

## 12. Řízené markovské řetězce

### 12.1. Definice: Definice optimální strategie

Nechť v každém kroku je možno ergodický homogenní markovský řetězec s konečnou množinou stavů  $J = \{0, 1, \dots, N\}$  definovat  $h$  různými maticemi přechodu  ${}^1\mathbf{P} = ({}^1p_{ij})$ , ...,  ${}^h\mathbf{P} = ({}^hp_{ij})$  a  $h$  různými maticemi výnosů  ${}^1\mathbf{R} = ({}^1r_{ij})$ , ...,  ${}^h\mathbf{R} = ({}^hr_{ij})$ . V každém kroku můžeme vybrat jednu matici přechodu a odpovídající matici výnosů. Možnosti 1, 2, ...,  $h$  se nazývají strategie. Označme  $d_i(n)$  strategii vybranou v  $n$ -tém kroku ve stavu  $i$ . Ta strategie  $d_i^*(n)$ , pro kterou dosahuje střední hodnota celkového výnosu svého maxima, se nazývá **optimální strategie**.

### 12.2. Věta: Věta o optimální strategii

Předpokládejme, že v krocích  $n-1, n-2, \dots, 1$  byla nalezena optimální strategie. Označme  $v_i(n)$  maximální střední hodnotu celkového výnosu v  $n$ -tém kroku ve stavu  $i$ , tj. 
$$v_i(n) = \max_{1 \leq k \leq h} \left\{ q_i^k + \sum_{j \in J} p_{ij}^k v_j(n-1) \right\}.$$
 Je-li maxima dosaženo pro  $k = k^*$ , pak optimální strategie v  $n$ -tém kroku ve stavu  $i$  je  $d_i^*(n) = k^*$ .

**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:

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

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

**Řešení:** Připomeňme, že střední hodnota výnosu při jednom přechodu ze stavu  $i$  se vypočte podle vzorce  $q_i = \sum_{j \in J} p_{ij} r_{ij}$ .

A dále, maximální střední hodnota výnosu v  $n$ -tém kroku ve stavu  $i$  je  $v_i(n) = \max_{1 \leq k \leq h} \left\{ k q_i + \sum_{j \in J}^k p_{ij} v_j(n-1) \right\}$ . Je-li maxima

dosaženo pro  $k = k^*$ , pak optimální strategie v  $n$ -tém kroku ve stavu  $i$  je právě  $k^*$ .

Přitom

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

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

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

$${}^1\mathbf{q} = \begin{pmatrix} 1 \\ -3,8 \end{pmatrix}, {}^2\mathbf{q} = \begin{pmatrix} 1,8 \\ -0,4 \end{pmatrix}$$

$$v_0(1) = \max\{1; 1,8\} = 1,8 \Rightarrow d_0^*(1) = 2, v_1(1) = \max\{-3,8; -0,4\} = -0,4 \Rightarrow d_1^*(1) = 2$$

$$v_0(2) = \max\{{}^1q_0 + {}^1p_{00}v_0(1) + {}^1p_{01}v_1(1); {}^2q_0 + {}^2p_{00}v_0(1) + {}^2p_{01}v_1(1)\} =$$

$$= \max\{1 + 0,5 \cdot 1,8 + 0,5 \cdot (-0,4); 1,8 + 0,4 \cdot 1,8 + 0,6 \cdot (-0,4)\} = \max\{1,7; 2,28\} = 2,28 \Rightarrow d_0^*(2) = 2$$

$$v_1(2) = \max\{{}^1q_1 + {}^1p_{10}v_0(1) + {}^1p_{11}v_1(1); {}^2q_1 + {}^2p_{10}v_0(1) + {}^2p_{11}v_1(1)\} =$$

$$= \max\{-3,8 + 0,2 \cdot 1,8 + 0,8 \cdot (-0,4); -0,4 + 0,4 \cdot 1,8 + 0,6 \cdot (-0,4)\} = \max\{-3,76; 0,08\} = 0,08 \Rightarrow d_1^*(2) = 2$$

$$v_0(3) = \max\{{}^1q_0 + {}^1p_{00}v_0(2) + {}^1p_{01}v_1(2); {}^2q_0 + {}^2p_{00}v_0(2) + {}^2p_{01}v_1(2)\} =$$

$$= \max\{1 + 0,5 \cdot 2,28 + 0,5 \cdot 0,08; 1,8 + 0,4 \cdot 2,28 + 0,6 \cdot 0,08\} = \max\{2,18; 2,76\} = 2,76 \Rightarrow d_0^*(3) = 2$$

$$v_1(3) = \max\{{}^1q_1 + {}^1p_{10}v_0(2) + {}^1p_{11}v_1(2); {}^2q_1 + {}^2p_{10}v_0(2) + {}^2p_{11}v_1(2)\} =$$

$$= \max\{-3,8 + 0,2 \cdot 2,28 + 0,8 \cdot 0,08; -0,4 + 0,4 \cdot 2,28 + 0,6 \cdot 0,08\} = \max\{-3,408; 0,56\} = 0,56 \Rightarrow d_1^*(3) = 2$$

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

**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

Ve větě 11.5. bylo dokázáno, že vytvořující funkce posloupnosti vektorů  $\{v(n)\}_{n=0}^{\infty}$  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 (tj. limitní matici  $\mathbf{A} = \begin{pmatrix} \mathbf{a} \\ \vdots \\ \mathbf{a} \end{pmatrix}$ , kde  $\mathbf{a}$  je

stacionární vektor matice  $\mathbf{P}$ , zde budeme matici  $\mathbf{A}$  značit  $\mathbf{S}$ ) a tranzientní složku  $\mathbf{T}$ . Lze tedy psát

$$G_v(z) = \frac{z}{(1-z)^2}\mathbf{S}\mathbf{q} + \frac{z}{(1-z)\left(1-\frac{z}{c_1}\right)\dots\left(1-\frac{z}{c_k}\right)}\mathbf{T}\mathbf{q} = \frac{z}{(1-z)^2}\mathbf{S}\mathbf{q} + \left(\frac{\mathbf{A}}{1-z} + \frac{\mathbf{B}_1}{1-\frac{z}{c_1}} + \dots + \frac{\mathbf{B}_k}{1-\frac{z}{c_k}}\right)\mathbf{T}\mathbf{q}, \text{ kde } c_i > 1, i = 1, \dots, k.$$

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

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

$\mathbf{B}_i\frac{1}{1-\frac{z}{c_i}}$  je vytvořující funkce posloupnosti  $a_n = \mathbf{B}_i\left(\frac{1}{c_i}\right)^n$ . 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:

$\mathbf{v}(n) = n\mathbf{S}\mathbf{q} + \mathbf{A}\mathbf{T}\mathbf{q}$ . Označíme-li  $\mathbf{S}\mathbf{q} = \mathbf{g}$  a  $\mathbf{A}\mathbf{T}\mathbf{q} = \mathbf{v}$ , lze psát pro dostatečně velká  $n$ :

$\mathbf{v}(n) = n\mathbf{g} + \mathbf{v}$  neboli  $v_i(n) = ng + v_i, i = 0, 1, \dots, N$ . Dosadíme-li toto vyjádření do rekurentního vztahu

$$v_i(n) = q_i + \sum_{j=0}^N p_{ij}v_j(n-1), \text{ dostaneme}$$

$ng + v_i = q_i + \sum_{j=0}^N p_{ij}[(n-1)g + v_j]$ . 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  $g, v_0, v_1, \dots, v_N$ . Protože neznámých je o jednu více než rovnic, nelze určit skutečné hodnoty neznámých  $v_0, v_1, \dots, v_N$ , kterým se říká váhy. Howard navrhl položit  $v_N = 0$ . Tím se počet neznámých sníží a lze vypočítat relativní váhy  $v_0, v_1, \dots, v_{N-1}$ , 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 vážně 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  ${}^1\mathbf{q} = ({}^1q_0, {}^1q_1)^T$  a  ${}^2\mathbf{q} = ({}^2q_0, {}^2q_1)^T$ .

$$\text{Přitom } {}^1\mathbf{P} = \begin{pmatrix} 0,5 & 0,5 \\ 0,4 & 0,6 \end{pmatrix}, {}^1\mathbf{R} = \begin{pmatrix} 9 & 3 \\ 3 & -7 \end{pmatrix}, {}^2\mathbf{P} = \begin{pmatrix} 0,8 & 0,2 \\ 0,7 & 0,3 \end{pmatrix}, {}^2\mathbf{R} = \begin{pmatrix} 4 & 4 \\ 1 & -19 \end{pmatrix}$$

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

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

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

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

1. iterace: zvolíme  $d_0(1) = 1$ ,  $d_1(1) = 1$  a pro  $i = 0, 1$  vyřešíme systém rovnic  $g + v_i = q_i + \sum_{j=0}^N p_{ij} v_j$  (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:

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

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

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

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

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

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

Výsledek 2. iterace dává vektor strategií  $(2, 2)^T$ .



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  $(2, 2)^T$ .

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.

**12.7. Poznámka:** Pro libovolný počet stavů a strategií lze použít funkce `howardn.m` (autor Jakub Buček, 2012):

```
function D = howardn(P,R)
```

```
% Algoritmus Howardova iteracního postupu pro libovolný počet stavů a strategií
```

```
% Vstupní parametry:
```

```
% P...matice přechodu všech strategií naskladané do jedné matice
```

```
% R...matice výnosů všech strategií naskladané do jedné matice
```

```
% Máme-li jednotlivé matice  $P_1, P_2, \dots, P_k$  a  $R_1, R_2, \dots, R_k$ , doporučuje se
```

```
% uvádět funkci ve tvaru: howardn([P1,P2,...,Pk],[R1,R2,...,Rk])
```

```
% Výstup:
```

```
% D...matice obsahující optimální strategii pro jednotlivé období,
```

```
% jednotlivá období jsou označena číslem řádku
```