdots\cdot\left(1-\frac{z}{c_k}\right)}\mathbf{Tq} =$$

$$= \frac{z}{(1-z)^2}\mathbf{Sq} + \left(\frac{A}{1-z} + \frac{B_1}{1-\frac{z}{c_1}} + \cdots + \frac{B_k}{1-\frac{z}{c_k}}\right)\mathbf{Tq}$$

kde $c_i > 1$, $i = 1, ..., k$.

$\frac{z}{(1-z)^2}$ je vytvořující funkce posloupnosti $a_n = n$.

$A\frac{1}{1-z}$ je vytvořující funkce posloupnosti $a_n = A$ (nezávisí tedy na n).

$B_i\frac{1}{1-\frac{z}{c_i}}$ je vytvořující funkce posloupnosti $a_n = B_i\left(\frac{1}{c_i}\right)^n$. Protože $c_i > 1$, budou výrazy obsahující $\left(\frac{1}{c_i}\right)^n$ pro dostatečně velká n velmi malé. Pro dostatečně velká n tedy platí:

$\mathbf{v}(n) = n\mathbf{Sq} + A\mathbf{Tq}$. Označíme-li $\mathbf{Sq} = \mathbf{g}$ a $A\mathbf{Tq} = \mathbf{v}$, lze psát pro dostatečně velká n:
$\mathbf{v}(n) = n\mathbf{g} + \mathbf{v}$ neboli $v_i(n) = ng + v_i$, $i = 0, 1, ..., N$. Dosadíme-li toto vyjádření do rekurentního vztahu $v_i(n) = q_i + \sum_{j=0}^{N} p_{ij}v_j(n-1)$, dostaneme

$ng + v_i = q_i + \sum_{j=0}^{N} p_{ij}\left[(n-1]) + v_j\right]$. Upravíme:

$$ng + v_i = q_i + ng\sum_{j=0}^{N} p_{ij} - g\sum_{j=0}^{N} p_{ij} + \sum_{j=0}^{N} p_{ij}v_j = q_i + ng - g + \sum_{j=0}^{N} p_{ij}v_j \Rightarrow g + v_i = q_i + \sum_{j=0}^{N} p_{ij}v_j$$

Dostali jsme systém N + 1 rovnic pro N + 2 neznámých g, v_0, v_1, ..., v_N. Protože neznámých je o jednu více než rovnic, nelze určit skutečné hodnoty neznámých v_0, v_1, ..., v_N, kterým se říká váhy. Howard navrhl položit $v_N = 0$. Tím se počet neznámých sníží a lze vypočítat relativní váhy v_0, v_1, ..., v_{N-1}, které se od skutečných vah budou lišit jenom o konstantu, tedy jejich rozdíly budou stejné jako u skutečných vah. Např. rozdíl $v_0 - v_1$ lze interpretovat jako rozdíl mezi střední hodnotou výnosu procesu, který vyšel ze stavu 0 a střední hodnotou výnosu procesu, který vyšel ze stavu 1. Kritériem pro volbu strategie je výraz ${}^{k}q_i + \sum_{j=0}^{N} {}^{k}p_{ij}v_j$.

Algoritmus Howardova iteračního postupu:

1. krok: Pomocí veličin p_{ij}, q_i určíme veličiny g, v_i z rovnic $g + v_i = q_i + \sum_{j=0}^{N} p_{ij} v_j$, přičemž $v_N = 0$.
2. krok: Pro každý stav $i \in J$ najdeme strategii k^*, která maximalizuje výraz ${}^k q_i + \sum_{j=0}^{N} {}^k p_{ij} v_j$.
3. krok: Nalezená strategie k^* poskytne hodnoty ${}^{k*} q_i, {}^{k*} p_{ij}$ pro opakování kroků 1 a 2.

Algoritmus končí, jakmile vektor strategií je stejný ve dvou po sobě následujících iteracích. (Zpravidla stačí provést jen několik málo iterací.)

12.6. Příklad: Závod produkuje nějaký spotřební výrobek, u něhož lze rozeznat dva stavy: stav 0 – výrobek je úspěšný s dobrým odbytem a cenou, stav 1 – výrobek je neúspěšný, odbyt vázne a cena je nízká. Při 1. strategii vedení závodu neinvestuje ani do technického rozvoje ani do reklamy. Při této strategii je matice přechodu ${}^1\mathbf{P} = \begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}4 & 0{,}6 \end{pmatrix}$ a matice výnosů ${}^1\mathbf{R} = \begin{pmatrix} 9 & 3 \\ 3 & -7 \end{pmatrix}$. Při 2. strategii vedení závodu zajistí technický rozvoj a investuje do reklamy. Matice přechodu: ${}^2\mathbf{P} = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}7 & 0{,}3 \end{pmatrix}$, matice výnosů: ${}^2\mathbf{R} = \begin{pmatrix} 4 & 4 \\ 1 & -19 \end{pmatrix}$. (Při 2. strategii se vyšší náklady promítnou do zisku, proto výnos ${}^2r_{00}$ musí být nižší než ${}^1r_{00}$, stejně tak ${}^2r_{11}$ musí být nižší než ${}^1r_{11}$.) Pomocí iterační metody je třeba zjistit, jakou strategii doporučit vedení závodu, aby střední hodnota celkového výnosu byla maximální.

Řešení: Nejprve vypočítáme vektory ${}^1\mathbf{q} = ({}^1q_0, {}^1q_1)$ a ${}^2\mathbf{q} = ({}^2q_0, {}^2q_1)$.

${}^1q_0 = {}^1p_{00}\,{}^1r_{00} + {}^1p_{01}\,{}^1r_{01} = 0{,}5 \cdot 9 + 0{,}5 \cdot 3 = 6, {}^1q_1 = {}^1p_{10}\,{}^1r_{10} + {}^1p_{11}\,{}^1r_{11} = 0{,}4 \cdot 3 + 0{,}6 \cdot (-7) = -3$

${}^1\mathbf{q} = (6, -3)$

${}^2q_0 = {}^2p_{00}\,{}^2r_{00} + {}^2p_{01}\,{}^2r_{01} = 0{,}8 \cdot 4 + 0{,}2 \cdot 4 = 4, {}^2q_1 = {}^2p_{10}\,{}^2r_{10} + {}^2p_{11}\,{}^2r_{11} = 0{,}7 \cdot 1 + 0{,}3 \cdot (-19) = -5$

${}^2\mathbf{q} = (4, -5)$

1. iterace: zvolíme $d_0(1) = 1$, $d_1(1) = 1$ a vyřešíme systém rovnic (přitom položíme $v_1 = 0$)

$g + v_0 = {}^1q_0 + {}^1p_{00} v_0 + {}^1p_{01} v_1 : g + v_0 = 6 + 0{,}5 \cdot v_0$

$g + v_1 = {}^1q_1 + {}^1p_{10} v_0 + {}^1p_{11} v_1 : g = -3 + 0{,}4 \cdot v_0$

Řešením tohoto systému obdržíme $v_0 = 10$, $g = 1$.

2. iterace:

$i = 0$, $k = 1$: ${}^1q_0 + {}^1p_{00} v_0 + {}^1p_{01} v_1 = 6 + 0{,}5 \cdot 10 = 11$

$k = 2$: ${}^2q_0 + {}^2p_{00} v_0 + {}^2p_{01} v_1 = 4 + 0{,}8 \cdot 10 = 12$

$\max\{11, 12\} = 12$, tedy $d_0(2) = 2$

$i = 1$, $k = 1$: ${}^1q_1 + {}^1p_{10} v_0 + {}^1p_{11} v_1 = -3 + 0{,}4 \cdot 10 = 1$

$k = 2$: ${}^2q_1 + {}^2p_{10} v_0 + {}^2p_{11} v_1 = -5 + 0{,}7 \cdot 10 = 2$

$\max\{1, 2\} = 2$, tedy $d_1(2) = 2$

Výsledek 2. iterace dává vektor strategií (2, 2).

S tímto vektorem vyřešíme systém rovnic (přitom položíme $v_1 = 0$)

$g + v_0 = {}^2q_0 + {}^2p_{00}v_0 + {}^2p_{01}v_1 : g + v_0 = 4 + 0{,}8 \cdot v_0$
$g + v_1 = {}^2q_1 + {}^2p_{10}v_0 + {}^2p_{11}v_1 : g = -5 + 0{,}7 \cdot v_0$
Řešením tohoto systému obdržíme $v_0 = 10$, $g = 2$.

3. iterace:
$i = 0$, $k = 1$: ${}^1q_0 + {}^1p_{00}v_0 + {}^1p_{01}v_1 = 6 + 0{,}5 \cdot 10 = 11$
$k = 2$: ${}^2q_0 + {}^2p_{00}v_0 + {}^2p_{01}v_1 = 4 + 0{,}8 \cdot 10 = 12$
max{11, 12} = 12, tedy $d_0(3) = 2$
$i = 1$, $k = 1$: ${}^1q_1 + {}^1p_{10}v_0 + {}^1p_{11}v_1 = -3 + 0{,}4 \cdot 10 = 1$
$k = 2$: ${}^2q_1 + {}^2p_{10}v_0 + {}^2p_{11}v_1 = -5 + 0{,}7 \cdot 10 = 2$
max{1, 2} = 2, tedy $d_1(3) = 2$
Výsledek 3. iterace dává vektor strategií (2, 2).

Protože ve dvou po sobě jdoucích iteracích jsme dostali stejný vektor strategií, výpočet končí.
Interpretace: Kromě počátečního kroku přinese větší zisk ta strategie, která zahrnuje náklady na technický rozvoj a reklamu výrobku.