Kybernetika

Řízené Markovovy řetězce

PETR MANDL

ACADEMIA
PRAHA

ÚVOD

Stať obsahuje výklad metod dynamického programování v Markovových řetězcích. U čtenáře se předpokládá asi taková znalost těchto řetězců, jakou je možno získat studiem knihy J. G. Kemeny, J. L. Snell: Finite Markov chains. Ten, kdo má definice teorie Markovových řetězců dostatečně promyšleny a má jistou zběhlost v matematickém způsobu uvažování, mohl by vystačiti s informací, obsaženou v odstavci 2 této práce. Tam jsou základní pojmy stručně zopakovány a dokázány věty, na něž se další výklad odvolává.

Definici řízeného řetězce v odstavci 5 předchází odstavec o úloze nalezení nejlepší cesty. Současně jsou zavedeny různé typy výnosu ze řízeného řetězce: očekávaný výnos do předem zvolené doby, očekávaný diskontovaný výnos, očekávaný výnos do dosažení zvolené množiny, průměrný výnos na jednotku času. Další odstavce postupně pojednávají o optimalizačních metodách při jednotlivých typech výnosu. Řízení Markovových procesů s konečným počtem stavů je věnován odstavec 11, řízení větvicích procesů odstavec 12. Všude s výjimkou posledního odstavce se požaduje, aby při řízení byla k dispozici úplná informace o stavu soustavy. Řízení řetězce s neúplnou informací se převede na řízení řetězce, jehož stavem je aposteriorní rozložení pravděpodobností.

Lineární programování a metody vzniklé jeho zobecněním lze při řízení Markovových řetězců, zejména při maximalizaci průměrného výnosu na jednotku času, rovněž použít. Tato práce vychází z jednotného hlediska dynamického programování a o lineárním programování nepojednává.

Seznam literatury je umístěn na konci statě a nejsou na něj odvolání v textu. V úvodu je vhodné uvésti dvě základní knihy:

R. Bellman: Dynamic programming. Princeton 1957,

R. A. Howard: Dynamic programming and Markov processes. New York, London 1960.

Při odkazech v témže odstavci jsou formule označovány pořadovými čísly, při odkazech na formule jiného odstavce též jeho číslem. Např. (13 § 4) značí formuli 13 odstavce 4. Na konci důkazu je umístěn symbol □.

Při psaní práce jsem vyšel ze zápisu části přednášky o Markovových procesech a jejich řízení, kterou jsem měl v zimním semestru 1966–67 na universitě v Heidelbergu.

1. MARKOVSKÉ MODELY

Uveďme několik všeobecných poznámek o markovských modelech. Přitom použijeme následujícího příkladu: V práci Formulae for Projecting Enrolments and Degrees Awarded in Universities [Journal Royal Stat. Soc. Ser. A, 126 (1963), 400–409] se J. Gani zabývá závislostí mezi počtem studentů v jednotlivých ročnících australských universit a počtem studentů, kteří se v předchozích letech zapsali do prvého ročníku. Předpokládá, že poměrné počty studentů, kteří v ročníku zůstanou, postoupí dále nebo ze studia vystoupí (resp. studium ukončí), se velmi málo mění v závislosti na čase.

Pro jednoduchost označení si budeme všímat pouze prvních tří ročníků universitního studia. Budiž

N^n – počet studentů, kteří se v roce n nově zapsali do prvého ročníku,
q_{ii} – poměrná část studentů, kteří zůstali v i-tém ročníku, $i = 1, 2, 3$,
q_{ii+1} – poměrná část studentů, kteří postoupili z i-tého ročníku do dalšího, $i = 1, 2$,
q_{i0} – poměrná část studentů i-tého ročníku, kteří opustili studium (resp. dosáhli hodnosti bakaláře).

O počtu studentů v i-tém ročníku v roce n, S_i^n, platí tedy podle modelu

$$\begin{aligned} S_1^n &= N^n + q_{11}S_1^{n-1}, \\ S_2^n &= q_{12}S_1^{n-1} + q_{22}S_2^{n-1}, \\ S_3^n &= q_{23}S_2^{n-1} + q_{33}S_3^{n-1}. \end{aligned} \tag{1}$$

Srovnání se statistikou australských universit ukázalo, že tyto rovnice dobře vystihují skutečný stav.

Rovnice (1) představují model počtu studentů na universitách jako celku, řekli bychom model makroskopický. Obvykleji je takový model označován jako *deterministický*, protože v něm pojem pravděpodobnosti nevystupuje. Nejjednodušší předpoklad o detailní struktuře uvažovaného jevu (*pravděpodobnostní, stochastický model*), který by současně byl v souladu s rovnicemi (1) je tento: Nechť každý student v i-tém ročníku má pravděpodobnost q_{ii} ročník opakovat, q_{ii+1} postoupit do dalšího ročníku. Přitom tato pravděpodobnost nezávisí na způsobu (zda s opakováním či bez opakování nějakého ročníku) jakým se student do i-tého ročníku dostal a na postupu ostatních studentů. Předpokládá se tedy, že studenti na universitách tvoří soubor, jehož členové se vyvíjejí podle Markovova řetězce s maticí pravděpodobností přechodu

$$\boldsymbol{P} = \|p_{ij}\|_{i,j=0}^3 = \begin{pmatrix} 1, & 0, & 0, & 0 \\ q_{10}, & q_{11}, & q_{12}, & 0 \\ q_{20}, & 0, & q_{22}, & q_{23} \\ q_{30}, & 0, & 0, & q_{33} \end{pmatrix}.$$

Ve stavu 0 se nacházejí studenti vystoupivší ze studia.

Je ihned patrné, že předpoklad o homogenitě souboru a o tom, že se jeho členové vyvíjejí nezávisle na sobě, málo souhlasí se skutečností. Na druhé straně o tomto stochastickém modelu vzhledem k velkému rozsahu souboru platí, že celkové počty v jednotlivých ročnících se velmi málo liší od očekávaných a splňují proto relace (1) dostatečně přesně. Nedopustíme se tedy velké chyby, budeme-li předpokládat, že empirické hodnoty veličin S_i^n, $i = 1, 2, 3$, jsou výsledkem pozorování právě popsaného souboru.

Ta vlastnost, že stochastický model není v rozporu s rovnicemi (1), jejichž platnost byla ověřena srovnáním se skutečností, nikterak neodůvodňuje nahrazení jednoduchého deterministického modelu složitějším modelem pravděpodobnostním. Potřeba pravděpodobnostního modelu se projeví až tehdy, máme-li co činit nikoliv se souborem jako celkem, ale s jednotlivými jeho členy, kteří byli vybráni způsobem dostatečně blízkým nezávislému náhodnému výběru se stejnými pravděpodobnostmi. Přitom výběr může být proveden nějakými vnějšími okolnostmi, nikoliv pouze námi.

Vhodnost pravděpodobnostních modelů byla ověřena např. mnohaletou zkušeností životních pojišťoven. Pozorováním nějakého dostatečně rozsáhlého souboru získá pojišťovna tabulku úmrtnosti, řekněme mužů v Československu. Tabulku tvoří hodnoty $l_1, l_2, \ldots, l_x, \ldots, l_\omega$, kde l_x značí počet mužů z tisíce živě narozených, kteří se dožili x let. Při stanovení pojistného pojišťovna předpokládá, že x-letý pojištěnec má pravděpodobnost $q_x = (l_x - l_{x+1})/l_x$ zemřít do jednoho roku. Obdobně jako v předchozím modelu se tedy připouští, že soubor pojištěnců je tvořen jednotlivci, vyvíjejícími se podle stejných pravděpodobnostních zákonů.

Tato stať je z největší části věnována Markovovým řetězcům, tedy pravděpodobnostním modelům. V odstavci 12 se k deterministickým modelům vrátíme. Tam bude také poukázáno na interpretaci vyložené teorie řízení v deterministických modelech.

2. KONEČNÉ MARKOVOVY ŘETĚZCE

Markovovy řetězce jsou matematické modely soustav (fyzikálních, biologických, ekonomických atd.) měnících svůj stav nebo pozorovaných v dané posloupnosti časových okamžiků a takových, že okamžitý stav soustavy plně určuje pravděpodobnostní vlastnosti jejího dalšího vývoje. Je-li tato závisiost budoucnosti soustavy na okamžitém stavu neměnná s časem, hovoříme o *homogenních* řetězcích.

Vývoj soustavy je popsán posloupností náhodných veličin $\{X_n, n = 0, 1, 2, \ldots\}$. X_n značí stav soustavy (řetězce) v čase n. Budeme se zabývat řetězci s konečným počtem stavů. Budiž $I = \{1, 2, \ldots, r\}$ množina stavů řetězce. Markovská vlastnost znamená, že platí

$$(1) \quad P(X_{n+1} = j_1, X_{n+2} = j_2, \ldots, X_{n+m} = j_m \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \\ = P(X_{n+1} = j_1, X_{n+2} = j_2, \ldots, X_{n+m} = j_m \mid X_n = i_n)$$

pro libovolná n, m celá a $j_1, \ldots, j_m, i_0, \ldots, i_n \in I$. Rozložení pravděpodobností řetězce je tedy určeno počátečním stavem (hodnotou X_0) a *pravděpodobnostmi přechodu*

$$P(X_{n+1} = j \mid X_n = i), \quad i, j \in I, \quad n = 0, 1, \ldots$$

Pravděpodobnost a střední hodnotu za podmínky, že počáteční stav soustavy je j, budeme označovat

$$P_j(\cdot) = P(\cdot \mid X_0 = j), \quad E_j \cdot = E\{\cdot \mid X_0 = j\}.$$

Pravděpodobnosti přechodu homogenních řetězců nezávisejí na n, tj.

$$P(X_{n+1} = j \mid X_n = i) = p_{ij}, \quad i, j \in I, \quad n = 0, 1, \ldots$$

Matice veličin p_{ij},

$$\mathbf{P} = \|p_{ij}\|_{i,j=1}^{r},$$

se nazývá maticí pravděpodobností přechodu řetězce. Označme

$$\mathbf{P}^m = \|p_{ij}^{(m)}\|_{i,j=1}^{r}.$$

$p_{ij}^{(m)}$ jsou pravděpodobnosti přechodu za m kroků, tj.

$$p_{ij}^{(m)} = P(X_{n+m} = j \mid X_n = i).$$

$\mathbf{P}^0$ je jednotková matice.

Ve zbylé části tohoto odstavce shrneme některé v dalším potřebné vlastnosti homogenních řetězců.

Věta 1. *Nechť o matici* $\mathbf{P}$ *platí: Lze nalézt* $j_0 \in I$, $d > 0$ *a celé kladné* ν *tak, že*

$$p_{ij_0}^{(\nu)} \geqq d \quad pro \quad i \in I.$$

Potom existují limitní pravděpodobnosti $\pi_1, \ldots, \pi_r$ *tak, že je splněna nerovnost*

$$|p_{ij}^{(n)} - \pi_j| \leqq (1 - d)^{n/\nu - 1} \quad pro \quad i, j \in I, \quad n = 0, 1, \ldots$$

Důkaz. Označme

$$M_j^{(n)} = \max_i p_{ij}^{(n)}, \quad m_j^{(n)} = \min_i p_{ij}^{(n)},$$

maximum a minimum prvků j-tého sloupce matice $\mathbf{P}^n$. Platí

$$m_j^{(n)} \leqq p_{ij}^{(n+1)} = \sum_k p_{ik} p_{kj}^{(n)} \leqq M_j^{(n)}.$$

Odkud

(2) $$m_j^{(n)} \leqq m_j^{(n+1)}, \quad M_j^{(n)} \geqq M_j^{(n+1)}.$$

Ze vztahu

$$p_{ij}^{((m+1)\nu)} = dp_{j_0 j}^{(m\nu)} + (p_{ij_0}^{(\nu)} - d)\, p_{j_0 j}^{(m\nu)} + \sum_{k \neq j_0} p_{ik}^{(\nu)} p_{kj}^{(m\nu)}$$

pro $i, j \in I$, $m = 0, 1, 2, \ldots$, plyne

$$dp_{j_0 j}^{(m\nu)} + (1 - d)\, m_j^{(m\nu)} \leqq p_{ij}^{((m+1)\nu)} \leqq dp_{j_0 j}^{(m\nu)} + (1 - d)\, M_j^{(m\nu)} .$$

Odtud snadno vidíme, že je

$$M_j^{((m+1)\nu)} - m_j^{((m+1)\nu)} \leqq (1 - d)\,(M_j^{(m\nu)} - m_j^{(m\nu)})$$

a tedy

(3) $$M_j^{(m\nu)} - m_j^{(m\nu)} \leqq (1 - d)^m ,$$

neboť

$$M_j^{(0)} - m_j^{(0)} = 1 .$$

Z (2) a (3) dostáváme

$$\lim_{n\to\infty} M_j^{(n)} = \lim_{n\to\infty} m_j^{(n)} = \pi_j .$$

Mějme nyní libovolné přirozené n a budiž $m = [n/\nu]$, ([ ] značí celou část). Potom

$$|p_{ij}^{(n)} - \pi_j| \leqq M_j^{(n)} - m_j^{(n)} \leqq M_j^{(m\nu)} - m_j^{(m\nu)} \leqq$$
$$\leqq (1 - d)^m \leqq (1 - d)^{n/\nu - 1} . \quad \square$$

Věta 2. *Pro libovolné* $\mathbf{P}$ *existují limity*

(4) $$\lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} p_{ij}^{(n)} = \pi_{ij} , \quad i, j \in I .$$

Důkaz. Vztah (4) bude ověřen, dokážeme-li, že pro libovolné dvě posloupnosti přirozených čísel

$$N'_1, N'_2, \ldots, N'_m \to \infty , \quad N''_1, N''_2, \ldots, N''_m \to \infty ,$$

pro něž existují

$$\lim_{m\to\infty} (N'_m)^{-1} \sum_{n=1}^{N'_m} p_{ij}^{(n)} = \pi'_{ij} , \quad \lim_{m\to\infty} (N''_m)^{-1} \sum_{n=1}^{N''_m} p_{ij}^{(n)} = \pi''_{ij} ,$$

pro $i, j \in I$, platí

$$\pi'_{ij} = \pi''_{ij} .$$

Máme

$$\sum_k \pi'_{ik} p_{kj} = \sum_k p_{ik} \pi'_{kj} = \lim_{m\to\infty} (N'_m)^{-1} \sum_{m=2}^{N'_m + 1} p_{ij}^{(n)} = \pi'_{ij} .$$

Odtud plyne

$$\sum_k \pi'_{ik} p_{kj}^{(n)} = \pi'_{ij} = \sum_k p_{ik}^{(n)} \pi'_{kj}$$

pro $n = 1, 2, \ldots$ a tedy také

$$\sum_k \pi'_{ik} \pi''_{kj} = \lim_{m \to \infty} (N''_m)^{-1} \sum_{n=1}^{N''_m} \sum_k \pi'_{ik} p_{kj}^{(n)} = \pi'_{ij} . \tag{5}$$

Obdobně

$$\sum_k \pi''_{ik} \pi'_{kj} = \pi'_{ij} . \tag{6}$$

Záměnou $\pi'_{..}$ za $\pi''_{..}$ v (5) a ze (6) dostáváme

$$\pi'_{ij} = \sum_k \pi''_{ik} \pi'_{kj} = \pi''_{ij} . \quad \square$$

Řekneme, že stav j je dosažitelný ze stavu i, existuje-li n tak, že platí $p_{ij}^{(n)} > 0$. V tom případě budeme psáti $i \to j$. Je-li současně $i \to j$ a $j \to i$, nazývají se stavy i a j sousledné. Stav i (konečného) řetězce se nazývá *rekurentním*, platí-li $j \to i$ pro všechna j, pro něž $i \to j$. V opačném případě je i *tranzientním* stavem.

Navzájem sousledné rekurentní stavy tvoří tzv. *isolované třídy rekurentních stavů*. Jakmile řetězec přešel do takové třídy, zůstává v ní trvale. Vhodným přečíslováním stavů lze matici $\boldsymbol{P}$ uvést na tvar

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{R}_1, & \boldsymbol{0}, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{R}_2, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{T}_1, & \boldsymbol{T}_2, & \boldsymbol{T}_3, & \ldots, & \boldsymbol{T} \end{pmatrix},$$

kde $\boldsymbol{R}_i$ jsou stochastické matice, utvořené z pravděpodobností přechodu v i-té izolované rekurentní třídě I_i. $\boldsymbol{T}_i$ značí matici pravděpodobností přechodu z tranzientních stavů do I_i. $\boldsymbol{T}$ je matice pravděpodobností přechodu mezi tranzientními stavy, $\boldsymbol{0}$ označuje nulové matice.

Věta 3. *Je-li j tranzientním stavem řetězce $\{X_n, n = 0, 1, \ldots\}$, potom řada $\sum_{n=0}^{\infty} p_{ij}^{(n)}$ konverguje pro $i \in I$.*

Důkaz. Položme

$$1 - f_{ij} = P_i \quad (X_n \neq j, n = 1, 2, \ldots) .$$

Z předpokladu o tranzientnosti stavu j vyplývá, že existuje $k \in I$, pro něž $j \to k$, ale neplatí $k \to j$. Tedy $f_{jj} < 1$.

Budiž $\mathscr{E}_{ij}$ očekávaný počet výskytů soustavy ve stavu j, je-li počátečním stav i. Soustava vycházející ze stavu j se s pravděpodobností f_{jj} opět do tohoto stavu vrátí. Za podmínky, že se soustava do stavu j skutečně vrátila, je pravděpodobnost dalšího návratu opět f_{jj}, protože $\{X_n, n = 0, 1, \ldots\}$ je homogenní Markovův řetězec. Vidíme, že soustava vycházející ze stavu j se vyskytne v tomto stavu aspoň n-krát s pravděpodobností f_{jj}^{n-1}, právě n-krát s pravděpodobností $f_{jj}^{n-1}(1 - f_{jj})$. Tedy

$$(7) \qquad \mathscr{E}_{jj} = \sum_{n=1}^{\infty} n f_{jj}^{n-1}(1 - f_{jj}) = (1 - f_{jj})^{-1}.$$

Dále zřejmě

$$(8) \qquad \mathscr{E}_{ij} = f_{ij}\mathscr{E}_{jj} \leqq \mathscr{E}_{jj}.$$

Nechť

$$(9) \qquad Y_n^{(j)} = 1 \quad \text{když} \quad X_n = j, \quad Y_n^{(j)} = 0 \quad \text{když} \quad X_n \neq j.$$

Pak platí

$$(10) \qquad \mathscr{E}_{ij} = E_i \sum_{n=0}^{\infty} Y_n^{(j)} = \sum_{n=0}^{\infty} E_i Y_n^{(j)} = \sum_{n=0}^{\infty} p_{ij}^{(n)}.$$

Z rovnic (7), (8) a (10) vidíme, že je $\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$. □

Věta 4. *Budiž*

$$N^{(j)} = \inf\{n : n \neq 0, X_n = j, n = 1, 2, \ldots\}$$

doba prvního přechodu řetězce $\{X_n, n = 0, 1, \ldots\}$ *do stavu* j. *Je-li* I *izolovanou třídou rekurentních stavů, pak*

$$E_i N^{(j)} < \infty \quad \text{pro} \quad i, j \in I.$$

Důkaz. Volme pro určitost $j = 1$. Z přechodové matice $\mathbf{P}$ utvořme matici

$$\tilde{\mathbf{P}} = \|\tilde{p}_{ij}\|_{i,j=1}^{r} = \begin{pmatrix} 1, & 0, & \ldots, & 0 \\ p_{21}, & p_{22}, & \ldots, & p_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ p_{r1}, & p_{r2}, & \ldots, & p_{rr} \end{pmatrix}.$$

Snadno vidíme, že $\tilde{\mathbf{P}}$ je přechodovou maticí řetězce $\{\tilde{X}_n, n = 0, 1, \ldots\}$, který vznikne z $\{X_n, n = 0, 1, \ldots\}$ zastavením soustavy ve stavu 1, tj. $\tilde{X}_n \equiv 1$, když $X_0 = 1$, jinak $\tilde{X}_n = X_n$ pro $n \leqq N^{(1)}$, $\tilde{X}_n = 1$ pro $n > N^{(1)}$.

Zřejmě stavy $k \neq 1$ jsou tranzientními stavy řetězce $\{\tilde{X}_n, n = 0, 1, \dots\}$, neboť $k \to 1$, ale nikoliv $1 \to k$. Dále, pro $i \neq 1$,

$$P_i(N^{(1)} > n) = P_i(\tilde{X}_n \neq 1) = \sum_{k \neq 1} \tilde{p}_{ik}^{(n)},$$

$$E_i N^{(1)} = \sum_{n=1}^{\infty} n[P_i(N^{(1)} > n - 1) - P_i(N^{(1)} > n)] =$$

$$= \sum_{n=0}^{\infty} P_i(N^{(1)} > n) = \sum_{k \neq 1} \sum_{n=0}^{\infty} \tilde{p}_{ik}^{(n)}.$$

Řada $\sum_{n=0}^{\infty} \tilde{p}_{ik}^{(n)}$ konverguje dle věty 3. Vidíme, že $E_i N^{(1)} < \infty$ pro $i \neq 1$. Dále

$$E_1 N^{(1)} = 1 + \sum_{k \neq 1} p_{ik} E_k N^{(1)} < \infty. \quad \square$$

Připomeňme *silný zákon velkých čísel*: Budiž $\{Z_k, k = 1, 2, \dots\}$ posloupnost nezávislých stejně rozložených náhodných veličin s konečnou střední hodnotou m. Potom

$$\lim_{k \to \infty} k^{-1} \sum_{i=1}^{k} Z_i = m$$

s pravděpodobností 1.

Věta 5. *Za předpokladů věty* 4

$$\lim_{N \to \infty} N^{-1} \sum_{n=0}^{N-1} p_{ij}^{(n)} = (E_j N^{(j)})^{-1} \quad \textit{pro} \quad i, j \in I.$$

Důkaz. Označme $N_k^{(j)}$ dobu k-tého přechodu do stavu j, tj.

$$N_k^{(j)} = \inf\{n : n > N_{k-1}^{(j)}, X_n = j\},$$

kde $N_0^{(j)} = 0$. Všimněme si této důležité vlastnosti homogenních řetězců: Pro libovolná $i, j, i_1, \dots, i_s \in I$, $k = 1, 2, \dots$ platí

$$\begin{aligned} & P_i(X_{N_k(j)+1} = i_1, X_{N_k(j)+2} = i_2, \dots, X_{N_k(j)+s} = i_s | \\ (11) \qquad & X_1 = j_1, X_2 = j_2, \dots, X_{N_k(j)-1} = j_{N_k(j)-1}) = \\ & = P_j(X_1 = i_1, X_2 = i_2, \dots, X_s = i_s). \end{aligned}$$

Vztah (11) říká, že vývoj soustavy od okamžiku $N_k^{(j)}$ je stejný jako vývoj soustavy, která byla ve stavu j na počátku. (11) je důsledkem (1), časové homogenity procesu a skutečnosti, že hodnota $N_k^{(j)}$ je nezávislá na vývoji řetězce po době $N_k^{(j)}$.

Z (11) vyplývá, že $\{N_{k+1}^{(j)} - N_k^{(j)},\ k = 1, 2, \ldots\}$ tvoří posloupnost nezávislých stejně rozložených náhodných veličin se střední hodnotou $E_j N^{(j)}$. Podle zákona velkých čísel

$$\lim_{k\to\infty} \frac{N_k^{(j)}}{k} = \lim_{k\to\infty} \frac{\sum_{m=0}^{k-1} N_{m+1}^{(j)} - N_m^{(j)}}{k} = E_j N^{(j)}$$

s pravděpodobností 1. Definujme opět veličiny $Y_n^{(j)}$ vztahem (9). Snadno se nahlédne, že je s pravděpodobností 1

$$(12) \qquad (E_j N^{(j)})^{-1} = \lim_{k\to\infty} \frac{k}{N_k^{(j)}} = \lim_{k\to\infty} \frac{\sum_{n=1}^{N_k(j)} Y_n^{(j)}}{N_k^{(j)}} = \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} Y_n^{(j)} .$$

Přejdeme-li ve (12) ke střední hodnotě, dostáváme

$$(E_j N^{(j)})^{-1} = \lim_{N\to\infty} N^{-1} E_i \sum_{n=0}^{N-1} Y_n^{(j)} = \lim_{N\to\infty} N^{-1} \sum_{n=0}^{N-1} p_{ij}^{(n)} . \quad \square$$

3. MARKOVOVY PROCESY S KONEČNÝM POČTEM STAVŮ

O *Markovových procesech* na rozdíl od řetězců hovoříme, mění-li se parametr času, v němž vývoj soustavy probíhá spojitě. Homogennímu Markovovu procesu $\{X_t, t \geqq 0\}$ s konečnou množinou stavů $I = \{1, \ldots, r\}$ odpovídá vektor $(\mu_1, \ldots, \mu_r)$ intenzit výstupu a stochastická matice

$$\mathbf{P} = \begin{pmatrix} 0, & p_{12}, & \ldots, & p_{1r} \\ p_{21}, & 0, & \ldots, & p_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ p_{r1}, & p_{r2}, & \ldots, & 0 \end{pmatrix}$$

pravděpodobností přechodu v okamžiku výstupu. Je-li soustava ve stavu i v čase t, potom pravděpodobnost změny jejího stavu v infinitezimálním časovém intervalu $(t, t + \mathrm{d}t)$ je rovna $\mu_i \, \mathrm{d}t$. Za podmínky, že přeskok nastal, přejde soustava do stavu j s pravděpodobností p_{ij}. Tyto pravděpodobnosti jsou nezávislé na vývoji soustavy do doby t.

Posloupnost stavů, jak za sebou následují bez ohledu na čas, tvoří Markovův řetězec s přechodovou maticí $\mathbf{P}$. V jednotlivých stavech setrvává soustava náhodnou dobu. Je-li posloupnost stavů známa, jsou tyto doby setrvání nezávislé náhodné veličiny. Budiž $F_i(t)$ distribuční funkce doby setrvání ve stavu i. Platí

$$(1) \qquad F_i(t + \mathrm{d}t) = F_i(t) + (1 - F_i(t))\, \mu_i \, \mathrm{d}t\,, \quad F(0) = 0\,.$$

Z (1) plyne snadno, že rozložení doby setrvání ve stavu i je exponenciální se střední hodnotou μ_i^{-1}, tj.

$$F_i(t) = 1 - e^{-\mu_i t}, \quad i \in I .$$

Utvořme matici

$$\boldsymbol{M} = \begin{pmatrix} \mu_{11}, \mu_{12}, \ldots, \mu_{1r} \\ \mu_{21}, \mu_{22}, \ldots, \mu_{2r} \\ \ldots\ldots\ldots\ldots\ldots \\ \mu_{r1}, \mu_{r2}, \ldots, \mu_{rr} \end{pmatrix} = \begin{pmatrix} -\mu_1, & \mu_1 p_{12}, \ldots, & \mu_1 p_{1r} \\ \mu_2 p_{21}, & -\mu_2, \quad \ldots, & \mu_2 p_{2r} \\ \ldots\ldots & \ldots\ldots\ldots & \ldots\ldots \\ \mu_r p_{r1}, & \mu_r p_{r2}, \ldots, & -\mu_r \end{pmatrix}.$$

Matice $\boldsymbol{M}$ se nazývá *maticí intenzit přechodu* procesu $\{X_t, t \geqq 0\}$. Je-li soustava ve stavu i v čase t, potom pravděpodobnost přechodu do stavu j v časovém intervalu $(t, t + dt)$ je rovna $\mu_{ij}\, dt$.

Buďtež

$$P_{ij}^{(t)} = P(X_{s+t} = j \mid X_s = i), \quad i, j \in I,$$

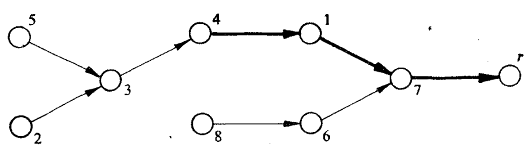
pravděpodobnosti přechodu procesu. Je splněna soustava Kolmogorovových diferenciálních rovnic

$$\frac{d}{dt} P_{ij}^{(t)} = \sum_k P_{ik}^{(t)} \mu_{kj} = \sum_k \mu_{ik} P_{kj}^{(t)}, \quad i, j \in I .$$

Nehomogenní Markovův proces má matici intenzit přechodu závislou na časovém parametru t.

4. ÚLOHA NALEZENÍ NEJLEPŠÍ CESTY

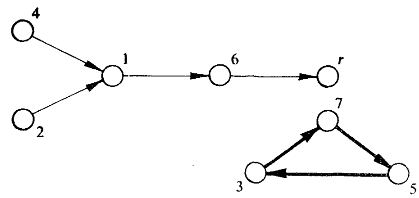
Nenáhodným případem úloh, kterým budou věnovány další odstavce, je následující problém: Předpokládejme, že máme r míst, označených čísly $\{1, 2, \ldots, r\} = I$ a kladná čísla $c(i, j)$, kde $i \neq j$, $i = 1, 2, \ldots, r - 1$, $j \in I$. Čísla $c(i, j)$ nechť představují náklad na pohyb z místa i do místa j. Naším cílem budiž dosažení místa r s nejmenším nákladem.



Obr. 1.

Ve shodě s terminologií použitou dále budeme řízením rozumět funkci $z(i)$ definovanou pro $i = 1, 2, \ldots, r - 1$ a nabývající hodnot z I. $z(i)$ udává místo, do něhož postupujeme z místa i. Funkci $z(i)$ budeme zkráceně označovati ω. Řízení ω odpovídá orientovaný graf Γ s množinou uzlů I a hranami $\overrightarrow{i\, z(i)}$, $i = 1, 2, \ldots, r - 1$

(obr. 1). Řízení nazveme přípustným, když z každého z uzlů $1, 2, \ldots, r-1$ vede orientované spojení do uzlu r. To nastává právě když Γ neobsahuje žádný cyklus (obr. 2). Přitom orientované spojení z libovolného uzlu do r je jediné.



Obr. 2.

Budiž ω přípustné řízení. Definujme $u(i; \omega)$, $i \in I$, rovnicemi

$$(1) \qquad \begin{aligned} u(i; \omega) &= c(i, z(i)) + u(z(i); \omega)\,, \quad i = 1, \ldots, r-1\,, \\ u(r; \omega) &= 0\,. \end{aligned}$$

Z následující věty vyplývá, že soustava (1) má jednoznačné řešení. Z jejího důkazu je pak patrné, že $u(i; \omega)$ představuje náklad na dosažení r z místa i při řízení ω. Poznamenejme, že vzhledem k pozitivnosti čísel $c(i, j)$ přípustnost řízení ω je nutná k řešitelnosti (1).

Věta 1. *Nechť ω je přípustné řízení, γ_j, $j \in I$, nechť jsou libovolná čísla. Potom soustava rovnic*

$$(2) \qquad \begin{aligned} u(i) &= \gamma_i + u(z(i)),\ i = 1, 2, \ldots, r-1\,, \\ u(r) &= \gamma_r\,, \end{aligned}$$

má jediné řešení. Je-li $\gamma_i \geqq 0$, $i \in I$, je $u(i) \geqq 0$, $i \in I$.

Důkaz. Uvažujme graf Γ odpovídající řízení ω. Mějme libovolné k. Jelikož ω je přípustné řízení, existuje v Γ orientované spojení $k = k_0$, $\overrightarrow{k_0 k_1}$, $\overrightarrow{k_1 k_2}$, $\ldots$, $\overrightarrow{k_{m-1} k_m}$, $k_m = r$. Takové spojení je jediné. Vidíme, že můžeme postupně klást

$$(3) \qquad \begin{aligned} u(k_m) &= \gamma_r\,, \\ u(k_{j-1}) &= \gamma_{k_{j-1}} + u(k_j)\,, \quad j = m, m-1, \ldots, 1\,. \end{aligned}$$

Volíme-li $k = 1, 2, \ldots, r-1$, dostáváme řešení soustavy (2). Z (3) vyplývá, že toto řešení je jednoznačné a nezáporné, kdykoliv $\gamma_i \geqq 0$, $i \in I$. $\square$

Základem iterační metody k nacházení optimálního řízení bude postupné zlepšování řízení způsobem popsaným ve větě 2.

Věta 2. *Nechť $\omega' \sim z'(i)$ splňuje vztah*

$$c(i, z'(i)) + u(z'(i); \omega) = \min_{j \neq i} \{c(i, j) + u(j; \omega)\} \tag{4}$$

pro $i = 1, \ldots, r - 1$. Potom ω' je přípustné řízení a platí

$$u(i; \omega') \leqq u(i; \omega) \quad pro \quad i \in I\,. \tag{5}$$

Přitom

$$u(i; \omega') = u(i; \omega) \quad pro \quad i \in I \tag{6}$$

právě když

$$c(i, z(i)) + u(z(i); \omega) = \min_{j \neq i} [c(i, j) + u(j; \omega)] \tag{7}$$

pro $i = 1, \ldots, r - 1$.

Důkaz. Položme

$$h(i; \omega') = c(i, z'(i)) + u(z'(i); \omega)\,. \tag{8}$$

Vzhledem k (1) a (4) platí

$$h(i; \omega') \leqq u(i; \omega)\,. \tag{9}$$

Z rovnosti (8) vyplývá, že $h(i; \omega')$ je náklad na dosažení r v případě, že v prvním kroku se pohybujeme dle řízení ω' a potom dle řízení ω. Předpokládejme, že ω' není přípustným řízením a tedy jeho graf Γ' obsahuje cyklus s uzly $k_0, k_1, \ldots, k_m = k_0$. Budiž

$$u(k_s; \omega) = \min_{j=0,\ldots,m-1} u(k_j; \omega)\,.$$

Potom

$$h(k_s; \omega') = c(k_s, k_{s+1}) + u(k_{s+1}; \omega) > u(k_s; \omega)\,.$$

To je spor s (9). Vidíme, že řízení ω' musí být přípustným.

Označme nyní

$$\bar{u}(i) = u(i; \omega) - u(i; \omega')\,.$$

Dle (1) a (4)

$$u(i; \omega) = \gamma_i + c(i, z'(i)) + u(z'(i); \omega)\,, \tag{10}$$

kde $\gamma_i \geqq 0$. Dále

$$u(i; \omega') = c(i, z'(i)) + u(z'(i); \omega')\,. \tag{11}$$

Odečtením (11) od (10) dostáváme

$$(12)\qquad \bar{u}(i) = \gamma_i + \bar{u}(z'(i))\,, \quad i = 1, \ldots, r-1\,,$$
$$\bar{u}(r) = 0\,.$$

Z věty 1 vyplývá $\bar{u}(i) \geqq 0$ pro $i \in I$, což je totéž jako (5).

Všimněme si nyní posledního tvrzení věty. Nechť neplatí (7). Potom ze (4) plyne, že aspoň pro jedno i je $\gamma_i > 0$ v rovnosti (10). Tedy, dle (12), $\bar{u}(i) > 0$. Vidíme, že také (6) neplatí. Platí-li (7), potom z následující věty plyne

$$u(i;\omega') \geqq u(i;\omega) \quad \text{pro} \quad i \in I\,,$$

což společně s (5) dává (6). □

Věta 3. *Nechť přípustné řízení $\hat{\omega} \sim \hat{z}(i)$ splňuje*

$$(13)\qquad c(i, \hat{z}(i)) + u(\hat{z}(i); \hat{\omega}) = \min_{j \neq i} \left[c(i, j) + u(j; \hat{\omega})\right],$$

pro $i = 1, \ldots, r-1$. Potom $\hat{\omega}$ je optimální řízení, tj.

$$u(i;\hat{\omega}) \leqq u(i;\omega)\,, \quad i \in I\,,$$

pro každé přípustné řízení ω.

Důkaz. Nechť $\hat{\omega}$ splňuje předpoklady věty. Mějme libovolné přípustné řízení ω. Ze (13) vyplývá

$$(14)\qquad u(i;\hat{\omega}) = c(i, \hat{z}(i)) + u(\hat{z}(i);\hat{\omega}) =$$
$$= \gamma_i + c(i, z(i)) + u(z(i);\hat{\omega})\,,$$

kde $\gamma_i \leqq 0$. Odečteme-li (1) od (14), dostáváme

$$u(i;\hat{\omega}) - u(i;\omega) = \gamma_i + u(z(i);\hat{\omega}) - u(z(i);\omega)\,.$$

Odtud a z věty 1

$$u(i;\hat{\omega}) - u(i;\omega) \leqq 0\,, \quad \text{tj.} \quad u(i,\hat{\omega}) \leqq u(i;\omega)$$

pro $i \in I$, neboť γ_i jsou nekladná. □

Iterační postup nacházení optimálního řízení. Vyjdeme z libovolného přípustného řízení $\omega^0 \sim z^0(i)$. Řízení $\omega^{m+1} \sim z^{m+1}(i)$ konstruujeme na základě ω^m takto:

1. Určíme $u(i;\omega^m)$ z rovnic

$$(15)\qquad u(i;\omega^m) = c(i, z^m(i)) + u(z^m(i);\omega^m)\,, \quad i = 1, \ldots, r-1\,,$$
$$u(r;\omega^m) = 0\,.$$

2. $z^{m+1}(i)$ volíme pro $i = 1, \ldots, r - 1$ tak, aby platilo

$$c(i, z^{m+1}(i)) + u(z^{m+1}(i); \omega^m) = \min_{j \neq i} [c(i, j) + u(j; \omega^m)] .$$

Z věty 2 vyplývá, že je

$$u(i; \omega^{m+1}) \leqq u(i; \omega^m) \quad \text{pro} \quad i \in I .$$

Přitom $u(i; \omega^{m+1}) \not\equiv u(i; \omega^m)$ dokud ω^m není optimální řízení.

Poznamenejme, že $u(i; \omega^m)$ v rovnicích (15) napočítáváme tím způsobem, že postupujeme proti směru orientovaného spojení z i do r v grafu Γ_m řízení ω^m.

Zmiňme se ještě o dvou modifikacích iteračního postupu k nacházení minimálního nákladu

$$\hat{u}(i) = \min_{\omega} u(i; \omega)$$

na dosažení místa r. Obě v podstatě používají pouze část 2 iteračního kroku.

I. Položme

$$u'_0(i) = c(i, r) , \quad z'_0(i) \equiv r , \quad i = 1, \ldots, r - 1 , \quad u'_0(r) = 0 ,$$

a postupně

$$u'_m(i) = \min_{j \neq i} [c(i, j) + u'_{m-1}(j)] , \tag{16}$$

$$u'_m(r) = 0 , \quad i = 1, \ldots, r - 1 ; \quad m = 1, 2, \ldots$$

$z'_m(i)$ se volí tak, že platí

$$u'_m(i) = c(i, z'_m(i)) + u'_{m-1}(z'_m(i)) , \quad i = 1, \ldots, r - 1 . \tag{17}$$

Z rovnic (17) vidíme, že $u'_m(i)$ je náklad na dosažení místa r, postupujeme-li z místa i nejprve do místa $z'_m(i) = k_1$. Dále z k_1 do $z'_{m-1}(k_1) = k_2$, z k_2 do $z'_{m-2}(k_2)$ atd. Místa r dosáhneme nejpozději v $m + 1$ krocích. Není obtížné dokázati úplnou indukcí, že $u'_m(i)$ je nejmenší náklad, se kterým lze dosáhnouti místa r v nejvýše $m + 1$ krocích. $u'_m(i)$, $m = 0, 1, \ldots$, tvoří tedy nerostoucí posloupnost a $u'_m(i) = \hat{u}(i)$, $i \in I$, pro dostatečně velká m.

II. Označme

$$u''_0(i) = \min_{j \neq i} c(i, j) , \quad i = 1, \ldots, r - 1 , \quad u''_0(r) = 0 ,$$

a volme $z''_0(i)$ tak, aby platilo

$$u''_0(i) = c(i, z''_0(i)) , \quad i = 1, \ldots, r - 1 .$$

Dále postupujeme stejně jako v (16) a (17), tj. klademe

$$u''_m(i) = \min_{j \neq i} [c(i, j) + u''_{m-1}(j)] ,$$

$$u''_m(r) = 0 , \quad i = 1, \ldots, r - 1 ; \quad m = 1, 2, \ldots ,$$

a určujeme $z''_m(i)$ tak, že platí

$$u''_m(i) = c(i, z''_m(i)) + u''_{m-1}(z''_m(i)) , \quad i = 1, \ldots, r - 1 .$$

$u''_m(i)$ je nejmenší náklad, za který můžeme buď uskutečniti $m + 1$ kroků nebo dosáhnouti místa r v nejvýše m krocích. $u''_m(i)$, $m = 0, 1, \ldots,$ tvoří neklesající posloupnost a platí $u''_m(i) = \hat{u}(i)$, $i \in I$ pro dostatečně velká m.

Tabulka 1.

i \ j	1	2	3	4	5	6	7
1	×	4	—	9	—	—	—
2	—	×	4	6	—	—	—
3	—	—	×	—	—	—	12
4	—	—	—	×	4	—	16
5	—	—	—	—	×	2	8
6	—	—	—	—		×	2

Tabulka 2.

i	1	2	3	4	5	6	7
$u_0 = u'_0$	—	—	12	16	8	2	0
$z^1 = z'_1$	4	3	7	5	6	7	
u_1	25	16	12	8	4	2	0
$z^2 = z'_3$	4	4	7	5	6	7	
$u_2 = u'_3 = \hat{u}$	17	14	12	8	4	2	0

i	1	2	3	4	5	6	7
u'_1	25	16	12	12	4	2	0
z'_2	2	3	7	5	6	7	
u'_2	20	16	12	8	4	2	0
z''_0	2	3	7	5	6	7	
$z''_1 = z''_2 = z''_3$	2	4	7	5	6	7	
z''_4	4	4	7	5	6	7	

Pro doplnění uveďme následující numerický příklad.

Příklad 1. Budiž $r = 7$ a veličiny $c(i, j)$ buďtež dány tabulkou 1. V tabulce 1 symbol — představuje dostatečně velké číslo, takže odpovídající přechody mezi místy je možno z úvah vyloučit. Veličiny $z^m(i)$, $u_m(i) = u(i; \omega^m)$ a $z'_m(i)$, $u'_m(i)$ při hledání optimálního řízení iteračním postupem a jeho modifikací I a veličiny $z''_m(i)$ pro modifikaci II jsou uvedeny v tabulce 2. Přitom je zvoleno $z^0(i) \equiv z'_0(i) \equiv r$.

5. DEFINICE ŘÍZENÉHO MARKOVOVA ŘETEZCE

Definici řízeného Markovova řetězce předešleme dva příklady. První z nich je velmi schematický a bude sloužit k numerické ilustraci výpočtových metod popsaných v dalších odstavcích.

Příklad 2. *Schematický příklad o údržbě výrobního zařízení.* Na konci každého období je stav jistého výrobního zařízení klasifikován do tří stupňů 1, 2, 3. Potom je na zařízení provedena údržba buď běžná ($z = 1$ podle označení zavedeného dále), nebo generální ($z = 2$). V tabulce 3 je uvedeno, s jakou pravděpodobností bude zařízení ve stavu 1, 2, 3 v závislosti na posledním stavu a na stupni údržby. Průměrný čistý výnos ze zařízení za jedno období je určen stavem zařízení na konci předchozího období, způsobem údržby a stavem na konci období. Je uveden v tabulce 4 (např. v tisících Kčs).

Příklad 3. *Seřizování stroje.* Uvažujme blok, skládající se ze dvou strojů vyrábějících výrobky stejného druhu. Předpokládá se, že pravděpodobnost vyrobení dobrého výrobku závisí (při

Tabulka 3.

stav na konci období:

stav na konci předchozího období:

$z = 1$	1	2	3
1	1	—	—
2	0,3	0,7	—
3	0,1	0,8	0,1

$z = 2$	1	2	3
1	0,4	0,5	0,1
2	0,2	0,7	0,1
3	—	0,9	0,1

Tabulka 4.

	1	2	3
1	50	—	—
2	55	120	—
3	60	125	150

	1	2	3
1	40	90	120
2	45	110	125
3	—	115	135

dostatečně dobrém přiblížení skutečnosti) pouze na počtu vadných výrobků (zmetků), které stroj od posledního seřízení již vyrobil. Nechť p_k, $k = 0, 1, 2, \ldots$, značí tuto pravděpodobnost v závislosti na počtu zmetků k a nechť platí $p_h = p_{h+1} = \ldots = 0$ pro nějaké číslo h. Oba dva stroje dokončují výrobky současně. Doba práce na výrobku budiž δ. Při seřízení se musí celý blok zastaviti na dobu Δ. Dobrý výrobek představuje příjem velikosti α, zmetek náklad velikosti β. Náklad na seřízení jednoho stroje budiž γ. Naším cílem je nalézt optimální plán seřizování strojů.

Zformulujeme úlohu v pojmech Markovových řetězců. Stav bloku je určen dvojicí čísel (l, m), kde l resp. m je počet zmetků vyrobených od posledního seřízení strojem 1 resp. 2, bylo-li těchto zmetků méně než h. Jinak je l resp. m rovno h. Stav bloku sledujeme vždy v okamžiku dokončení výrobků nebo, je-li blok seřizován, v okamžiku dokončení seřízení. Časový parametr řetězce udává tedy počet dokončených dvojic výrobků, zvětšený o počet seřízení bloku od počátku.

Při každém kroku máme k dispozici tato čtyři rozhodnutí: 1. Seřídit stroj 1. 2. Seřídit stroj 2. 3. Seřídit oba stroje. 4. Nechat blok v provozu. Pravděpodobnosti přechodu $p(i, j; z)$ ze stavu i do stavu j při jednotlivých rozhodnutích (z — pořadové číslo rozhodnutí) a výnosy $c(i, j; z)$ vyplývající z těchto přechodů, jsou uvedeny v tabulce 5. Přechody v tabulce neuvažované mají pravděpodobnost 0.

Řízený Markovův řetězec, jehož stavy jsou označeny celými čísly $\{1, 2, \ldots, r\} = I$, je definován soustavou

$$\|p(i, j; z)\|_{i,j=1}^{r}, \quad z \in J = \{1, 2, \ldots, s\}, \tag{1}$$

Tabulka 5.

$p(i, j; z)$:
$i = (l, m)$

z	1	2	3	4			
j	$(0, m)$	$(l, 0)$	$(0, 0)$	(l, m)	$(l + 1, m)$	$(l, m + 1)$	$(l + 1, m + 1)$
$p(i, j; z)$	1	1	1	$p_l p_m$	$(1 - p_l) p_m$	$p_l(1 - p_m)$	$(1 - p_l)(1 - p_m)$

Je-li $m = h$ ($l = h$) klademe $m + 1 = h$ ($l + 1 = h$).

$c(i, j; z)$:

z	1	2	3	4			
j	$(0, m)$	$(l, 0)$	$(0, 0)$	(l, m)	$(l + 1, m)$	$(l, m + 1)$	$(l + 1, m + 1)$
$l \neq h \neq m$	$-\gamma$	$-\gamma$	-2γ	2α	$\alpha - \beta$	$\alpha - \beta$	-2β
$l = h \neq m$	$-\gamma$	$-\gamma$	-2γ	$\alpha - \beta$	—	-2β	—
$l \neq h = m$	$-\gamma$	$-\gamma$	-2γ	$\alpha - \beta$	-2β	—	—
$l = h = m$	$-\gamma$	$-\gamma$	-2γ	-2β	—	—	—

matic pravděpodobností přechodu, závisejících na parametru řízení z a soustavou výnosových matic

$$(2) \qquad \|c(i, j; z)\|_{i,j=1}^{r}, \quad z \in J = \{1, 2, \ldots, s\}.$$

$p(i, j; z)$ je pravděpodobnost přechodu ze stavu i do stavu j za podmínky, že parametr řízení nabyl hodnoty z. $c(i, j; z)$ je výnos, plynoucí z takového přechodu.

Řízením řetězce rozumíme posloupnost

$$\omega = \{z_n(j_0, \ldots, j_n), \quad n = 0, 1, \ldots\}$$

funkcí na $I^{n+1} = I \times \ldots \times I$ s hodnotami v J. $z_n(j_0, \ldots, j_n)$ udává hodnotu parametru řízení, kterou volíme po n-tém kroku za podmínky, že v prvních n krocích řetězec prošel posloupností stavů $j_0, j_1, \ldots, j_n$. Řízení ω se nazývá *markovským*, když $z_n(j_0, j_1, \ldots, j_n)$ závisí pouze na j_n, tj. $\omega = \{z_n(j_n), n = 0, 1, \ldots\}$. Markovské řízení ω se nazývá *homogenním*, když $z_n(j_n) = z(j_n)$ pro $n = 0, 1, 2, \ldots$ Homogenní markovské řízení je jednoznačně určeno funkcí $z(j)$. Budeme proto psát $\omega \sim z(j)$.

Při zvoleném řízení ω a počátečním stavu $j \in I$ soustava (1) současně se vztahem

$$(3) \qquad P(X_{n+1} = j_{n+1} \mid X_0 = j, X_1 = j_1, \ldots, X_n = j_n) = \\ = p(j_n, j_{n+1}; z_n(j, j_1, \ldots, j_n))$$

určují rozložení pravděpodobností posloupnosti $\{X_n, n = 0, 1, \ldots\}$ náhodných veličin. Střední hodnotu vzhledem k tomuto rozložení budeme označovati symbolem E_j^{ω}. Je-li ω markovské řízení, tvoří $\{X_n, n = 0, 1, \ldots\}$ Markovův řetězec s pravděpodobnostmi přechodu $p(i, j; z_n(i))$, jak plyne srovnáním (3) a (1 § 2).

Celkový výnos z posloupnosti $\{X_n, n = 0, 1, \ldots\}$ do doby N je roven

$$\sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1})).$$

Při hledání optimálního řízení se snažíme maximalizovat vhodně definovaný očekávaný (průměrný) výnos z posloupnosti $\{X_n, n = 0, 1, \ldots\}$. Vcelku lze uvésti čtyři základní způsoby definice očekávaného výnosu.

1. *Očekávaný výnos do předem zvolené doby.* Je dána pevná doba N. Celkový očekávaný výnos do této doby je

$$(4) \qquad E_j^{\omega} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1})).$$

2. *Očekávaný diskontovaný výnos.* Pro zvolené číslo β, $0 < \beta < 1$, tzv. diskontní faktor, očekávaným diskontovaným výnosem se rozumí střední hodnota

$$E_j^{\omega} \sum_{n=1}^{\infty} \beta^{n-1} c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1})).$$

3. *Očekávaný výnos do dosažení zvolené množiny* je definován výrazem (4), kde N značí nyní dobu prvního dosažení zvolené množiny stavů I_0 (popř. zvoleného stavu), tj.

$$N = \inf\{n : X_n \in I_0\}\,.$$

4. *Průměrný výnos na jednotku času* je definován jako limita

$$\lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1}))$$

za předpokladu, že tato s pravděpodobností 1 existuje a je rovna konstantě. V obecném případě hledáme řízení, které maximalizuje

$$E_j^{\omega} \liminf_{N\to\infty} N^{-1} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1}))\,.$$

Otázkám nacházení optimálního řízení při různých definicích výnosu jsou věnovány další odstavce.

6. OČEKÁVANÝ VÝNOS DO PŘEDEM ZVOLENÉ DOBY

Řízený Markovův řetězec budiž definován soustavami matic (1 § 5), (2 § 5). Zaveďme očekávaný výnos z jednoho přechodu

$$r(j, z) = \sum_k p(j, k; z)\, c(j, k; z)\,. \tag{1}$$

Budiž dále $\{X_n, n = 0, 1, \ldots\}$ náhodová posloupnost, odpovídající řízení ω. Označme

$$u_m^N(j_0, \ldots, j_m; \omega) = E_{j_0}^{\omega}\{\sum_{n=m+1}^{N} c(X_{n-1}, X_n; z_{n-1}) \mid X_1 = j_1, \ldots, X_m = j_m\}\,.$$

Z (3 § 5) plyne

$$\begin{aligned} u_{m-1}^N(j_0, \ldots, j_{m-1}; \omega) &= E_{j_0}^{\omega}\{c(X_{m-1}, X_m; z_{m-1}) \mid X_1 = j_1, \ldots, X_{m-1} = j_{m-1}\} + \\ &+ E_{j_0}^{\omega}\{\sum_{n=m+1}^{N} c(X_{n-1}, X_n; z_{n-1}) \mid X_1 = j_1, \ldots, X_{m-1} = j_{m-1}\} = \\ &= r(j_{m-1}, z_{m-1}(j_0, \ldots, j_{m-1})) + \sum_k p(j_{m-1}, k; z_{m-1}(j_0, \ldots, j_{m-1}))\, u_m^N(j_0, \ldots, j_{m-1}, k; \omega). \end{aligned} \tag{2}$$

Z tohoto rekurentního vztahu pro $u_m^N(j_0, \ldots, j_m; \omega)$ budeme vycházet při určování maximálního výnosu a optimálního řízení.

Položme

$$\hat{u}_{N-1}^N(j) = \max_z r(j, z)\,.$$

Zřejmě to nejlepší, co lze v posledním kroku udělat je zvolit tu hodnotu parametru řízení, která maximalizuje očekávaný zisk na jeden krok. Platí proto

$$u_{N-1}^N(j_0, \ldots, j_{N-1}; \omega) = r(j_{N-1}, z_{N-1}(j_0, \ldots, j_{N-1})) \leqq \hat{u}_{N-1}^N(j_{N-1}) \,. \tag{3}$$

Definujme postupně pro $m = N - 1, N - 2, \ldots, 1$

$$\hat{u}_{m-1}^N(j) = \max_z \left[r(j, z) + \sum_k p(j, k; z)\, \hat{u}_m^N(k)\right], \quad j \in I \,. \tag{4}$$

S použitím (2) ověříme indukcí, že platí (pro libovolné ω)

$$u_m^N(j_0, \ldots, j_m; \omega) \leqq \hat{u}_m^N(j_m) \tag{5}$$

pro $m = N - 1, N - 2, \ldots, 0$ a všechna $j_0, \ldots, j_m$. Pro $m = N - 1$ je (5) totožné s (3). Platí-li (5), můžeme psát

$$\begin{aligned} u_{m-1}^N(j_0, \ldots, j_{m-1}; \omega) &\leqq r(j_{m-1}, z_{m-1}(j_0, \ldots, j_{m-1})) + \\ &+ \sum_k p(j_{m-1}, k; z_{m-1}(j_0, \ldots, j_{m-1}))\, \hat{u}_m^N(k) \leqq \\ &\leqq \max_z \left[r(j_{m-1}, z) + \sum_k p(j_{m-1}, k; z)\, \hat{u}_m^N(k)\right] = \hat{u}_{m-1}^N(j_{m-1}) \,. \end{aligned}$$

Stejně pokračujeme dále až dostaneme nerovnost

$$u_0^N(j; \omega) \leqq \hat{u}_0^N(j) \quad \text{pro} \quad j \in I \,.$$

To znamená, že $\hat{u}_0^N(j)$ je maximum očekávaného výnosu.

Ze vztahu (4) obdržíme také optimální řízení. Budiž $\hat{z}_{m-1}(j)$ ta hodnota parametru řízení, pro níž výraz v hranaté závorce ve (4) nabývá maxima, tj.

$$\left[r(j, \hat{z}_{m-1}(j)) + \sum_k p(j, k; \hat{z}_{m-1}(j))\, \hat{u}_m^N(k)\right] = \max_z \left[r(j, z) + \sum_k p(j, k; z)\, \hat{u}_m^N(k)\right].$$

Potom, jak se snadno vidí, $\hat{\omega} = \{\hat{z}_n(j), n = 0, \ldots, N - 1\}$ je optimální řízení, tj. $u_0^N(j; \hat{\omega}) = \hat{u}_0^N(j)$ pro $j \in I$. Je to markovské řízení.

Vztah (4) je jedním z příkladů *Bellmanovy rovnice*, která je základem dynamického programování. Má tento názorný smysl: Výraz v hranaté závorce je střední hodnota výnosu z m-tého přechodu a ze zbývajících $N - m$ přechodů za těchto podmínek: 1. $X_{m-1} = j$. 2. Při m-tém přechodu volíme parametr řízení roven z. 3. V posledních $N - m$ krocích docílíme optimálního výnosu. Odtud je patrné, že k dosažení maximálního očekávaného výnosu v posledních $N - m + 1$ krocích je třeba v čase $m - 1$ voliti parametr řízení tak, aby výraz v hranaté závorce ve (4) byl co největší.

Příklad 2a. Za předpokladů uvedených v příkladu 2 stanovíme způsob údržby, který maximalizuje očekávaný výnos za 5 období. Postupujeme tak, že pro $m = 5, 4, 3, 2, 1$ napočítáváme veličiny

$$r(j, z) + \sum_k p(j, k; z)\, \hat{u}_m^5(k) \,.$$

Výsledek je uveden v tabulce 6, kde hodnoty optimálního očekávaného výnosu $\hat{u}^5_{m-1}(j)$ jsou vytištěny polotučně.

Z tabulky 6 vyplývá, že optimální řízení je $\hat{\omega} = \{\hat{z}_n(j), n = 0, 1, 2, 3, 4\}$, kde $\hat{z}_n(1) = 2$, $\hat{z}_n(2) = 2$, $\hat{z}_n(3) = 1$ pro $n = 0, 1, 2, 3$; $\hat{z}_4(1) = 2$, $\hat{z}_4(2) = 1$, $\hat{z}_4(3) = 1$.

Tabulka 6.

z	m \ j	1	2	3
1	5	50,00	**100,50**	**121,00**
2		**73,00**	98,50	117,00
1	4	123,00	192,55	**220,80**
2		**164,55**	**195,55**	219,55
1	3	214,55	286,75	**315,98**
2		**258,68**	**290,38**	315,08
1	2	308,68	381,37	**410,77**
2		**353,26**	**385,10**	409,94
1	1	403,26	476,04	**505,48**
2		**447,93**	**479,79**	504,66

7. OČEKÁVANÝ DISKONTOVANÝ VÝNOS

Budiž $\omega = \{z_0(j_0), z_1(j_0, j_1), \ldots\}$ řízení řetězce definovaného soustavami (1 § 5), (2 § 5) a $\{X_n, n = 0, 1, \ldots\}$ budiž odpovídající náhodová posloupnost. Očekávaný diskontovaný výnos, přiřazený tomuto řízení, označme

$$v(j; \omega) = v(j, \beta; \omega) = E_j^{\omega} \sum_{n=1}^{\infty} \beta^{n-1} c(X_{n-1}, X_n; z_{n-1}) .$$

Symbol $v(j, \beta; \omega)$ budeme užívat, bude-li třeba závislost na diskontním faktoru β, $0 < \beta < 1$, zvláště vyznačit.

Nechť řízení $\omega^{(k)}$, $k \in I$, jsou odvozena z řízení ω posunutím o jeden krok do stavu k, tj.

$$\omega^{(k)} = \{z_1(k, j_0), z_2(k, j_0, j_1), \ldots\} .$$

rv č. 1/1969

Z definice řízeného řetězce a diskontovaného výnosu plyne bezprostředně vztah

$$(1) \qquad v(j;\omega) = r(j, z_0(j)) + \beta \sum_k p(j, k; z_0(j))\, v(j; \omega^{(k)}), \quad j \in I .$$

$r(j, z)$ je definováno v (1 § 6).

Zaveďme maximální očekávaný diskontovaný výnos

$$\hat{v}(j) = \max_{\omega} v(j;\omega), \quad j \in I .$$

Řízení $\hat{\omega}$ se nazývá optimálním, platí-li

$$v(j;\hat{\omega}) = \hat{v}(j), \quad j \in I .$$

Dokážeme tuto větu:

Věta 1. *Maximální výnos $\hat{v}(j)$ je jediným řešením rovnice*

$$(2) \qquad \hat{v}(j) = \max_z \left[r(j, z) + \beta \sum_k p(j, k; z)\, \hat{v}(k)\right], \quad j \in I .$$

Je-li $\hat{z}(j)$ hodnota parametru řízení, pro niž výraz v hranaté závorce ve (2) nabývá maximální hodnoty, potom $\hat{\omega} \sim \hat{z}(j)$ je optimální (homogenní markovské) řízení.

Poznámka 1. Při srovnání Bellmanovy rovnice (2) s rovnicí (4 § 6) vyniká rozdíl mezi optimalizační úlohou § 6 a nynější úlohou, ve které po jednom kroku řetězce stojíme před stejným problémem maximalizace diskontovaného výnosu jako na počátku.

Důkaz věty 1. Optimální řízení lze sestrojiti diagonálním postupem. Budiž

$$(3) \qquad \omega^k = \{z_0^k(j_0), z_1^k(j_0, j_1), \ldots\}, \quad k = 1, 2, \ldots,$$

posloupnost řízení taková, že

$$\lim_{k \to \infty} v(j; \omega^k) = \hat{v}(j), \quad j \in I .$$

Jelikož zobrazení I do J je konečně mnoho, existuje $\bar{z}_0(j_0)$ tak, že v posloupnosti (3) $\bar{z}_0(j_0) \equiv z_0^k(j_0)$ pro nekonečně mnoho k. Vyberme z (3) podposloupnost

$$(4) \qquad {}^0\omega^k = \{{}^0z_0^k(j_0), {}^0z_1^k(j_0, j_1), \ldots\}, \quad k = 1, 2, \ldots,$$

ve které ${}^0z_0^k(j_0) \equiv \bar{z}_0(j_0)$ pro $k = 1, 2, \ldots$ Existuje dále $\bar{z}_1(j_0, j_1)$ tak, že $\bar{z}_1(j_0, j_1) \equiv$ $\equiv {}^0z_1^k(j_0, j_1)$ pro nekonečně mnoho k. Lze tedy ze (4) vybrat podposloupnost

$${}^1\omega^k = \{{}^1z_0^k(j_0), {}^1z_1^k(j_0, j_1), \ldots\}, \quad k = 1, 2, \ldots,$$

ve které ${}^1z_1^k(j_0, j_1) \equiv \bar{z}_1(j_0, j_1)$ pro $k = 1, 2, \ldots$ Tímto postupem získáme pro každé m podposloupnost

$${}^m\omega^k = \{{}^mz_0^k(j_0), {}^mz_1^k(j_0, j_1), \ldots\}, \quad k = 1, 2, \ldots,$$

splňující

$$ {}^m z_0^k(j_0) \equiv \tilde{z}_0(j_0)\,,\quad {}^m z_1^k(j_0, j_1) \equiv \tilde{z}_1(j_0, j_1)\,,\ \ldots\,,\ {}^m z_m^k(j_0, \ldots, j_m) \equiv $$
$$ \equiv \tilde{z}_m(j_0, \ldots, j_m) \quad \text{pro} \quad k = 1, 2, \ldots $$

Ukážeme, že $\tilde{\omega} = \{\tilde{z}_n(j_0, \ldots, j_n), n = 0, 1, \ldots\}$ je optimální řízení. Označme

$$ \bar{c} = \max_{i,j,z} |c(i, j; z)|\,. $$

Platí

$$ (5) \qquad |v(j; \tilde{\omega}) - v(j; {}^m\omega^k)| = \Big| \sum_{n=m+2}^{\infty} \beta^{n-1}(E_j^{\tilde{\omega}} c(X_{n-1}, X_n; \tilde{z}_{n-1}) - $$
$$ - E_j^{m\omega^k} c(X_{n-1}, X_n; {}^m z_{n-1}^k))\Big| \leqq 2\beta^{m+1}\bar{c}/(1 - \beta)\,. $$

Odtud limitním přechodem pro $k \to \infty$

$$ |v(j; \tilde{\omega}) - \hat{v}(j)| \leqq 2\beta^{m+1}\bar{c}/(1 - \beta) $$

a dále pro $m \to \infty$

$$ v(j; \tilde{\omega}) = \hat{v}(j) \quad \text{pro} \quad j \in I\,. $$

Z (1) vyplývá

$$ (6) \qquad \hat{v}(j) = v(j; \tilde{\omega}) = r(j, \tilde{z}_0(j)) + \beta \sum_k p(j, k; \tilde{z}_0(j))\, v(k; \tilde{\omega}^{(k)}) \leqq $$
$$ \leqq r(j, \tilde{z}_0(j)) + \beta \sum_k p(j, k; \tilde{z}_0(j))\, \hat{v}(k) \leqq $$
$$ \leqq \max_z \left[r(j, z) + \beta \sum_k p(j, k; z)\, \hat{v}(k)\right], \quad j \in I\,. $$

Budiž $\hat{z}(j)$ funkce definovaná ve znění věty. Utvořme řízení $\tilde{\omega}_0 = \{\hat{z}(j_0), \tilde{z}_0(j_1), \tilde{z}_1(j_1, j_2), \ldots\}$. Potom dle definice $\tilde{\omega}_0$ a dle (1)

$$ (7) \qquad \hat{v}(j) \geqq v(j; \tilde{\omega}_0) = r(j, \hat{z}(j)) + \beta \sum_k p(j, k; \hat{z}(j))\, \hat{v}(k; \tilde{\omega}_0^{(k)}) = $$
$$ = r(j, \hat{z}(j)) + \beta \sum_k p(j, k; \hat{z}(j))\, \hat{v}(k) = $$
$$ = \max_z \left[r(j, z) + \beta \sum_k p(j, k; z)\, \hat{v}(k)\right], \quad j \in I\,. $$

Ze (6) a (7) plyne (2) a také

$$ v(j; \tilde{\omega}_0) = \hat{v}(j) \quad \text{pro} \quad j \in I\,. $$

Obdobně se zjistí, že pro řízení $\tilde{\omega}_1 = \{\hat{z}(j_0), \hat{z}(j_1), \tilde{z}_0(j_2), \tilde{z}_1(j_2, j_3), \ldots\}$ platí

$$ v(j; \tilde{\omega}_1) = \hat{v}(j)\,, \quad j \in I\,, $$

a obecně pro $\tilde{\omega}_m = \{\hat{z}(j_0), \ldots, \hat{z}(j_m), \tilde{z}_0(j_{m+1}), \tilde{z}_1(j_{m+1}, j_{m+2}), \ldots\}$

$$ v(j; \tilde{\omega}_m) = \hat{v}(j)\,, \quad j \in I\,. $$

Použitím odhadu analogického (5) ověříme, že je

$$v(j; \hat{\omega}) = \lim_{m \to \infty} \hat{v}(j; \tilde{\omega}_m) = \hat{v}(j) \quad \text{pro} \quad j \in I .$$

Vidíme, že $\hat{\omega} \sim \hat{z}(j)$ je optimální řízení.

Zbývá dokázat jednoznačnost rovnice (2). Nechť platí

$$\bar{v}(j) = \max_z \left[r(j, z) + \beta \sum_k p(j, k; z)\, \bar{v}(k)\right], \quad j \in I .$$

Potom

$$\bar{v}(j) \geqq r(j, \hat{z}(j)) + \beta \sum_k p(j, k; \hat{z}(j))\, \bar{v}(k) .$$

Odečteme-li od této nerovnosti vztah

$$\hat{v}(j) = r(j, \hat{z}(j)) + \beta \sum_k p(j, k; \hat{z}(j))\, \hat{v}(k) ,$$

dostáváme

$$\bar{v}(j) - \hat{v}(j) \geqq \beta \sum_k p(j, k; \hat{z}(j)) \left(\bar{v}(k) - \hat{v}(k)\right) \geqq$$

$$\geqq \beta \min_k \left[\bar{v}(k) - \hat{v}(k)\right], \quad j \in I .$$

Odtud plyne

(8) $$\min_j \left[\bar{v}(j) - \hat{v}(j)\right] \geqq \beta \min_k \left[\bar{v}(k) - \hat{v}(k)\right] .$$

Vzhledem k tomu, že je $\beta < 1$, může (8) platit pouze tehdy, je-li

$$\min_j \left[\bar{v}(j) - \hat{v}(j)\right] \geqq 0 , \quad \text{tj.} \quad \bar{v}(j) \geqq \hat{v}(j) , \quad j \in I .$$

Stejně lze ovšem dokázat $\hat{v}(j) \geqq \bar{v}(j)$, $j \in I$, odkud $\bar{v}(j) = \hat{v}(j)$ pro $j \in I$. □

Poznámka 2. Tímtéž způsobem, jako byla dokázána jednoznačnost rovnice (2) ve větě 1, lze ukázat, že při $\omega \sim z(j)$ soustava rovnic

$$v(j; \omega) = r(j, z(j)) + \beta \sum_k p(j, k; z(j))\, v(k; \omega) , \quad j \in I ,$$

jednoznačně určuje $v(j; \omega)$.

K řešení rovnice (2) a k nalezení optimálního řízení se užívá *Howardova iteračního postupu*: Vychází se z libovolného řízení $\omega^0 \sim z^0(j)$ a určuje se postupně $\omega^1 \sim z^1(j)$, $\omega^2 \sim z^2(j), \ldots, \omega^n \sim z^n(j), \ldots$ Na základě $z^n(j)$ se definuje $z^{n+1}(j)$ takto: Ze systému rovnic

(9) $$v(j; \omega^n) = r(j, z^n(j)) + \beta \sum_k p(j, k; z^n(j))\, v(k; \omega^n) , \quad j \in I ,$$

se stanoví výnos odpovídající řízení $\omega^n \sim z^n(j)$. $z^{n+1}(j)$ se klade rovným hodnotě parametru z, pro kterou výraz

$$r(j, z) + \beta \sum_k p(j, k; z)\, v(k; \omega^n) \tag{10}$$

nabývá maxima. Platí tedy

$$v(j; \omega^n) \leqq r(j, z^{n+1}(j)) + \beta \sum_k p(j, k; z^{n+1}(j))\, v(k; \omega^n), \tag{11}$$

$$v(j; \omega^{n+1}) = r(j, z^{n+1}(j)) + \beta \sum_k p(j, k; z^{n+1}(j))\, v(k; \omega^{n+1}), \tag{12}$$

odkud

$$v(j; \omega^n) - v(j; \omega^{n+1}) \leqq \beta \sum_k p(j, k; z^{n+1}(j)) \left(v(k; \omega^n) - v(k; \omega^{n+1})\right) \leqq$$

$$\leqq \beta \max_k \left[v(k; \omega^n) - v(k; \omega^{n+1})\right],$$

neboli

$$\max_j \left[v(j; \omega^n) - v(j; \omega^{n+1})\right] \leqq \beta \max_k \left[v(k; \omega^n) - v(k; \omega^{n+1})\right].$$

Poslední nerovnost může být při $0 < \beta < 1$ splněna jen tehdy, když $v(j; \omega^{n+1}) \geqq$ $\geqq v(j; \omega^n)$ pro $j \in I$.

$v(j; \omega^n)$ neklesá, roste-li n. Jelikož množina homogenních markovských řízení je konečná, platí pro nějaké n rovnost $v(j; \omega^n) = v(j; \omega^{n+1})$ pro $j \in I$. Z (12) potom plyne

$$v(j; \omega^n) = r(j, z^{n+1}(j)) + \beta \sum_k p(j,k;z^{n+1}(j))\, v(k; \omega^n) =$$

$$= \max_z \left[r(j, z) + \beta \sum_k p(j, k; z)\, v(k; \omega^n)\right], \quad j \in I.$$

Tedy $v(j; \omega^n)$ vyhovuje rovnici (2) a ω^n je optimální řízení.

Příklad 2b. Za podmínek příkladu 2 určíme způsob údržby, který maximalizuje očekávaný diskontovaný výnos při diskontním faktoru $\beta = 0{,}95$. Z tabulek 3, 4 vypočteme $r(j, z)$:

$j =$	1	2	3
$r(j, 1)$	50	100,5	121
$r(j, 2)$	73	98,5	117

Volme $z^0(j) \equiv 1$. Soustava rovnic (9) je potom

$$\begin{aligned} 0{,}050v(1;\omega^0) &= 50 \\ -0{,}285v(1;\omega^0) + 0{,}335v(2;\omega^0) &= 100{,}5 \\ -0{,}095v(1;\omega^0) - 0{,}760v(2;\omega^0) + 0{,}905v(3;\omega^0) &= 121\,. \end{aligned}$$

Řešením dostáváme

$$v(1;\omega^0) = 1000{,}0\,;\quad v(2;\omega^0) = 1150{,}7\,;\quad v(3;\omega^0) = 1205{,}0\,.$$

Veličina (10) nabývá pro $n = 0$ těchto hodnot:

z \ j	1	2	3
1	1000,0	1150,7	1205,0
2	1114,1	1168,2	1215,4

Odtud vyplývá, že je třeba volit $z^1(j) \equiv 2$. Potom soustava (9) je

$$\begin{aligned} 0{,}620v(1;\omega^1) - 0{,}475v(2;\omega^1) - 0{,}095v(3;\omega^1) &= 73 \\ -0{,}190v(1;\omega^1) + 0{,}335v(2;\omega^1) - 0{,}095v(3;\omega^1) &= 98{,}5 \\ - 0{,}855v(2;\omega^1) + 0{,}905v(3;\omega^1) &= 117\,. \end{aligned}$$

Řešení soustavy je

$$v(1;\omega^1) = 1865{,}4\,;\quad v(2;\omega^1) = 1896{,}9\,;\quad v(3;\omega^1) = 1921{,}4\,.$$

Dále určíme hodnoty veličiny (10) pro $n = 1$:

z \ j	1	2	3
1	1822,1	1893,6	1922,4
2	1865,4	1897,0	1921,4

Klademe tedy $z^2(1) = 2$, $z^2(2) = 2$, $z^2(3) = 1$.

Platí

$$\begin{aligned} 0{,}620v(1;\omega^2) - 0{,}475v(2;\omega^2) - 0{,}095v(3;\omega^2) &= 73 \\ -0{,}190v(1;\omega^2) + 0{,}335v(2;\omega^2) - 0{,}095v(3;\omega^2) &= 98{,}5 \\ -0{,}095v(1;\omega^2) - 0{,}760v(2;\omega^2) + 0{,}905v(3;\omega^2) &= 121\,. \end{aligned}$$

Odtud

$$v(1;\omega^2) = 1867{,}3\,;\quad v(2;\omega^2) = 1898{,}8\,;\quad v(3;\omega^2) = 1924{,}3\,.$$

Z tabulky veličin (10) pro $n = 2$:

$z \backslash j$	1	2	3
1	1824,0	1895,4	1924,3
2	1867,3	1898,8	1923,3

vyplývá, že $\omega^2 \sim z^2(j)$ maximalizuje očekávaný diskontovaný výnos.

Chceme-li ke dvěma řízením $\omega \sim z(j)$, $\omega' \sim z'(j)$ sestrojit řízení $\omega'' \sim z''(j)$, které je stejnoměrně (tj. pro všechny počáteční polohy) lepší, můžeme použíti *Eatonovu-Zadehovu metodu.* Klademe

$$z''(i) = z(i) \quad \text{pro } i, \text{ pro něž } \; v(i; \omega) \geqq v(i; \omega'),$$
$$z''(i) = z'(i) \quad \text{pro } i, \text{ pro něž } \; v(i; \omega) < v(i; \omega').$$

Potom platí

$$v(j; \omega'') \geqq \max\{v(j; \omega), v(j; \omega')\}, \quad j \in I. \tag{13}$$

Abychom se o tom přesvědčili, označme $\tilde{v}(j) = \max\{v(j; \omega), v(j; \omega')\}$ a zvolme $i \in I$. Budiž např. $z''(i) = z(i)$. Potom

$$\tilde{v}(i) = v(i; \omega) = r(i, z(i)) + \beta \sum_k p(i, k; z(i))\, v(k; \omega) \leqq$$
$$\leqq r(i, z''(i)) + \beta \sum_k p(i, k; z''(i))\, \tilde{v}(k).$$

Dále

$$v(i; \omega'') = r(i, z''(i)) + \beta \sum_k p(i, k; z''(i))\, v(k; \omega'').$$

Odtud způsobem použitým na (11), (12) usoudíme, že je $v(i; \omega'') \geqq \tilde{v}(i)$ pro $i \in I$. Čímž je (13) dokázáno.

8. OČEKÁVANÝ VÝNOS DO DOSAŽENÍ ZVOLENÉ MNOŽINY

Nestochastickou úlohou dosažení zvoleného stavu jsme se zabývali v odstavci 4. Jednalo se o speciální případ řízeného řetězce, kdy v soustavě (1 § 5) $p(i, z; z) = 1$, $p(i, j; z) = 0$ při $j \neq z$; pro $i \in I$, $z \in J = \{1, 2, \ldots, r\}$.

Zvolme (neprázdnou) množinu I_0 stavů řízeného řetězce, definovaného soustavami (1 § 5), (2 § 5). Budiž $\{X_n, n = 0, 1, \ldots\}$ jeho trajektorie a

$$N = \inf\{n : X_n \in I_0\}$$

doba prvého dosažení množiny I_0. Řízení ω nazveme přípustným, je-li $E_j^\omega N < \infty$ pro všechna $j \in I$. Je-li $\omega = \{z_n(j_0, \ldots, j_n), n = 0, 1, \ldots\}$ přípustným řízením, potom očekávaný výnos do dosažení množiny I_0

$$u(j; \omega) = E_j^\omega \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}), \quad j \in I,$$

má smysl.

Věta 1. *Je-li každé homogenní markovské řízení přípustným, je každé řízení přípustným.*

Důkaz. Pro jednoduchost označení budeme předpokládat, že I_0 je tvořeno jediným stavem 1. Budiž $\{\tilde{X}_n, n = 0, 1, \ldots\}$ trajektorie, která vznikne z $\{X_n, n = 0, 1, \ldots\}$ zastavením soustavy ve stavu 1 jako při důkazu věty 4 § 2. Systém přechodových matic odpovídající zastavení soustavy ve stavu 1 je

$$\|\tilde{p}(i, j; z)\|_{i,j=1}^r = \begin{pmatrix} 1, & 0, & \ldots, & 0 \\ p(2, 1; z), & p(2, 2; z), & \ldots, & p(2, r; z) \\ \ldots & \ldots & \ldots & \ldots \\ p(r, 1; z), & p(r, 2; z), & \ldots, & p(r, r; z) \end{pmatrix}, \quad z \in J.$$

Dále zaveďme výnosovou matici

$$\|\tilde{c}(i, j)\|_{i,j=1}^r = \begin{pmatrix} 0, & 0, & \ldots, & 0 \\ 1, & 1, & \ldots, & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 1, & 1, & \ldots, & 1 \end{pmatrix}.$$

Předpokládejme, že všechna homogenní markovská řízení jsou přípustná a označme

(1) $$\hat{E} = \max\{E_j^\omega N : \omega \text{ homogenní markovské řízení}, j \in I\}.$$

Budiž ω libovolné řízení. Zřejmě platí

$$\tilde{E}_j^\omega \sum_{n=1}^{\infty} \tilde{c}(\tilde{X}_{n-1}, \tilde{X}_n) = E_j^\omega N.$$

Zaveďme diskontovaný výnos

$$\tilde{v}(j, \beta; \omega) = \tilde{E}_j^\omega \sum_{n=1}^{\infty} \beta^{n-1} \tilde{c}(\tilde{X}_{n-1}, \tilde{X}_n) = E_j^\omega \frac{1 - \beta^N}{1 - \beta}.$$

Z věty 1 § 7 vyplývá, že existuje homogenní markovské řízení ω_β tak, že platí

$$\tilde{v}(j, \beta; \omega) \leqq \tilde{v}(j, \beta; \omega_\beta), \quad j \in I,$$

a tedy

$$E_j^\omega \frac{1-\beta^N}{1-\beta} \leqq E_j^{\omega_\beta} \frac{1-\beta^N}{1-\beta} \leqq E_j^{\omega_\beta} N \leqq \hat{E}\,,$$

odkud

$$E_j^\omega N = \lim_{\beta \to 1} E_j^\omega \frac{1-\beta^N}{1-\beta} \leqq \hat{E} < \infty\,. \quad \square \tag{2}$$

Z věty 1 a věty 4 § 2 vyplývá tento důsledek: Tvoří-li I při každém homogenním markovském řízení $\omega \sim z(j)$ izolovanou třídu rekurentních stavů (vzhledem k přechodové matici $\|p(i,j;z(i))\|_{i,j=1}^r$), potom jsou všechna řízení přípustná. S použitím vět 3 § 2 a 4 § 2 lze také snadno dokázat, že všechna řízení jsou přípustná právě když pro každé homogenní markovské řízení $\omega \sim z(j)$ platí: I_0 obsahuje aspoň jeden stav z každé izolované třídy rekurentních stavů vzhledem k přechodové matici

$$\|p(i,j;z(i))\|_{i,j=1}^r\,.$$

Metody maximalizace $u(j;\omega)$ jsou obdobné metodám předchozího odstavce a nebudeme se o nich podrobně šířit. Lze je odvodit z výsledků § 7 limitním přechodem $\beta \to 1$, kterého bylo použito při důkazu věty 1. Zejména platí: Nechť všechna řízení jsou přípustná pro I_0, potom maximální výnos

$$\hat{u}(j) = \max_\omega u(j;\omega)\,, \quad j \in I - I_0\,,$$

splňuje Bellmanovu rovnici

$$\hat{u}(j) = \max_z \left[r(j,z) + \sum_{k \in I - I_0} p(j,k;z)\,\hat{u}(k)\right], \quad j \in I - I_0\,. \tag{3}$$

Uveďme ještě příklad.

Příklad 4. *Strategie hráče.* Hráč mající k stokorun hraje hazardní hru s protivníkem, který je vždy ochoten ke hře. S pravděpodobností p získá hráč výhru, rovnou vsazené částce, s pravděpodobností q vsazenou částku prohraje. Rozhodl se skončit hru, když vyhraje $r - k$ stokorun, nebude-li dříve ruinován. Chce hráti tak, aby pravděpodobnost dosažení kapitálu r stokorun byla co největší. Jeho strategie $\omega \sim z(j)$ udává velikost sázky v závislosti na počtu stokorun j, které má v průběhu hry k dispozici.

Ukažme nejprve, že v případě $p = q = \frac{1}{2}$ jsou všechny strategie rovnocenné. Budiž $u(j)$ pravděpodobnost dosažení kapitálu r za předpokladu, že výchozí kapitál hráče byl j a že sází pokaždé jednu stokorunu, pokud nedosáhl kapitálu r. Zřejmě platí

$$u(j) = \tfrac{1}{2}[u(j-1) + u(j+1)]\,, \quad j = 2, \dots, r-2\,,$$
$$u(1) = \tfrac{1}{2}u(2)\,, \quad u(r-1) = \tfrac{1}{2} + \tfrac{1}{2}u(r-2)\,.$$

Odtud

$$u(j) = j/r\,, \quad j = 1, \dots, r-1\,. \tag{4}$$

Položme $u(0) = 0$, $u(r) = u(r + 1) = \ldots = 1$. Potom podmínku (3) maximálnosti $u(j)$ lze psáti

(5) $$u(j) = \max_{0<z\leqq j} \tfrac{1}{2}(u(j - z) + u(j + z)), \quad j = 1, \ldots, r - 1,$$

kde z značí velikost sázky. Z (4) vyplývá, že (5) je splněno, a také, že platí

(6) $$\tfrac{1}{2}[u(j - \dot{z}(j)) + u(j + \dot{z}(j))] = \max_{0<z\leqq j} \tfrac{1}{2}(u(j - z) + u(j + z)), \quad j = 1, \ldots, r - 1,$$

kdykoliv $\dot{z}(j) \leqq \min\{j, r - j\}$. Rovnice (6) charakterizuje optimální řízení, takže skutečně všechny strategie, při nichž sázka nepřekračuje $r - j$, jsou rovnocenné.

Volme nyní $r = 6$ a budiž $0 < p < q$. Dokážeme, že $\hat{\omega} \sim \hat{z}(j)$ je optimální strategií právě když

$$\hat{z}(j) \leqq \min\{j, r - j\}, \quad \hat{z}(3) = 3.$$

Stanovme nejprve pravděpodobnost $\bar{u}(j)$ dosažení kapitálu r při strategii $\bar{\omega} \sim \bar{z}(j)$, definované tabulkou:

j	1	2	3	4	5
$\bar{z}(j)$	1	2	3	2	1

$\bar{u}(j)$ splňuje rovnice

(7) $$\bar{u}(0) = 0, \quad \bar{u}(1) = p\,\bar{u}(2), \quad \bar{u}(2) = p\,\bar{u}(4), \quad \bar{u}(3) = p,$$
$$\bar{u}(4) = p + q\,\bar{u}(2), \quad \bar{u}(5) = p + q\,\bar{u}(4), \quad \bar{u}(6) = \bar{u}(7) = \ldots = 1.$$

Řešením (7) dostáváme:

$\bar{u}(0)$	$\bar{u}(1)$	$\bar{u}(2)$	$\bar{u}(3)$	$\bar{u}(4)$	$\bar{u}(5)$	$\bar{u}(6)$
0	$\dfrac{p^3}{1 - pq}$	$\dfrac{p^2}{1 - pq}$	p	$\dfrac{p}{1 - pq}$	$p + \dfrac{pq}{1 - pq}$	1

Bezprostředním výpočtem se pak ověří, že platí

$$\bar{u}(j) = \max_{0<z\leqq j} [q\,\bar{u}(j - z) + p\,\bar{u}(j + z)], \quad j = 1, \ldots, 5.$$

Přitom rovnost

$$\bar{u}(j) = q\,\bar{u}(j - z) + p\,\bar{u}(j + z)$$

platí při $j \neq 3$ pro $z \leqq \min\{j, r - j\}$. Dále pro $z = 1, 2$

$$q\,\bar{u}(3 - z) + p\,\bar{u}(3 + z) = \frac{1 - q^2}{1 - pq}\bar{u}(3) < \bar{u}(3),$$

takže strategie $\omega \sim z(j)$, pro které $z(3) = 1$ nebo $z(3) = 2$ nejsou optimální.

9. PRŮMĚRNÝ VÝNOS ZE ŘETĚZCE S JEDINOU IZOLOVANOU TŘÍDOU REKURENTNÍCH STAVŮ

V tomto odstavci budeme předpokládat, že soustava přechodových matic (1 § 5) řízeného řetězce splňuje následující podmínku: Označme

$$\bar{p}_{ij} = \min_z p(i, j; z), \quad i, j \in I,$$

a budiž

$$\bar{\mathbf{P}} = \|\bar{p}_{ij}\|_{i,j=1}^r, \quad \bar{\mathbf{P}}^n = \|\bar{p}_{ij}^{(n)}\|_{i,j=1}^r.$$

Existují $j_0 \in I$, $d > 0$ a celé kladné ν tak, že

$$\bar{p}_{ij_0}^{(\nu)} \geqq d \quad \text{pro} \quad i \in I. \tag{1}$$

Tato podmínka má za následek, že pro každé homogenní markovské řízení $\omega \sim z(j)$ přechodová matice

$$\|p(i, j; z(i))\|_{i,j=1}^r \tag{2}$$

splňuje předpoklady věty 1 § 2.

Mějme dánu dále soustavu výnosných matic (2 § 5) a označme opětně $\{X_n, n = 0, 1, \ldots\}$ trajektorii řízeného řetězce. Průměrným výnosem na jednotku času se rozumí limita

$$\lim_{N \to \infty} N^{-1} \sum_{n=1}^N c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1})) \tag{3}$$

pokud existuje s pravděpodobností 1. Ve větě 1 bude dokázáno, že limita (3) existuje, je-li řízení homogenní markovské. Abychom mohli brát v úvahu i ta řízení, pro něž průměrný výnos neexistuje, nahradíme (3) ve větě 2 veličinou

$$\liminf_{N \to \infty} N^{-1} \sum_{n=1}^N c(X_{n-1}, X_n; z_{n-1}(X_0, \ldots, X_{n-1})).$$

Věta 1. *K libovolnému homogennímu markovskému řízení $\omega \sim z(j)$ existuje konstanta $\Theta(\omega)$ tak, že*

$$\Theta(\omega) = \lim_{N \to \infty} N^{-1} \sum_{n=1}^N c(X_{n-1}, X_n; z(X_{n-1})) \tag{4}$$

skoro všude vzhledem k P_j^ω, $j \in I$.

Důkaz věty 1 je obdobný důkazu věty 5 § 2. Z (1) vyplývá, že j_0 patří do jediné izolované třídy rekurentních stavů vzhledem k matici (2). Do této izolované třídy řetězec $\{X_n, n = 0, 1, \ldots\}$ s pravděpodobností 1 přejde ze stavů tranzientních, existují-li takové stavy (viz též větu 3 § 2). Budiž N_k, $k = 1, 2, \ldots$ doba k-tého přechodu řetězce do stavu j_0. Dle (11 § 2)

$$\sum_{n=N_i+1}^{N_{i+1}} c(X_{n-1}, X_n; z(X_{n-1}))\,, \quad i = 1, 2, \ldots,$$

tvoří posloupnost nezávislých stejně rozložených náhodných veličin s konečnou střední hodnotou jak plyne z věty 4 § 2. Ze silného zákona velkých čísel tedy dostáváme

$$\lim_{k\to\infty} N_k^{-1} \sum_{n=1}^{N_k} c(X_{n-1}, X_n; z(X_{n-1})) = \tag{5}$$

$$= \lim_{k\to\infty} (N_k/k)^{-1} \Big[\sum_{n=1}^{N_1} c(X_{n-1}, X_n; z(X_{n-1})) +$$

$$+ \sum_{i=1}^{k-1} \sum_{n=N_i+1}^{N_{i+1}} c(X_{n-1}, X_n; z(X_{n-1}))\Big]/k =$$

$$= E_{j_0}^{\omega} \sum_{n=1}^{N_1} c(X_{n-1}, X_n; z(X_{n-1}))/E_{j_0}^{\omega} N_1 = \Theta(\omega)$$

s pravděpodobností 1. Z (5) není obtížné nahlédnout, že platí (4). □

Průměrný výnos odpovídající homogennímu markovskému řízení ω budeme v tomto odstavci ztotožňovat s konstantou $\Theta(\omega)$ definovanou ve větě 1.

Pomocná věta 1. *Budiž a_n, $n = 1, 2, \ldots$, ohraničená posloupnost čísel a budiž*

$$\liminf_{N\to\infty} N^{-1} \sum_{k=1}^{N} a_k = A\,.$$

Potom je

$$\liminf_{\beta\to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} a_n \geqq A\,. \tag{6}$$

Důkaz. Zvolme libovolné $\varepsilon > 0$ a označme $S_n = \sum_{k=1}^{n} a_k$. Existuje n_0 tak, že je

$$S_n \geqq (A - \varepsilon)\, n \quad \text{pro} \quad n \geqq n_0\,.$$

Potom platí pro $0 < \beta < 1$

$$(1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} a_n = (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} (S_n - S_{n-1}) =$$

$$= (1 - \beta) \sum_{n=1}^{\infty} (\beta^{n-1} - \beta^n)\, S_n \geqq (1 - \beta) \sum_{n=1}^{n_0} (\beta^{n-1} - \beta^n)\,(S_n - nA + n\varepsilon) +$$

$$+ (1 - \beta)^2 \sum_{n=1}^{\infty} n \beta^{n-1} (A - \varepsilon)\,.$$

Poslední výraz je roven $A - \varepsilon$ a předposlední konverguje k nule pro $\beta \to 1$. Vidíme tedy, že je

$$\liminf_{\beta \to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} a_n \geqq A - \varepsilon .$$

Jelikož ε bylo libovolné, plyne odtud (6). □

Důsledek. Existuje-li

$$\lim_{N \to \infty} N^{-1} \sum_{k=1}^{N} a_k = A , \tag{7}$$

potom je

$$\lim_{\beta \to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} a_n = A .$$

Abychom to dokázali, je třeba nahlédnout, že za předpokladu (7) je také

$$\limsup_{\beta \to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} a_n \leqq A .$$

Můžeme postupovat buď obdobně jako při důkazu (6), anebo použít pomocnou větu 1 na posloupnost $-a_n$, $n = 1, 2, \ldots$.

V následující větě se, stručně řečeno, tvrdí, že lze nalézt optimální homogenní markovské řízení.

Věta 2. *Existuje homogenní markovské řízení $\hat{\omega} \sim \hat{z}(j)$ tak, že pro libovolné řízení*

$$\omega = \{z_n(j_0, \ldots, j_n),\ n = 0, 1, \ldots\}$$

platí

$$\Theta(\hat{\omega}) \geqq \liminf_{N \to \infty} N^{-1} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}) \tag{8}$$

skoro všude vzhledem P_j^{ω}, $j \in I$.

Důkaz: Náhodnou veličinu na pravé straně (8) označme $\varphi(\omega)$. V důkaze budeme používat očekávaný diskontovaný výnos

$$v(j, \beta; \bar{\omega}) = E_j^{\bar{\omega}} \sum_{n=1}^{\infty} \beta^{n-1}\, c(X_{n-1}, X_n; \bar{z}_{n-1}) ,$$

kde $\bar{\omega} = \{\bar{z}_n(j_0, \ldots, j_n),\ n = 0, 1, \ldots\}$ značí libovolné řízení řetězce. Budiž

$$\hat{v}(j, \beta) = \max_{\bar{\omega}} v(j, \beta; \bar{\omega}) , \quad j \in I .$$

Dle věty 1 § 7 existuje homogenní markovské řízení ω_β tak, že

$$\hat{v}(j, \beta) = v(j, \beta; \omega_\beta) , \quad j \in I .$$

Jelikož zobrazení I do J je konečně mnoho, lze nalézt $\hat{\omega} \sim \hat{z}(j)$ tak, že $\hat{\omega} = \omega_{\beta_m}$ pro vhodnou posloupnost β_m, $m = 1, 2, \ldots, \beta_m \to 1$. Tvrdíme, že $\hat{\omega}$ je hledané optimální řízení.

Ze (4) plyne

$$\Theta(\hat{\omega}) = \lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} E_j^{\hat{\omega}}\, c(X_{n-1}, X_n; \hat{z}(X_{n-1}))$$

a tedy, dle důsledku k pomocné větě 1,

$$\Theta(\hat{\omega}) = \lim_{\beta \to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} E_j^{\hat{\omega}}\, c(X_{n-1}, X_n; \hat{z}) = \tag{9}$$

$$= \lim_{\beta \to 1-} (1 - \beta)\, E_j^{\hat{\omega}} \sum_{n=1}^{\infty} \beta^{n-1}\, c(X_{n-1}, X_n; \hat{z}) = \lim_{\beta \to 1-} (1 - \beta)\, v(j, \beta; \hat{\omega})\,, \quad j \in I\,.$$

Budiž ω libovolné řízení. Dokážeme nejprve, že nemůže platit $E_j^{\omega}\, \varphi(\omega) > \Theta(\hat{\omega})$. Použitím Fatouova lemmatu a pomocné věty 1 dostáváme

$$\begin{aligned} E_j^{\omega}\, \varphi(\omega) &= E_j^{\omega} \liminf_{N \to \infty} N^{-1} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}) \leqq \\ &\leqq \liminf_{N \to \infty} N^{-1} E_j^{\omega} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}) \leqq \\ &\leqq \liminf_{\beta \to 1-} (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1}\, E_j^{\omega}\, c(X_{n-1}, X_n; z_{n-1}) = \\ &= \liminf_{\beta \to 1-} (1 - \beta)\, v(j, \beta; \omega)\,. \end{aligned} \tag{10}$$

Z (9), (10) vidíme, že z předpokladu $E_j^{\omega}\, \varphi(\omega) > \Theta(\hat{\omega})$ vyplývá

$$v(j, \beta; \omega) > v(j, \beta; \hat{\omega})$$

pro β dostatečně blízká 1. To je spor, neboť $\hat{\omega}$ maximalizuje diskontovaný výnos při $\beta_m \to 1$.

Důkaz věty ukončíme nepřímo. Předpokládejme, že je

$$P_j^{\omega}(\varphi(\omega) > \Theta(\hat{\omega})) > 0\,.$$

Sestrojíme řízení $\omega' = \{z'_n(j_0, \ldots, j_n),\, n = 0, 1, \ldots\}$, pro které musí platit

$$E_j^{\omega'}\, \varphi(\omega') = E_j^{\omega'} \liminf_{N \to \infty} N^{-1} \sum_{n=1}^{N} c(X_{n-1}, X_n; z'_{n-1}) > \Theta(\hat{\omega})\,. \tag{11}$$

Nerovnost (11) je ve sporu s tím, co jsme právě dokázali.

Zřejmě

$$E_j^{\omega}\{\varphi(\omega) \mid \varphi(\omega) > \Theta(\hat{\omega})\} > \Theta(\hat{\omega})\,.$$

Lze tedy nalézt číslo m a množinu A $(m + 1)$-tic čísel z I, tj. $A \subset I \times \ldots \times I$ tak, že platí

$$E_j^{\omega}\{\varphi(\omega) \mid [X_0, \ldots, X_m] \in A\} > \Theta(\hat{\omega}) . \tag{12}$$

Definujeme

$$z'_0(j_0) = z_0(j_0), \ldots, z'_m(j_0, \ldots, j_m) = z_m(j_0, \ldots, j_m)$$

a pro $n > m$

$$z'_n(j_0, \ldots, j_n) = z_n(j_0, \ldots, j_n), \quad \text{když} \quad [j_0, \ldots, j_m] \in A ,$$
$$z'_n(j_0, \ldots, j_n) = \hat{z}(j_n), \quad \text{když} \quad [j_0, \ldots, j_m] \notin A .$$

Z definice řízení ω' vyplývá, že je

$$\varphi(\omega') = \varphi(\omega) \quad \text{kdykoliv} \quad [X_0, \ldots, X_m] \in A ,$$
$$\varphi(\omega') = \Theta(\hat{\omega}) \quad \text{skoro všude na množině} \quad [X_0, \ldots, X_m] \notin A .$$

Odtud a z (12) snadno dostáváme $E_j^{\omega'} \varphi(\omega') > \Theta(\hat{\omega})$, tedy nerovnost (11). □

Poznámka. Poznamenejme, že z (10) vyplývá nejen $E_j^{\omega} \varphi(\omega) \leqq \Theta(\hat{\omega})$, ale také

$$\liminf_{N \to \infty} N^{-1} E_j^{\omega} \sum_{n=1}^{N} c(X_{n-1}, X_n; z_{n-1}) \leqq \Theta(\hat{\omega}) .$$

Z věty 2 vidíme, že při maximalizaci průměrného výnosu na jednotku času se můžeme omeziti na homogenní markovská řízení. Zvolme řízení $\omega \sim z(j)$ a označme pro stručnost

$$\|p(i, j; z(i))\|_{i,j=1}^{r} = \|p_{ij}\|_{i,j=1}^{r} = \mathbf{P} , \quad \mathbf{P}^n = \|p_{ij}^{(n)}\|_{i,j=1}^{r} , \tag{13}$$
$$\|c(i, j; z(i))\|_{i,j=1}^{r} = \|c_{ij}\|_{i,j=1}^{r} .$$

Dle (1) matice $\mathbf{P}$ splňuje podmínky věty 1 § 2. Zejména tedy existují limitní pravděpodobnosti $\pi_1, \ldots, \pi_r$. Dále je

$$\Theta(\omega) = \lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} E_j^{\omega} c(X_{n-1}, X_n; z(X_{n-1})) =$$
$$= \lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} \sum_{i} \sum_{k} p_{ji}^{(n-1)} p_{ik} c_{ik} = \sum_{i} \sum_{k} \pi_i p_{ik} c_{ik} .$$

$\Theta(\omega)$ představuje tedy očekávaný výnos z jednoho kroku ve stacionárním stavu řetězce.

Odvodíme nyní soustavu rovnic k určování $\Theta(\omega)$. Platí

$$E_j^{\omega} \sum_{n=1}^{N} c(X_{n-1}, X_n; z(X_{n-1})) - N\,\Theta(\omega) = \sum_{n=1}^{N} \sum_{i} \sum_{k} (p_{ji}^{(n-1)} - \pi_i)\, p_{ik} c_{ik} .$$

Dle věty 1 § 2 konverguje nekonečná řada

$$\sum_{n=1}^{\infty} \sum_i \sum_k (p_{ji}^{(n-1)} - \pi_i) \, p_{ik} c_{ik} = w_j \, .$$

Není obtížné ověřit, že je

$$\sum_k p_{jk} w_k = w_j - \sum_k p_{jk} c_{jk} + \sum_i \sum_k \pi_i p_{ik} c_{ik} = w_j - r(j; z(j)) + \Theta(\omega) \, , \tag{14}$$

kde $r(j, z)$ je definováno vztahem (1 § 6). Výsledek vyslovíme jako větu.

Věta 3. *$\Theta(\omega)$ je jediné číslo, k němuž lze nalézt $w_1, \ldots, w_r$ tak, že platí*

$$w_j + \Theta(\omega) = r(j, z(j)) + \sum_k p(j, k; z(j)) \, w_k \, , \quad j \in I \, . \tag{15}$$

Přitom $w_1, \ldots, w_r$ jsou určeny jednoznačně až na aditivní konstantu.

Důkaz. Vzhledem ke (14) je třeba dokázat pouze jednoznačnost. Předpokládejme, že Θ' a $w'_1, \ldots, w'_r$ splňují

$$w'_j + \Theta' = r(j, z(j)) + \sum_k p_{jk} w'_k \, , \quad j \in I \, .$$

Odečtením od (15) dostáváme

$$w_j - w'_j + \Theta(\omega) - \Theta' = \sum_k p_{jk}(w_k - w'_k) \, , \quad j \in I \, . \tag{16}$$

Budiž např. $\Theta(\omega) - \Theta' > 0$. Potom (16) má za následek

$$w_j - w'_j < \max_k (w_k - w'_k) \, , \quad j \in I \, ,$$

což je spor. Obdobně se nahlédne, že nemůže být $\Theta(\omega) - \Theta' < 0$. Tedy $\Theta(\omega) = \Theta'$. Máme proto

$$w_j - w'_j = \sum_k p_{jk}(w_k - w'_k) \, , \quad j \in I \, ,$$

odkud plyne

$$w_j - w'_j = \sum_k p_{jk}^{(n)}(w_k - w'_k) \quad \text{pro} \quad j \in I \, , \quad n = 1, 2, \ldots$$

Limitní přechod pro $n \to \infty$ dává

$$w_j - w'_j = \sum_k \pi_k(w_k - w'_k) = \text{const.} \, , \quad j \in I \, . \ \square$$

Zaveďme nyní maximální výnos

$$\hat{\Theta} = \max \{\Theta(\omega) : \omega \text{ homogenní markovské řízení}\} \, .$$

Řízení $\hat{\omega}$ nazveme optimálním, je-li $\hat{\Theta} = \Theta(\hat{\omega})$. Maximální výnos je charakterizován větou 4, v jejímž důkaze bude také popsána *Howardova iterační metoda* nacházení $\hat{\Theta}$ a odpovídajícího optimálního řízení.

Věta 4. *$\hat{\Theta}$ je jediné číslo, k němuž lze nalézt $\hat{w}_1, \ldots, \hat{w}_r$ tak, že platí*

$$\hat{w}_j + \hat{\Theta} = \max_z \left[r(j, z) + \sum_k p(j, k; z)\, \hat{w}_k \right], \quad j \in I. \tag{17}$$

Je-li $\hat{z}(j)$ hodnota parametru řízení, pro kterou výraz v hranaté závorce v (17) je maximální, potom $\hat{\omega} \sim \hat{z}(j)$ je optimální řízení.

Důkaz: Nejprve pomocí Howardova iteračního postupu dokážeme existenci řešení rovnice (17). Zvolíme libovolné $\omega^0 \sim z^0(j)$ a nacházíme postupně $\omega^1 \sim z^1(j)$, $\omega^2 \sim z^2(j), \ldots, \omega^n \sim z^n(j), \ldots$, takto: Řešíme soustavu rovnic

$$w_j^n + \Theta(\omega^n) = r(j, z^n(j)) + \sum_k p(j, k; z^n(j))\, w_k^n, \quad j \in I. \tag{18}$$

Přitom klademe $w_{j_0}^n = 0$, kde j_0 splňuje (1). Za této podmínky má (18) dle věty 3 jednoznačné řešení. $z^{n+1}(j)$ volíme rovným takové hodnotě z, pro níž výraz

$$r(j, z) + \sum_k p(j, k; z)\, w_k^n$$

je maximální.

Platí potom

$$\begin{aligned} w_j^n + \Theta(\omega^n) &\leqq r(j; z^{n+1}(j)) + \sum_k p(j, k; z^{n+1}(j))\, w_k^n, \\ w_j^{n+1} + \Theta(\omega^{n+1}) &= r(j, z^{n+1}(j)) + \sum_k p(j, k; z^{n+1}(j))\, w_k^{n+1}. \end{aligned} \tag{19}$$

Jako v důkaze předchozí věty se přesvědčíme, že nemůže být $\Theta(\omega^n) - \Theta(\omega^{n+1}) > 0$. Je tedy $\Theta(\omega^0) \leqq \Theta(\omega^1) \leqq \ldots \leqq \Theta(\omega^n), \ldots$

Jelikož homogenních markovských řízení je konečně mnoho, existuje l tak, že $\Theta(\omega^l) = \Theta(\omega^{l+1}) = \ldots$

Pro $n \geqq l$ dostáváme z (19)

$$w_j^n - w_j^{n+1} = \varrho_j + \sum_k p(j, k; z^{n+1}(j))\,(w_k^n - w_k^{n+1}), \quad j \in I, \tag{20}$$

kde $\varrho_j \leqq 0$, $j \in I$. Ukážeme, že z (20) plyne

$$w_j^n \leqq w_j^{n+1}, \quad j \in I. \tag{21}$$

Zaveďme opětně zkrácené označení

$$\| p(i, j; z^{n+1}(i)) \|_{i,j=1}^r = \| p_{ij} \|_{i,j=1}^r = \mathbf{P}, \quad \mathbf{P}^m = \| p_{ij}^{(m)} \|_{i,j=1}^r.$$

$\mathbf{P}$ vyhovuje větě 1 § 2. Budiž I_1 izolovaná třída rekurentních stavů, I' množina tranzientních stavů vzhledem k matici $\mathbf{P}$. O limitním rozložení $\pi_1, \ldots, \pi_r$ platí $\pi_i > 0$ pro $i \in I_1$, $\pi_i = 0$ pro $i \in I'$ (viz též věty 3 § 2, 5 § 2).

Vynásobením (20) π_j, sečtením pro $j \in I$ a použitím vztahu $\pi_k = \sum_j \pi_j p_{jk}$ dostáváme

$$\sum_j \pi_j(w_j^n - w_j^{n+1}) = \sum_j \pi_j \varrho_j + \sum_k \pi_k(w_k^n - w_k^{n+1}) .$$

Odtud vidíme, že musí být $\varrho_j = 0$ pro $j \in I_1$. Můžeme tedy z (20) psát

$$w_j^n - w_j^{n+1} = \sum_{k \in I_1} p_{jk}(w_k^n - w_k^{n+1}) , \quad j \in I_1 .$$

Z věty 3 pak vyplývá, že $w_j^n - w_j^{n+1}$ je rovno konstantě pro $j \in I_1$. Dále $j_0 \in I_1$ a $w_{j_0}^n = w_{j_0}^{n+1} = 0$. Odkud

$$w_j^n = w_j^{n+1} \quad \text{pro} \quad j \in I_1 . \tag{22}$$

Z (20) a (22) vidíme, že pro $j \in I'$ můžeme psát

$$w_j^n - w_j^{n+1} = \varrho_j + \sum_{k \in I'} p_{jk}(w_k^n - w_k^{n+1}) .$$

Odtud plyne

$$w_j^n - w_j^{n+1} = \varrho_j + \sum_{k \in I'} p_{jk}\varrho_k + \sum_{k \in I'} p_{jk}^{(2)}(w_k^n - w_k^{n+1})$$

a dále stejným postupem

$$w_j^n - w_j^{n+1} = \sum_{m=0}^{N} \sum_{k \in I'} p_{jk}^{(m)}\varrho_k + \sum_{k \in I'} p_{jk}^{(N+1)}(w_k^n - w_k^{n+1}) .$$

Vzhledem k větě 3 § 2

$$w_j^n - w_j^{n+1} = \sum_{m=0}^{\infty} \sum_{k \in I'} p_{jk}^{(m)}\varrho_k \leqq 0 \quad \text{pro} \quad j \in I' . \tag{23}$$

(23) spolu s (22) dává (21).

Z (21) a z konečnosti množiny homogenních markovských řízení vyplývá, že lze nalézt $m \geqq l$ tak, že

$$w_j^m = w_j^{m+1} , \quad j \in I .$$

Jelikož je také $\Theta(\omega^m) = \Theta(\omega^{m+1})$, můžeme psát, použijeme-li (18) a definice $z^{m+1}(j)$

$$w_j^m + \Theta(\omega^m) = w_j^{m+1} + \Theta(\omega^{m+1}) = r(j, z^{m+1}(j)) +$$
$$+ \sum_k p(j, k; z^{m+1}(j))\, w_k^{m+1} = r(j, z^{m+1}(j)) + \sum_k p(j, k; z^{m+1}(j))\, w_k^m =$$
$$= \max_z \left[r(j, z) + \sum_k p(j, k; z)\, w_k^m\right] , \quad j \in I .$$

Vidíme, že $\hat{\Theta} = \Theta(\omega^m)$ a $\hat{w}_j = w_j^m$, $j \in I$, je řešení rovnice (17).

Ověřme nyní, že $\hat{\Theta}$ je určeno rovnicí (17) jednoznačně. Předpokládejme, že existuje $\bar{\Theta} \neq \hat{\Theta}$, pro něž

$$\bar{w}_j + \bar{\Theta} = \max_z \left[r(j, z) + \sum_k p(j, k; z) \, \bar{w}_k \right], \quad j \in I . \tag{24}$$

Budiž např. $\hat{\Theta} - \bar{\Theta} > 0$ a budiž $\hat{\omega} \sim \hat{z}(j)$ řízení, definované ve znění věty. Potom platí

$$\hat{w}_j + \hat{\Theta} = r(j, \hat{z}(j)) + \sum_k p(j, k; \hat{z}(j)) \, \hat{w}_k , \quad j \in I , \tag{25}$$

a dle (24)

$$\bar{w}_j + \bar{\Theta} \geqq r(j, \hat{z}(j)) + \sum_k p(j, k; \hat{z}(j)) \, \bar{w}_k , \quad j \in I .$$

Odečtením dostáváme

$$\hat{w}_j - \bar{w}_j + \hat{\Theta} - \bar{\Theta} \leqq \sum_k p(j, k; \hat{z}(j)) \, (\bar{w}_k - \bar{w}_k) , \quad j \in I .$$

Odtud vyvodíme spor stejným způsobem jako ze (16). Obdobně nemůže být $\bar{\Theta} - \hat{\Theta} > 0$. Tedy $\bar{\Theta} = \hat{\Theta}$.

Zřejmě platí $\Theta(\omega) \leqq \hat{\Theta}$ pro libovolné $\omega \sim z(j)$, neboť v Howardově iteračním postupu jsme vycházeli z libovolného řízení a průměrné výnosy $\Theta(\omega^n)$ tvořily neklesající posloupnost. Zbývá ještě ověřiti, že $\hat{\omega} \sim \hat{z}(j)$ je optimální řízení. Věta 3 říká, že soustava rovnic

$$w_j + \Theta(\hat{\omega}) = r(j, \hat{z}(j)) + \sum_k p(j, k; \hat{z}(j)) \, w_k , \quad j \in I , \tag{26}$$

určuje $\Theta(\hat{\omega})$ jednoznačně. Srovnáním (25) a (26) obdržíme $\Theta(\hat{\omega}) = \hat{\Theta}$. □

Příklad 2c. Použijeme Howardova iteračního postupu k maximalizaci průměrného výnosu na jednotku času za podmínek příkladu 2. Volíme $z^0(j) = 1$. Klademe-li $w_1^0 = 0$, můžeme psát soustavu rovnic (18) ve tvaru

$$\begin{aligned}
\Theta(\omega^0) & & & = 50 \\
\Theta(\omega^0) & + 0{,}3 w_2^0 & & = 100{,}5 \\
\Theta(\omega^0) & - 0{,}8 w_2^0 & + 0{,}9 w_3^0 & = 121 .
\end{aligned}$$

Řešení této soustavy je

$$\Theta(\omega^0) = 50{,}00 \, ; \quad w_2^0 = 168{,}33 \, ; \quad w_3^0 = 228{,}52 .$$

Hodnoty veličiny

$$r(j, z) + \sum_k p(j, k; z) \, w_k^n \tag{27}$$

jsou pro $n = 0$ uvedeny v následující tabulce:

z \ j	1	2	3
1	50,00	218,33	278,52
2	180,02	239,19	291,35

Vidíme, že má být $z^1(j) = 2$. Je $w_1^1 = 0$ a ostatní veličiny se určí z rovnic

$$\Theta(\omega^1) - 0{,}5w_2^1 - 0{,}1w_3^1 = 73$$
$$\Theta(\omega^1) + 0{,}3w_2^1 - 0{,}1w_3^1 = 98{,}5$$
$$\Theta(\omega^1) - 0{,}9w_2^1 + 0{,}9w_3^1 = 117\,.$$

Řešením dostáváme

$$\Theta(\omega^1) = 94{,}61\,; \quad w_2^1 = 31{,}88\,; \quad w_3^1 = 56{,}75\,.$$

Veličina (27) pro $n = 1$ nabývá těchto hodnot:

z \ j	1	2	3
1	50,00	122,81	152,18
2	94,61	126,49	151,36

Odtud dostáváme $z^2(1) = 2$, $z^2(2) = 2$, $z^2(3) = 1$. Soustava (18) je nyní ($w_1^2 = 0$):

$$\Theta(\omega^2) - 0{,}5w_2^2 - 0{,}1w_3^2 = 73$$
$$\Theta(\omega^2) + 0{,}3w_2^2 - 0{,}1w_3^2 = 98{,}5$$
$$\Theta(\omega^2) - 0{,}8w_2^2 + 0{,}9w_3^2 = 121\,.$$

Odkud

$$\Theta(\omega^2) = 94{,}69\,; \quad w_2^2 = 31{,}88\,; \quad w_3^2 = 57{,}56\,.$$

ω^2 je optimální řízení, jak vyplývá z tabulky veličiny (27) pro $n = 2$:

z \ j	1	2	3
1	50,00	122,81	152,26
2	94,69	126,57	151,44

Vraťme se nyní k očekávanému výnosu do pevné doby

$$u_0^M(j;\,\omega) = E_j^\omega \sum_{n=1}^{M} c\big(X_{n-1}, X_n;\, z_{n-1}(X_0, \ldots, X_{n-1})\big)\,,$$

kterému byl věnován § 6. Chceme ukázat, že maximální očekávaný výnos

$$\hat{u}_0^M(j) = \max_{\omega} u(j; \omega)$$

je pro velká M přibližně roven $M\hat{\Theta}$, tj. M-násobku maximálního průměrného výnosu na jednotku času. K tomu budeme potřebovat dvě pomocné věty.

Pomocná věta 2. *Nechť j_0 a v splňují* (1). *Budiž*

$$N = \inf \{n : X_n = j_0\}, \tag{28}$$

doba prvého dosažení stavu j_0. Potom

$$E_j^\omega N \leqq v/d \tag{29}$$

pro $j \in I$ a libovolné řízení ω.

Důkaz. V důkaze věty 1 § 8 byla dokázána nerovnost (2 § 8), kde $\hat{E}$ je definováno vztahem (1 § 8). Tedy (29) platí pro všechna řízení, platí-li pro všechna homogenní markovská řízení.

Budiž $\omega \sim z(j)$. Dokážeme pro $n = 0, 1, \ldots$, nerovnost

$$P_j^\omega(N > nv) \leqq (1 - d)^n, \quad j \in I. \tag{30}$$

Pro $n = 0$ je (30) zřejmě splněno. Dále, z (30) plyne vzhledem k (1)

$$P_j^\omega(N > (n + 1)\, v) \leqq \sum_{k \neq j_0} p_{jk}^{(v)}\, P_k^\omega(N > nv) \leqq (1 - d)^{n+1}, \quad j \in I.$$

Zde jsme použili označení (13). Vidíme, že platnost (30) pro $n = 0, 1, \ldots$, vyplývá úplnou indukcí.

Můžeme proto psát

$$E_j^\omega N = \sum_{m=1}^{\infty} m(P_j^\omega(N > m - 1) - P_j^\omega(N > m)) = \sum_{m=0}^{\infty} P_j^\omega(N > m) \leqq$$

$$\leqq \sum_{n=0}^{\infty} v\, P_j^\omega(N > nv) \leqq v \sum_{n=0}^{\infty} (1 - d)^n = v/d. \quad \square$$

Pomocná věta 3. *Nechť $\bar{c}$ splňuje*

$$|c(i, j; z)| \leqq \bar{c} \quad pro \quad i, j \in I, \quad z \in J,$$

potom

$$|\hat{u}_0^M(j) - \hat{u}_0^M(k)| \leqq 4\bar{c}v/d \quad pro \quad j, k \in I, \quad M = 1, 2, \ldots \tag{31}$$

Důkaz. (31) vyplývá ze vztahu

$$|\hat{u}_0^M(j) - \hat{u}_0^M(j_0)| \leqq 2\bar{c}v/d, \quad j \in I.$$

Dokážeme např.

$$\hat{u}_0^M(j) \leqq \hat{u}_0^M(j_0) + 2\bar{c}v/d \,. \tag{32}$$

Zřejmě je

$$\hat{u}_0^{M-m}(j_0) - m\bar{c} \leqq \hat{u}_0^M(j_0)\,, \quad \text{kde} \quad m = 0, 1, \ldots, M-1\,.$$

Budiž $\omega = \{z_n(j_n), n = 0, 1, \ldots\}$ markovské řízení, pro které platí

$$u_0^M(j; \omega) = \hat{u}_0^M(j)\,, \quad j \in I\,.$$

Dle § 6

$$E_j^\omega\{\sum_{n=m+1}^{M} c(X_{n-1}, X_n; z_{n-1}(X_{n-1})) \mid X_m = j_0\} = \hat{u}_0^{M-m}(j_0)$$

a tedy také, s použitím označení (28),

$$\hat{u}_0^M(j) = u_0^M(j; \omega) \leqq \sum_{m=0}^{M-1} P_j^\omega(N = m)\,[m\bar{c} +$$

$$+ E_j^\omega\{\sum_{n=m+1}^{M} c(X_{n-1}, X_n; z_{n-1}(X_{n-1})) \mid X_m = j_0\}] + P_j^\omega(N > M-1)\,M\bar{c} =$$

$$= \sum_{m=0}^{M-1} P_j^\omega(N = m)\,[m\bar{c} + \hat{u}_0^{M-m}(j_0)] + P_j^\omega(N > M-1)\,M\bar{c} \leqq$$

$$\leqq \sum_{m=0}^{M-1} P_j^\omega(N = m)\,[2m\bar{c} + \hat{u}_0^M(j_0)] + P_j^\omega(N > M-1)\,[2M\bar{c} + \hat{u}_0^M(j_0)] \leqq$$

$$\leqq \hat{u}_0^M(j_0) + 2\bar{c}E_j^\omega N \leqq \hat{u}_0^M(j_0) + 2\bar{c}v/d$$

dle pomocné věty 2. Tím je nerovnost (32) dokázána. □

Věta 5.

$$\lim_{M\to\infty} M^{-1}\,\hat{u}_0^M(j) = \hat{\Theta} \quad pro \quad j \in I\,.$$

Důkaz. Budiž $\hat{\omega} \sim \hat{z}(j)$ řízení, pro něž $\hat{\Theta} = \Theta(\hat{\omega})$. Jelikož je

$$\hat{u}_0^M(j) \geqq u_0^M(j; \hat{\omega}) \quad \text{pro} \quad M = 1, 2, \ldots,$$

platí

$$\liminf_{M\to\infty} M^{-1}\,\hat{u}_0^M(j) \geqq \lim_{M\to\infty} M^{-1}\,u_0^M(j; \hat{\omega}) = \hat{\Theta}\,. \tag{33}$$

Připusťme, že je

$$\limsup_{M\to\infty} M^{-1}\,\hat{u}_0^M(j) > \hat{\Theta}\,. \tag{34}$$

Z pomocné věty 3 vyplývá

$$(35) \qquad \hat{u}_0^M(i) > \hat{u}_0^M(j) - 4\bar{c}v/d\,, \quad i \in I\,, \quad M = 1, 2, \ldots$$

Z (34), (35) snadno nahlédneme, že existuje L takové, že je

$$(36) \qquad L^{-1}\,\hat{u}_0^L(i) \geqq \Theta' > \hat{\Theta} \quad \text{pro} \quad i \in I$$

a pro vhodné Θ'.

Budiž $\omega = \{z_n(j_n), n = 0, 1, \ldots\}$ řízení, pro které

$$\hat{u}_0^L(i) = u_0^L(i; \omega)\,, \quad i \in I\,.$$

S pomocí ω utvořme „periodické" řízení

$$\omega' = \{z_0(j_0), z_1(j_1), \ldots, z_{L-1}(j_{L-1}), z_0(j_L), \ldots, z_{L-1}(j_{2L-1}), \ldots\}\,.$$

Potom v důsledku (36)

$$u_0^{nL}(j; \omega') = E_j^{\omega'} \sum_{k=0}^{n-1} u_0^L(X_{kL}; \omega) \geqq nL\Theta' \quad \text{pro} \quad n = 1, 2, \ldots$$

Odtud plyne vztah

$$\liminf_{M\to\infty} M^{-1}\, u_0^M(j; \omega') = \liminf_{n\to\infty} (nL)^{-1}\, u_0^{nL}(j; \omega') \geqq \Theta' > \hat{\Theta}\,.$$

To je spor s poznámkou k větě 2. Vidíme, že platí (33) a neplatí (34). Tedy

$$\lim_{M\to\infty} M^{-1}\, \hat{u}_0^M(j) = \hat{\Theta}\,. \quad \square$$

Předpokládejme, že vedle výnosové matice (2 § 5) je dána matice

$$(37) \qquad \|d(i, j; z)\|_{i,j=1}^r\,,$$

splňující

$$(38) \qquad r_0(j, z) = \sum_k p(j, k; z)\, d(j, k; z) > 0\,, \quad j \in I\,, \quad z \in J\,.$$

Matici (37) budeme nazývat *nákladovou maticí.*

Budiž $\omega \sim z(j)$ homogenní markovské řízení. Náhodná veličina

$$(39) \qquad \sum_{n=1}^{N} d(X_{n-1}, X_n; z(X_{n-1}))$$

bude označována jako náklad na řetězec nebo zobecněný čas po N krocích. Z věty 1 vyplývá, že existuje konstanta $\Theta_0(\omega)$ tak, že platí

$$\lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} d(X_{n-1}, X_n; z(X_{n-1})) = \Theta_0(\omega) > 0$$

skoro jistě. Poslední nerovnost je důsledkem (38). Je tedy také

$$\lim_{N\to\infty} \sum_{n=1}^{N} c(X_{n-1}, X_n; z(X_{n-1})) / \sum_{n=1}^{N} d(X_{n-1}, X_n; z(X_{n-1})) =$$

$$= \Theta(\omega)/\Theta_0(\omega) = \vartheta(\omega)$$

skoro všude vzhledem k P_j^ω, $j \in I$. $\vartheta(\omega)$ představuje průměrný výnos na jednotku nákladu.

$\Theta(\omega)$ je charakterizováno soustavou (15) a stejně tak pro $\Theta_0(\omega)$ platí rovnice

$$w_j^0 + \Theta_0(\omega) = r_0(j, z(j)) + \sum_k p(j, k; z(j))\, w_k^0\,, \quad j \in I\,. \tag{40}$$

Položme

$$\tilde{w}_j = w_j - \vartheta(\omega)\, w_j^0\,.$$

Odečteme-li (40) vynásobené $\vartheta(\omega)$ od (15) dostáváme

$$\tilde{w}_j + \vartheta(\omega)\, r_0(j, z(j)) = r(j, z(j)) + \sum_k p(j, k; z(j))\, \tilde{w}_k\,, \quad j \in I\,. \tag{41}$$

Soustava (41) určuje $\vartheta(\omega)$ jednoznačně.

Vzhledem k obdobě mezi (15) a (41) lze téměř beze změn v důkaze věty 4 odvodit, že maximální výnos na jednotku nákladu,

$$\hat{\vartheta} = \max_\omega \vartheta(\omega)\,,$$

je jednoznačně určen rovnicí

$$\hat{w}_j = \max_z \left[r(j, z) + \sum_k p(j, k; z)\, \hat{w}_k - \hat{\vartheta}\, r_0(j, z)\right], \quad j \in I\,. \tag{42}$$

K stanovení maximálního výnosu na jednotku nákladu lze použít Howardova iteračního postupu, kde ovšem na základě $\omega^n \sim z^n(j)$ volíme za $z^{n+1}(j)$ hodnotu parametru řízení, pro niž

$$r(j, z) + \sum_k p(j, k; z)\, w_k^n - \vartheta(\omega^n)\, r_0(j, z)$$

nabývá maximální hodnoty.

Vraťme se ještě k příkladu 3 o seřizování stroje. Trvání jednoho kroku řetězce v příkladu popsaného je buď δ nebo Δ v závislosti na tom, zda blok vyrábí nebo je seřizován. Skutečný čas tedy není při $\delta \neq \Delta$ úměrný počtu kroků a je v naší terminologii vlastně zobecněným časem (39), kde

$$d(i, j; z) = \Delta \quad \text{pro} \quad z = 1, 2, 3\,; \quad d(i, j; 4) = \delta, \quad i, j \in I\,.$$

To je třeba při stanovení průměrného výnosu na jednotku času brát v úvahu.

10. PRŮMĚRNÝ VÝNOS Z OBECNÉHO ŘETĚZCE

Podmínka (1 § 9) má za následek, že rekurentní stavy řízeného řetězce tvoří jedinou izolovanou třídu a jsou aperiodické. V tomto odstavci nebudeme klást na soustavy matic (1 § 5), (2 § 5), definující řízený řetězec, žádné omezení. $\{X_n, n = 0, 1, \ldots\}$ bude jako dříve značit trajektorii řetězce.

Budiž $\omega \sim z(j)$ homogenní markovské řízení. Zaveďme zkrácené označení

$$\|p(i, j; z(i))\|_{i,j=1}^r = \|p_{ij}\|_{i,j=1}^r = \boldsymbol{P}, \quad \boldsymbol{P}^n = \|p_{ij}^{(n)}\|_{i,j=1}^r,$$

$$r_j = r(j, z(j)) = \sum_k p(j, k; z(j))\, c(j, k; z(j)), \quad j \in I.$$

Průměrný výnos na jeden krok může být nyní náhodnou veličinou, jejíž hodnota je rovna průměrnému výnosu v té izolované třídě rekurentních stavů, kterou trajektorie řetězce dostihne. Za charakteristiku kvality řízení budeme proto považovat očekávaný průměrný výnos

(1)
$$\Theta_j(\omega) = \lim_{N \to \infty} N^{-1} E_j^\omega \sum_{n=1}^N c(X_{n-1}, X_n; z(X_{n-1})) =$$
$$= \lim_{N \to \infty} N^{-1} \sum_{n=1}^N \sum_k p_{jk}^{(n-1)} r_k = \sum_k \pi_{jk} r_k$$

dle věty 2 § 2 s použitím označení tam zavedeného.

Nechť matice $\boldsymbol{P}$ definuje izolované rekurentní třídy $I_1, \ldots, I_m$ a množinu tranzientních stavů I'. Poznamenejme, že $\pi_{ik} = \pi_{jk}$ pokud i a j patří téže izolované rekurentní třídě I_l. To plyne z věty 5 § 2, je-li $k \in I_l$. Pro $k \notin I_l$ jsou obě limitní pravděpodobnosti nulové. Zejména také $\Theta_j(\omega)$ je rovno konstantě v izolované rekurentní třídě.

Odvodíme obdobu rovnic (15 § 9). V každé izolované rekurentní třídě zvolíme jeden stav $j_i \in I_i$, $i = 1, \ldots, m$. Budiž

$$N^i = \inf\{n : n \neq 0, X_n = j_i\}, \quad i = 1, \ldots, m,$$

doba prvního přechodu do stavu j_i. Z důkazu věty 1 § 9 (vztah 5 § 9) vyplývá, že je

$$\Theta_j(\omega) = E_{j_i}^\omega \sum_{n=1}^{N^i} c(X_{n-1}, X_n; z(X_{n-1})) / E_{j_i}^\omega N^i \quad \text{pro} \quad j \in I_i.$$

Položme

$$w_j = E_j^\omega \sum_{n=1}^{N^i} c(X_{n-1}, X_n; z(X_{n-1})) - \Theta_j(\omega)\, E_j^\omega N^i, \quad j \in I_i, \quad i = 1, \ldots, m.$$

Jelikož $p_{jk} = 0$ pro $j \in I_i$, $k \notin I_i$, a $w_{j_i} = 0$, můžeme bez ohledu na definici w_k pro

$k \in I'$ psát

$$\sum_k p_{jk} w_k = \sum_{k \in I_i - j_i} p_{jk} w_k = E_j^{\omega}\{\sum_{n=2}^{N^i} c(X_{n-1}, X_n; z(X_{n-1})) - \Theta_j(\omega)(N^i - 1)\} =$$

$$= E_j^{\omega}\{\sum_{n=1}^{N^i} c(X_{n-1}, X_n; z(X_{n-1})) - \Theta_j(\omega) N^i\} - E_j^{\omega} c(X_0, X_1; z(X_0)) +$$

$$+ \Theta_j(\omega) = w_j - r_j + \Theta_j(\omega), \quad j \in I_i.$$

Vidíme, že pro $j \in I - I'$ platí

$$w_j + \Theta_j(\omega) = r_j + \sum_k p_{jk} w_k. \tag{2}$$

Pro $j \in I'$ se můžeme na (2) dívati jako na soustavu rovnic k určení w_j, tj.

$$w_j - \sum_{k \in I'} p_{jk} w_k = r_j - \Theta_j(\omega) + \sum_{k \in I - I'} p_{jk} w_k, \quad j \in I'. \tag{3}$$

Označíme-li pravou stranu (3) s_j dostáváme řešení (viz větu 3 § 2)

$$w_j = \sum_{n=0}^{\infty} \sum_{k \in I'} p_{jk}^{(n)} s_k, \quad j \in I'.$$

Zobecněním věty 3 § 9 je tato věta:

Věta 1. $\Theta_1(\omega), \ldots, \Theta_r(\omega)$ *jsou jediná čísla taková, že platí*

$$\Theta_j(\omega) = \sum_k p(j, k; z(j))\, \Theta_k(\omega), \quad j \in I, \tag{4}$$

a že lze najít $w_1, \ldots, w_r$, *pro něž*

$$w_j + \Theta_j(\omega) = r(j, z(j)) + \sum_k p(j, k; z(j))\, w_k, \quad j \in I. \tag{5}$$

Důkaz. Existenci čísel $w_1, \ldots, w_r$ splňujících (5) jsme právě dokázali. Z (1) a ze vztahu

$$\pi_{ji} = \sum_k p_{jk} \pi_{ki}, \quad i, j \in I,$$

vyplývá

$$\sum_k p_{jk}\, \Theta_k(\omega) = \sum_k \sum_i p_{jk} \pi_{ki} r_i = \sum_i \pi_{ji} r_i = \Theta_j(\omega), \quad j \in I,$$

což jest (4). Zbývá tedy ověřit jednoznačnost.

Nechť čísla $\Theta_1, \ldots, \Theta_r$ splňují (4), (5). Ze (4) dostáváme

$$\Theta_j = \sum_k p_{jk}^{(n)} \Theta_k, \quad \Theta_j = N^{-1} \sum_{n=1}^{N} \sum_k p_{jk}^{(n)} \Theta_k,$$

odkud limitním přechodem dle věty 2 § 2

(6) $$\Theta_j = \sum_k \pi_{jk} \Theta_k \,, \quad j \in I \,.$$

Probíhá-li j izolovanou rekurentní třídu, je π_{jk} konstantní při pevném k a tedy také Θ_j je konstantní. Z věty 3 § 9 potom plyne, že soustava rovnic

$$w_j + \Theta_j = r_j + \sum_{k \in I_l} p_{jk} w_k \,, \quad j \in I_l \,, \quad l = 1, \ldots, m \,,$$

určuje Θ_j jednoznačně. Jelikož $\pi_{jk} = 0$ pro $k \in I'$, platí dle (6)

(7) $$\Theta_j = \sum_{k \in I - I'} \pi_{jk} \Theta_k \,, \quad j \in I \,.$$

Vztahem (7) je Θ_j také pro $j \in I'$ vyjádřeno pomocí Θ_k, $k \in I - I'$, která jsou určena jednoznačně. □

Nyní popíšeme *obecný Howardův iterační postup* k nacházení maximálního výnosu a optimálního řízení. Zvolíme počáteční řízení $\omega^0 \sim z^0(j)$ libovolné. Dále sestrojujeme postupně pro $n = 0, 1, 2, \ldots$ řízení $\omega^{n+1} \sim z^{n+1}(j)$ na základě $\omega^n \sim \sim z^n(j)$ takto:

1. Vypočteme $\Theta_1(\omega^n), \ldots, \Theta_r(\omega^n)$ a $w_1^n, \ldots, w_r^n$ z rovnic

(8) $$\Theta_j(\omega^n) = \sum_k p(j, k; z^n(j)) \, \Theta_k(\omega^n) \,, \quad j \in I \,,$$

(9) $$w_j^n + \Theta_j(\omega^n) = r(j, z^n(j)) + \sum_k p(j, k; z^n(j)) \, w_k^n \,, \quad j \in I \,.$$

Přitom, je-li $n \neq 0$ v každé izolované rekurentní třídě vzhledem k matici

$$\| p(i, j, z^n(i)) \|_{i,j=1}^r$$

volíme-li jeden stav k, pro který klademe $w_k^n = w_k^{n-1}$.

(Při řešení postupujeme tak, že nejprve řešíme (9) pro každou izolovanou rekurentní třídu, při čemž $\Theta_j(\omega^n)$ je neznámá, nezávislá na j. Dosazením již vypočtených neznámých do (8), dostáváme soustavu rovnic pro $\Theta_j(\omega^n)$, $j \in I'$. Posléze dosazením všech vypočtěných proměnných do (9) dostaneme soustavu rovnic pro w_j^n, $j \in I'$).

2. Určíme $z^{n+1}(j)$ pro $j \in I$:

a) Nabývá-li *hlavní kritérium*, tj. výraz

(10) $$\sum_k p(j, k; z) \, \Theta_k(\omega^n)$$

maxima pouze pro jedinou hodnotu parametru z, položíme $z^{n+1}(j)$ rovným této hodnotě.

b) Nabývá-li (10) maxima pro více hodnot z položíme $z^{n+1}(j)$ rovným takové z nich, pro kterou nabývá maxima také *pomocné kritérium*, tj. výraz

$$r(j, z) + \sum_k p(j, k; z)\, w_k^n . \tag{11}$$

Platí-li pro nějaké n

$$\Theta_j(\omega^{n+1}) = \Theta_j(\omega^n)\,, \quad w_j^{n+1} = w_j^n\,, \quad j \in I\,, \tag{12}$$

iterační postup zastavíme. Potom ω^n je optimální řízení, neboli

$$\Theta_j(\omega^n) = \max_\omega \Theta_j(\omega)\,, \quad j \in I\,. \tag{13}$$

Ověříme, že (12) musí skutečně nastat. Zaveďme zkrácené označení

$$\|p(i, j; z^{n+1}(i))\|_{i,j=1}^r = \|p_{ij}\|_{i,j=1}^r = \mathbf{P}\,, \quad \mathbf{P}^n = \|p_{ij}^{(n)}\|_{ij,=1}^r\,,$$
$$r(j, z^{n+1}(j)) = r_j\,, \quad \Theta_j(\omega^{n+1}) - \Theta_j(\omega^n) = \Theta_j'\,, \quad j \in I\,.$$

Opět budeme předpokládat, že matice $\mathbf{P}$ definuje izolované rekurentní třídy $I_1, \ldots, I_m$ a množinu stavů tranzientních I'.

Nejprve dokážeme, že $\Theta_j(\omega^n)$, $n = 0, 1, \ldots$, tvoří neklesající posloupnost. Platí

$$\Theta_j(\omega^{n+1}) = \sum_k p_{jk}\, \Theta_k(\omega^{n+1})$$

a podle (8) a definice $z^{n+1}(j)$

$$\Theta_j(\omega^n) = \sum_k p_{jk}\, \Theta_k(\omega^n) - d_j\,, \tag{14}$$

kde $d_j \geqq 0$, $j \in I$, odkud

$$\Theta_j' = d_j + \sum_k p_{jk} \Theta_k'\,, \quad j \in I\,. \tag{15}$$

Dále s použitím označení (4 § 2)

$$\sum_j \pi_{ij} \Theta_j' = \sum_j \pi_{ij} d_j + \sum_k \pi_{ik} \Theta_k'\,, \quad i \in I\,. \tag{16}$$

Jelikož je $\pi_{ii} > 0$ pro $i \in I - I'$ vyplývá z (16) $d_i = 0$ pro $i \in I - I'$. To znamená, kdykoliv j je rekurentním stavem vzhledem k $\mathbf{P}$, potom $z^n(j)$ maximalizuje hlavní kritérium (10). Bylo tedy buď položeno $z^{n+1}(j) = z^n(j)$, je-li hodnota maximalizující (10) jediná, nebo bylo použito pomocného kritéria. V obou případech můžeme psát

$$w_j^n + \Theta_j(\omega^n) = r(j, z^n(j)) + \sum_k p(j, k; z^n(j))\, w_k^n = r_j + \sum_k p_{jk} w_k^n - e_j\,, \tag{17}$$

kde $e_j \geqq 0$. Odečteme-li (17) od

$$w_j^{n+1} + \Theta_j(\omega^{n+1}) = r_j + \sum_k p_{jk} w_k^{n+1},$$

dostáváme

$$w_j^{n+1} - w_j^n + \Theta'_j = e_j + \sum_{k \in I_l} p_{jk}(w_k^{n+1} - w_k^n), \quad j \in I_l, \quad l = 1, \dots, m. \tag{18}$$

Z (18) vyplývá

$$\Theta'_j = \sum_k \pi_{jk} e_k \geqq 0, \quad j \in I_l, \quad l = 1, \dots, m. \tag{19}$$

Přitom, není-li $e_j = 0$ pro $j \in I - I'$ není také $\Theta'_j = 0$.

Podle (15) můžeme psát

$$\Theta'_j - \sum_{k \in I'} p_{jk} \Theta'_k = d_j + \sum_{i \in I - I'} p_{ji} \Theta'_i, \quad j \in I',$$

odkud řešením

$$\Theta'_j = \sum_{n=0}^{\infty} \sum_{k \in I'} p_{jk}^{(n)} [d_k + \sum_{i \in I - I'} p_{ki} \Theta'_i] \geqq 0, \quad j \in I'. \tag{20}$$

Přitom není-li $d_j = 0$ pro $j \in I'$, není také $\Theta'_j = 0$. (19) a (20) dává

$$\Theta_j(\omega^{n+1}) \geqq \Theta_j(\omega^n), \quad j \in I.$$

Z konečnosti množiny homogenních markovských řízení usoudíme, že existuje q tak, že

$$\Theta_j(\omega^{n+1}) = \Theta_j(\omega^n) \quad \text{pro} \quad j \in I, \quad n = q, q+1, \dots \tag{21}$$

Platí-li (21), potom

$$w_j^{n+1} - w_j^n = \sum_{k \in I_l} p_{jk}(w_k^{n+1} - w_k^n), \quad j \in I_l, l = 1, \dots, m,$$

neboť v (18) je $\Theta'_j = 0 = e_j$. Tedy $w_j^{n+1} - w_j^n$ je rovno konstantě pro $j \in I_l$. Jelikož v každé izolované rekurentní třídě existuje stav k, pro který bylo zvoleno $w_k^{n+1} = w_k^n$, platí

$$w_j^{n+1} - w_j^n = 0, \quad j \in I - I'. \tag{22}$$

Z (21) dále vyplývá, že ve (14) $d_j = 0$ pro $j \in I'$. Je tedy také pro $j \in I'$ buď $z^{n+1}(j) = z^n(j)$ nebo bylo použito pomocného kritéria. Můžeme proto psát s ohledem na (22)

$$w_j^{n+1} - w_j^n = e_j + \sum_{k \in I'} p_{jk}(w_k^{n+1} - w_k^n), \quad j \in I',$$

kde $e_j \geqq 0$, $j \in I'$. Odkud

$$(23) \qquad w_j^{n+1} - w_j^n = \sum_{n=0}^{\infty} \sum_{k \in I'} p_{jk}^{(n)} e_k \geqq 0 \,, \quad j \in I' \,.$$

Z (22) a (23) vidíme, že je

$$(24) \qquad w_j^{n+1} \geqq w_j^n \,, \quad j \in I \,, \quad n = q, q + 1, \ldots$$

Poznamenejme, že ve (24) platí rovnost pro všechna j, kdykoliv $\omega^{n+1} = \omega^n$. Použijeme-li opět skutečnosti, že homogenních markovských řízení je konečně mnoho, usoudíme, že lze nalézti $n \geqq q$ tak, že platí (12).

Dále bude třeba ověřit, že při zastavení obecného iteračního postupu dostaneme optimální řízení. Při důkazu se používá úvah, obdobných těm, jimiž jsme dokazovali, že $\Theta_j(\omega^n)$, $n = 0, 1, \ldots$, tvoří neklesající posloupnost. Abychom se mohli na ně odvolávat, přizpůsobíme nová označení označením dříve zavedeným. Nechť platí (12). Chceme dokázati (13). Budiž $\omega \sim z(j)$ libovolné homogenní markovské řízení. Položme

$$\|p(i, j; z(i))\|_{i,j=1}^r = \|p_{ij}\|_{i,j=1}^r = \mathbf{P} \,, \quad \mathbf{P}^n = \|p_{ij}^{(n)}\|_{i,j=1}^r \,,$$

$$r(j, z(j)) = r_j \,, \quad \Theta_j(\omega^n) - \Theta_j(\omega) = \Theta'_j \,, \quad j \in I \,.$$

$I_1, \ldots, I_m$ buďtež rekurentní třídy vzhledem k přechodové matici $\mathbf{P}$, I' množina stavů tranzientních.

Podle (12) a definice $z^{n+1}(j)$

$$\Theta_j(\omega^n) = \Theta_j(\omega^{n+1}) = \sum_k p(j, k; z^{n+1}(j)) \, \Theta_k(\omega^{n+1}) =$$

$$= \sum_k p(j, k; z^{n+1}(j)) \, \Theta_k(\omega^n) = \max_z \sum_k p(j, k; z) \, \Theta_k(\omega^n) \,,$$

odkud

$$(25) \qquad \Theta_j(\omega^n) = d_j + \sum_k p_{jk} \, \Theta_k(\omega^n) \,, \quad j \in I \,,$$

kde $d_j \geqq 0$, $j \in I$. Odečteme-li (4) od (25), obdržíme

$$\Theta'_j = d_j + \sum_k p_{jk} \Theta'_k \,, \quad j \in I \,,$$

a stejně jako z (15) vyvodíme, že $d_j = 0$, je-li j rekurentním stavem vzhledem k matici $\mathbf{P}$. Vidíme, že $z(j)$ maximalizuje (10) a je proto buď $z(j) = z^{n+1}(j)$, nebo $z^{n+1}(j)$ bylo získáno maximalizací pomocného kritéria. Platí tedy

$$w_j^n + \Theta_j(\omega^n) = w_j^{n+1} + \Theta_j(\omega^{n+1}) = r(j, z^{n+1}(j)) +$$

$$+ \sum_k p(j, k; z^{n+1}(j)) \, w_k^{n+1} = r(j, z^{n+1}(j)) + \sum_k p(j, k; z^{n+1}(j)) \, w_k^n =$$

$$= e_j + r_j + \sum_k p_{jk} w_k^n \,, \quad j \in I - I' \,,$$

kde $e_j \geqq 0$, $j \in I - I'$. Odečteme-li odtud (5), získáme soustavy rovnic

$$(26) \qquad w_j^n - w_j + \Theta_j' = e_j + \sum_{k \in I_l} p_{jk}(w_k^n - w_k)\,, \quad j \in I_l\,, \quad l = 1, \ldots, m\,.$$

Z (26) stejně jako z (18) vyvodíme (19) a potom (20), což je v nynějším označení

$$\Theta_j(\omega^n) \geqq \Theta_j(\omega)\,, \quad j \in I\,.$$

Tím je (13) dokázáno.

11. ŘÍZENÉ MARKOVOVY PROCESY S KONEČNÝM POČTEM STAVŮ

Tento odstavec obsahuje přehled výsledků, které jsou pro řízené Markovovy procesy s konečným počtem stavů obdobou tvrzení, odvozených v §§ 6–9 pro řízené řetězce.

Řízený Markovův proces s množinou stavů $I = \{1, 2, \ldots, r\}$ je definován soustavou matic přechodových intenzit

$$(1) \qquad \|\mu(i, j; z)\|_{i,j=1}^r\,, \quad z \in J = \{1, 2, \ldots, s\}\,,$$

soustavou matic výnosů z přechodů

$$(2) \qquad \|\gamma(i, j; z)\|_{i,j=1}^r\,, \quad z \in J\,,$$

a soustavou vektorů výnosů na jednotku času

$$(3) \qquad \|\delta(i; z)\|_{i=1}^r\,, \quad z \in J\,.$$

Zde $\mu(i, j; z)$ je intenzita pravděpodobnosti přechodu ze stavu i do stavu j za předpokladu, že parametr řízení nabývá hodnoty z. $\gamma(i, j; z)$ je výnos, vyplývající z takového přechodu. Klademe $\gamma(i, i; z) = 0$, $i \in I$, $z \in J$. $\delta(i; z)$ značí výnos na jednotku času ve stavu i při parametru řízení z.

Homogenní markovské řízení ω je opět určeno funkcí $z(j)$ udávající hodnotu parametru řízení v závislosti na stavu soustavy. Je-li navíc tato hodnota parametru po částech konstantní, zleva spojitou funkcí času, hovoříme o (nehomogenním) markovském řízení $\omega \sim z(j, t)$. Jiná řízení uvažovati nebudeme. Každému řízení $\omega \sim z(j, t)$ odpovídá Markovův proces $\{Y_t, t \geqq 0\}$ s maticí intenzit přechodu

$$\|\mu(i, j; z(i, t))\|_{i,j=1}^r\,.$$

Výnos z procesu $\{Y_t, t \geqq 0\}$ v časovém intervalu $(t, T]$ má potom v souhlase s předchozími definicemi tvar

$$(4) \qquad C_\omega(t, T) = \int_t^T \delta(Y_s; z(Y_s, s))\,\mathrm{d}s + \sum_j \sum_k \int_t^T \gamma(j, k; z(j, s))\,\mathrm{d}\varphi_{jk}(s)\,,$$

kde $\varphi_{jk}(s)$ je počet přechodů ze stavu j do stavu k, které se uskutečnily v časovém intervalu $(0, s]$.

Příklad 5. *Volba disciplíny hromadné obsluhy.* Opravna, mající n opravářů přijímá vedle zakázek s obvyklou dodací lhůtou (typ 1) zakázky na rychloopravu (typ 2), na nichž se začíná pracovat okamžitě. Pravděpodobnost příchodu zájemce o rychloopravu v nekonečně malém časovém intervalu $(t, t + \mathrm{d}t)$ budiž rovna $\lambda\,\mathrm{d}t$ nezávisle na příchodech vně tohoto intervalu (Poissonův proces). Není-li zakázka typu 2 okamžitě vzata do práce, zákazník buď odejde nebo je jeho zakázka změněna na typ 1. Oprava zakázky typu i $(i = 1, 2)$ trvá náhodnou dobu, mající exponenciální rozložení se střední hodnotou ϑ_i^{-1}. To znamená, že oprava zakázky typu i, na které se v čase t pracuje, bude skončena v časovém intervalu $(t, t + \mathrm{d}t)$ s pravděpodobností $\vartheta_i\,\mathrm{d}t$.

Předpokládá se, že zakázek typu 1 má opravna dostatek, takže opraváři, který se uvolní, může být okamžitě taková zakázka přidělena. Příjem za jednotku pracovního času na zakázce typu i budiž d_i, $i = 1, 2$. Úloha má smysl pouze pro $d_1 < d_2$. Jsou-li všichni opraváři zaměstnáni, je možno zařadit rychloopravu pouze tak, že se práce na nějaké zakázce typu 1 přeruší. Tím vznikne dodatečný náklad výše c. Poznamenejme, že zakázku, na níž byla práce přerušena, není třeba od ostatních zakázek typu 1 odlišovat. Pracuje-li se na takové zakázce znovu v čase t, potom pravděpodobnost ukončení práce v intervalu $(t, t + \mathrm{d}t)$ je opět $\vartheta_1\,\mathrm{d}t$.

Stav opravy je určen dvojicí čísel (p, q), $0 \leqq p + q \leqq n$, kde první číslo značí počet zakázek typu 1, druhé číslo počet zakázek typu 2 v opravě. Naší úlohou je stanovit, při kterých stavech je po ukončení opravy výhodné zařadit okamžitě zakázku typu 1 a při kterých je výhodné čekat na zájemce o rychloopravu. Dále, v případě, že jsou všichni opraváři zaměstnáni, je-li výhodné nahradit rozpracovanou zakázku typu 1 eventuální zakázkou typu 2, kterou by bylo třeba odmítnout. Cílem je maximalizace průměrného příjmu na jednotku času.

Hodnoty parametru řízení, odpovídající jednotlivým rozhodnutím jsou určeny tabulkou 7. Intenzity přechodu $\mu(i, j; z)$ jsou uvedeny v tabulce 8. Intenzity v tabulce neobsažené jsou rovné nule.

Tabulka 7.

z	Po ukončení opravy typu 1	Po ukončení opravy typu 2	Při příchodu typu 2 není-li volno
1	zařadit typ 1	zařadit typ 1	odmítnout
2	čekat		
3	zařadit typ 1	čekat	
4	čekat		
5	zařadit typ 1	zařadit typ 1	zařadit namísto typu 1
6	čekat		
7	zařadit typ 1	čekat	
8	čekat		

Všimněme si na příklad intenzit v prvním řádku. Předpokládejme, že opravna je ve stavu (p, q), kde $p + q < n$ v čase t. Při $z = 1$ může opravna změnit stav pouze dvěma způsoby:

1. Je ukončena oprava nějaké zakázky typu 2. Tato zakázka je okamžitě nahrazena zakázkou typu 1, nastane přechod do $(p + 1, q - 1)$. Dle předpokladu pracuje opravna v čase t na q zakázkách typu 2. Každá z nich má pravděpodobnost ukončení v intervalu $(t, t + \mathrm{d}t)$ rovnou $\vartheta_2\,\mathrm{d}t$. Pravděpodobnost, že bude některá z nich ukončena v tomto intervalu, je $q\vartheta_2\,\mathrm{d}t$, což dává intenzitu ve třetím sloupci.

2. Přišel zájemce o rychloopravu. Nastane přechod do stavu $(p, q + 1)$. Pravděpodobnost

Tabulka 8.

$i = (p, q)$				$p + q < n$	$p + q = n$, $p \neq 0$
z \ j	$(p - 1, q)$	$(p, q - 1)$	$(p + 1, q - 1)$	$(p, q + 1)$	$(p - 1, q + 1)$
1	—	—	$q\vartheta_2$	λ	—
2	$p\vartheta_1$	—	$q\vartheta_2$	λ	—
3	—	$q\vartheta_2$	—	λ	—
4	$p\vartheta_1$	$q\vartheta_2$	—	λ	—
5	—	—	$q\vartheta_2$	λ	λ
6	$p\vartheta_1$	—	$q\vartheta_2$	λ	λ
7	—	$q\vartheta_2$	—	λ	λ
8	$p\vartheta_1$	$q\vartheta_2$	—	λ	λ

takového přechodu v časovém intervalu $(t, t + dt)$ je $\lambda\, dt$. Proto ve čtvrtém sloupci je intenzita rovna λ. Obdobnými úvahami se získají i ostatní intenzity v tabulce 8.

Soustavy (2) a (3), popisující výnos, jsou v uvažovaném příkladě tyto

$$\gamma(i, j; z) = 0 \quad \text{pro} \quad i = (p, q)\,, \quad j \neq (p - 1, q + 1)\,, \quad z = 1, \dots, 8\,,$$

$$\gamma(i, j; z) = -c \quad \text{pro} \quad i = (p, q)\,, \quad j = (p - 1, q + 1)\,, \quad z = 1, \dots, 8\,,$$

$$\delta(i; z) = pd_1 + qd_2 \quad \text{pro} \quad i = (p, q)\,, \quad z = 1, \dots, 8\,.$$

Nejprve si všimneme *celkového očekávaného výnosu* ze řízeného procesu, definovaného soustavami (1), (2), (3). Mějme řízení $\omega \sim z(j, t)$. Zaveďme označení

$$u_t^T(j; \omega) = E_i^\omega\{C_\omega(t, T) \mid X_t = j\}\,, \quad t \leqq T\,, \quad j \in I\,. \tag{5}$$

$C_\omega(t, T)$ je určeno vztahem (4). Zřejmě je (5) nezávislé na počátečním stavu i.

Najdeme soustavu rovnic pro $u_t^T(j; \omega)$. Uvážíme-li, že v nekonečně krátkém časovém intervalu $(t - dt, t]$ soustava ze stavu j s pravděpodobností $\mu(j, k; z(j, t - dt))\,.\,dt = \mu(j, k; z(j, t))\, dt$ přejde do stavu k a s pravděpodobností $1 + \mu(j, j; z(j, t))\, dt$ ve stavu j setrvá, dostáváme

$$\begin{aligned} u_{t-dt}^T(j; \omega) &= E_i^\omega\{C_\omega(t - dt, t) \mid X_{t-dt} = j\} + \\ &+ E_i^\omega\{C_\omega(t, T) \mid X_{t-dt} = j\} = \delta(j; z(j, t))\, dt + \\ &+ \sum_{k \neq j} \mu(j, k; z(j, t))\, dt[\gamma(j, k; z(j, t)) + u_t^T(k; \omega)] + \\ &+ (1 + \mu(j, j; z(j, t))\, dt)\, u_t^T(j; \omega)\,. \end{aligned}$$

Položíme-li

$$\varrho(j, z) = \delta(j; z) + \sum_k \mu(j, k; z)\, \gamma(j, k; z)\,, \quad j \in I\,, \tag{6}$$

můžeme psát

$$\frac{\mathrm{d}}{\mathrm{d}t} u_t^T(j;\omega) = -\left[\varrho(j, z(j,t)) + \sum_k \mu(j,k; z(j,t))\, u(k;\omega)\right], \tag{7}$$

$$u_T^T(j;\omega) = 0\,, \quad j \in I\,.$$

V bodech nespojitosti řízení $z(j, t)$ je třeba derivaci v (7) chápat jako derivaci zleva.

Maximální výnos

$$\hat{u}_t^T(j) = \sup_\omega u_t^T(j;\omega)\,, \quad j \in I\,,$$

vyhovuje Bellmanově rovnici

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{u}_t^T(j) = -\max_z \left[\varrho(j, z) + \sum_k \mu(j,k;z)\, \hat{u}_t^T(k)\right], \quad j \in I\,, \tag{8}$$

s počáteční podmínkou $\hat{u}_T^T(j) = 0,\, j \in I$. Heuristicky lze (8) odvodit ze vztahu

$$\hat{u}_{t-\mathrm{d}t}^T(j) = \max_z \left[\varrho(j,z)\,\mathrm{d}t + \sum_{k\neq j} \mu(j,k;z)\,\mathrm{d}t\, \hat{u}_t^T(k) + \right.$$
$$\left. + \left(1 + \mu(j,j;z)\,\mathrm{d}t\right) \hat{u}_t^T(j)\right].$$

Volíme-li totiž v časovém intervalu $(t - \mathrm{d}t, t]$ parametr řízení roven z, potom můžeme očekávati nejvýše výnos rovný výrazu v hranaté závorce. Je nutno tedy volit z tak, aby tento výnos byl maximální.

Laplaceova-Stieltjesova transformace celkového výnosu do doby T

$$\int_0^\infty \mathrm{e}^{-\lambda T}\, \mathrm{d}C_\omega(0, T)\,, \quad \lambda > 0\,,$$

představuje diskontovaný výnos. *Očekávaný diskontovaný výnos* je pak roven

$$v(j;\omega) = E_j^\omega \int_0^\infty \mathrm{e}^{-\lambda T}\, \mathrm{d}C_\omega(0, T)\,, \quad j \in I\,.$$

Při vyšetřování diskontovaného výnosu se omezíme na homogenní markovská řízení $\omega \sim z(j)$. Obdobnou úvahou jako jsme získali (7), odvodíme

$$v(j;\omega) = E_j^\omega \left\{\int_0^{\mathrm{d}T} \mathrm{e}^{-\lambda T}\, \mathrm{d}C_\omega(0,T) + \mathrm{e}^{-\lambda \mathrm{d}T} \int_0^\infty \mathrm{e}^{-\lambda S}\, \mathrm{d}C_\omega(0, \mathrm{d}T + S)\right\} =$$
$$= \delta(j; z(j))\,\mathrm{d}T + \sum_{k \neq j} \mu(j,k; z(j))\,\mathrm{d}T\left[\gamma(j,k;z(j)) + \right.$$
$$\left. + v(k;\omega)\right] + \left(1 + \mu(j,j;z(j))\,\mathrm{d}T\right)\left(1 - \lambda\,\mathrm{d}T\right) v(j;\omega) + \ldots.$$

Odtud

$$0 = \left[\delta(j; z(j)) + \sum_k \mu(j,k;z(j))\left(\gamma(j,k;z(j)) + v(k;\omega)\right) - \right.$$
$$\left. - \lambda\, v(j;\omega)\right]\mathrm{d}T + \ldots,$$

kde jsme vynechali členy řádu aspoň $(\mathrm{d}T)^2$. S použitím označení (6) můžeme psát

$$0 = \varrho(j, z(j)) + \sum_k \mu(j, k; z(j))\, v(k; \omega) - \lambda\, v(j; \omega)\,, \quad j \in I\,.$$

Bellmanova rovnice pro maximální výnos

$$\hat{v}(j) = \max_\omega v(j; \omega)\,, \quad j \in I\,,$$

je nyní

$$\lambda\, \hat{v}(j) = \max_z \left[\varrho(j, z) + \sum_k \mu(j, k; z)\, \hat{v}(k)\right], \quad j \in I\,.$$

Při odvození soustavy rovnic, určující *průměrný výnos na jednotku času*, využijeme Markovova řetězce, který je tvořen posloupností stavů procesu $\{Y_t, t \geqq 0\}$. Označme

$$(9) \qquad \mu(i; z) = -\mu(i, i; z)\,, \quad i \in I\,, \quad z \in J\,,$$

a budiž

$$\mu(i; z) > 0\,, \quad i \in I\,, \quad z \in J\,.$$

Položme

$$(10) \qquad p(i, i; z) = 0\,, \quad p(i, j; z) = \mu(i, j; z)/\mu(i; z)\,, \quad i, j \in I\,, \quad i \neq j\,.$$

$p(i, j; z)$ jsou pravděpodobnosti přechodu v okamžiku výstupu za předpokladu, že parametr řízení je roven z. Tvoří soustavu matic přechodu

$$(11) \qquad \|p(i, j; z)\|_{i,j=1}^{r}\,, \quad z \in J\,,$$

řízeného řetězce. Při libovolném homogenním markovském řízení $\omega \sim z(j)$ představuje posloupnost stavů procesu $\{Y_t, t \geqq 0\}$ Markovův řetězec $\{X_n, n = 0, 1, \ldots\}$ řízený týmž řízením ω. Přitom (11) je soustavou matic přechodu řetězce $\{X_n, n = 0, 1, \ldots\}$. Nechť tato soustava splňuje (1 § 9).

Výnos z procesu $\{Y_t, t \geqq 0\}$ za jeden přechod řetězce $\{X_n, n = 0, 1, \ldots\}$ je náhodná veličina. Není obtížné stanovit její průměrnou hodnotu. Přešel-li proces do stavu i, setrvává v něm náhodnou dobu se střední hodnotou $\mu(i; z)^{-1}$, je-li parametr řízení roven z. Ze setrvání ve stavu i vyplyne tedy průměrný výnos $\delta(i; z) : \mu(i; z)$. Dále při přechodu do stavu j vznikne výnos $\gamma(i, j; z)$. Průměrný výnos z přechodu ze stavu i do stavu j v řetězci $\{X_n, n = 0, 1, \ldots\}$ je tedy

$$(12) \qquad c(i, j; z) = \delta(i; z)/\mu(i; z) + \gamma(i, j; z)\,, \quad i, j \in I\,, \quad i \neq j\,, \quad z \in J\,.$$

Matice veličin (12), mající na diagonále nuly, tvoří soustavu výnosových matic řízeného řetězce.

Průměrná doba setrvání ve stavu i je $\mu(i; z)^{-1}$. Zavedeme tedy v řetězci zobecněný čas pomocí veličin

$$(13) \qquad d(i, j; z) = \mu(i; z)^{-1}\,, \quad i, j \in I\,, \quad z \in J\,.$$

Z definic (12), (13) vyplývá, že průměrný výnos ze řízeného procesu na jednotku času při řízení ω

$$\vartheta(\omega) = \lim_{T \to \infty} C_\omega(0, T)/T$$

($\vartheta(\omega)$ značí konstantu, k níž konverguje výraz napravo skoro všude), je roven průměrnému výnosu z řetězce na jednotku zobecněného času, tj.

$$\vartheta(\omega) = \lim_{N \to \infty} \sum_{n=1}^{N} c(X_{n-1}, X_n; z(X_{n-1}))/\sum_{n=1}^{N} d(X_{n-1}, X_n; z(X_{n-1})) .$$

Dosaďme (10), (12), (13) do soustavy (41 § 9) pro určení $\vartheta(\omega)$. Dostáváme

$$(14) \qquad \tilde{w}_j + \vartheta(\omega)/\mu(j; z(j)) = [\delta(j; z(j)) + \sum_{k \neq j} \mu(j, k; z(j)) (\gamma(j, k; z(j)) + \\ + \tilde{w}_k)]/\mu(j; z(j)) , \quad j \in I .$$

Odkud, použijeme-li (6), (9),

$$\vartheta(\omega) = \varrho(j, z(j)) + \sum_k \mu(j, k; z(j)) \, \tilde{w}_k , \quad j \in I .$$

Ze (42 § 9) pak vidíme, že maximální výnos

$$\hat{\vartheta} = \max_\omega \vartheta(\omega)$$

je určen rovnicí

$$(15) \qquad \hat{\vartheta} = \max_z [\varrho(j, z) + \sum_k \mu(j, k; z) \, \tilde{w}_k] , \quad j \in I .$$

Howardův iterační postup pro nacházení $\hat{\vartheta}$ spočívá v konstrukci posloupnosti $\omega^0, \omega^1, \ldots$ následujícím způsobem: Je-li $\omega^n \sim z(j)$, potom $\omega^{n+1} \sim z'(j)$ kde $z'(j)$ je hodnota parametru řízení, pro kterou výraz

$$\varrho(j, z) + \sum_k \mu(j, k; z) \, \tilde{w}_k$$

nabývá maxima. Když se dojde k řešení rovnice (15), výpočet se zastaví. Tento postup není zcela totožný s postupem, který bychom dostali aplikací iterační metody z § 9 na řízený řetězec, určený soustavami (11), (12), (13).

Při odvození soustavy (14) jsme nepoužili vlastnosti, že doby setrvání procesu ve stavech mají exponenciální rozložení. Lze tedy provedené úvahy rozšířit na obecnější *semi-markovské procesy*.

12. ŘÍZENÍ EXPONENCIÁLNÍHO RŮSTU

Galtonův-Watsonův větvicí proces s r typy jedinců je matematickým modelem populace, jejíž členové se vyvíjejí navzájem nezávisle a rozmnožují se tak, že každý z jedinců typu i bude s pravděpodobností $p(i; n_1, n_2, \ldots, n_r)$ nahrazen v další generaci n_1 jedinci typu 1, n_2 jedinci typu 2, ..., n_r jedinci typu r. Galtonův-Watsonův proces je proto popsán soustavou rozložení pravděpodobností

$$\{p(i; n_1, n_2, \ldots, n_r), n_j = 0, 1, \ldots\}, \quad i \in I = \{1, \ldots, r\}.$$

V řízeném Galtonově-Watsonově procesu tato soustava závisí ještě na parametru řízení. Jedná se tedy o soustavu

$$\{p(i; n_1, n_2, \ldots, n_r; z), n_j = 0, 1, \ldots\}, \quad i \in I, \quad z \in J = \{1, \ldots, s\}. \tag{1}$$

Množinu parametrů J budeme interpretovat jako množinu způsobů ovlivnění (ošetření), kterým mohou být příslušníci populace podrobeni. Ve svorkách v (1) je tedy rozložení pravděpodobností potomstva jedince typu i, který byl ovlivněn způsobem z. Řízením rozumíme funkci $\omega \sim z_n(i)$, kde $z_n(i)$ značí, jakému ovlivnění jsou podrobeni jedinci typu i v n-té generaci, $n = 0, 1, \ldots$ Časově homogenní řízení nezávisí na generaci, tj. $\omega \sim z(i)$. Dosazením $z(i)$ do (1) dostáváme soustavu rozložení pravděpodobností

$$\{p(i; n_1, n_2, \ldots, n_r; z(i)), n_j = 0, 1, \ldots\}, \quad i \in I,$$

Galtonova-Watsonova procesu v obvyklém smyslu.

Budiž dále

$$c(i; n_1, \ldots, n_r; z), \quad n_j = 0, 1, \ldots, \quad i \in I, \quad z \in J,$$

výnos z jedince typu i, jemuž se dostalo ovlivnění z a který měl n_1 přímých potomků typu 1, ..., n_r přímých potomků typu r. Předpokládáme, že platí

$$\sum_{n_1, \ldots, n_r} p(i; n_1, \ldots, n_r; z)\,|c(i; n_1, \ldots, n_r; z)| < \infty,$$

$$\sum_{n_1, \ldots, n_r} p(i; n_1, \ldots, n_r; z)\,(n_1 + \ldots + n_r) < \infty$$

pro $i \in I$, $z \in J$.

Označme $u^N(i; \omega)$, $i \in I$, $N = 1, 2, \ldots$, očekávaný výnos z jednoho jedince typu i a z $N - 1$ generací jeho potomků při řízení ω. Je-li např.

$$c(i; n_1, \ldots, n_r; z) = n_1 + \ldots + n_r - 1,$$

potom $u^N(i; \omega)$ je očekávaný počet jedinců v N-té generaci zmenšený o 1.

Není obtížné usoudit, že $u^N(i; \omega)$ splňuje rekurentní vztah

$$u^N(i; \omega) = \sum_{n_1, \ldots, n_r} p(i; n_1, \ldots, n_r; z_0(i))\,[c(i; n_1, \ldots, n_r; z_0(i)) + \\ + \sum_j n_j\, u^{N-1}(j; \omega')], \tag{2}$$

kde $\omega' \sim z_{n+1}(i)$. Zavedeme-li očekávaný výnos z jedince typu i

$$r(i, z) = \sum_{n_1, \ldots, n_r} p(i; n_1, \ldots, n_r; z)\, c(i; n_1, \ldots, n_r; z)$$

a očekávaný počet jeho potomků typu j

$$m(i, j; z) = \sum_{n_1, \ldots, n_r} p(i; n_1, \ldots, n_j, \ldots, n_r; z)\, n_j\,,$$

můžeme (2) přepsat jako

$$u^N(i; \omega) = r(i, z_0(i)) + \sum_j m(i, j; z_0(i))\, u^{N-1}(j; \omega')\,. \tag{3}$$

Vztah (3) implikuje řadu bezprostředních analogií s teorií řízených řetězců, přičemž $m(i, j; z)$ zastupuje pravděpodobnost přechodu $p(i, j; z)$. Na příklad maximální výnos

$$\hat{u}^N(i) = \max_\omega u^N(i; \omega)\,, \quad i \in I\,,$$

vyhovuje Bellmanově rovnici

$$\hat{u}^N(i) = \max_z \{r(i, z) + \sum_j m(i, j; z)\, \hat{u}^{N-1}(j)\}\,, \quad i \in I\,,$$

která je obdobou (4 § 6).

Jelikož Markovův řetězec s r stavy je speciálním případem Galtonova-Watsonova procesu s r typy jedinců, není tato souvislost náhodná. Větší obecnost nynější formulace spočívá v tom, že matice

$$\|m(i, j; z)\|_{i,j=1}^r\,, \quad z \in J\,, \tag{4}$$

se nepředpokládají stochastické.

Nechť nultá generace populace, rozmnožující se dle (1), se skládá z S_1^0 jedinců typu 1, S_2^0 jedinců typu $2, \ldots, S_r^0$ jedinců typu r. Budiž $S_1^n, S_2^n, \ldots, S_r^n$ očekávané složení n-té generace. Zřejmě při řízení $\omega \sim z_n(i)$ platí vztah

$$S_j^{n+1} = \sum_i S_i^n\, m(i, j; z_n(i))\,, \quad j \in I\,. \tag{5}$$

Celkový očekávaný výnos z N generací populace je potom dle (3)

$$\sum_i S_i^0\, u^N(i; \omega) = \sum_{n=0}^{N-1} \sum_j S_j^n\, r(j, z_n(j))\,. \tag{6}$$

Budiž $\omega \sim z(i)$ homogenní řízení. Zaveďme označení

$$\|m(i, j; z(i))\|_{i,j=1}^r = \|m_{ij}\|_{i,j=1}^r = \boldsymbol{M}_\omega\,, \quad \boldsymbol{M}_\omega^n = \|m_{ij}^{(n)}\|_{i,j=1}^r\,. \tag{7}$$

Z (5) vyplývá, že je

$$S_j^n = \sum_i S_i^0 m_{ij}^{(n)}\,, \quad j \in I\,, \tag{8}$$

a tedy dle (6)

$$\sum_i S_i^0\, u^N(i;\, \omega) = \sum_{n=0}^{N-1} \sum_i \sum_j S_i^0\, m_{ij}^{(n)}\, r(j,\, z(j))\,. \tag{9}$$

Vztah (5) představuje návrat k deterministickým modelům, z nichž jsme v § 1 vyšli. Bylo by také možné všechna tvrzení §§ 6–11 s výjimkou vět 1 § 9, 2 § 9 vyslovit (pro markovská řízení) v řeči deterministických modelů. V § 1 byl soubor nezávislých jedinců vyvíjejících se podle Markovova řetězce popsán jako jeden z možných předpokladů o detailní struktuře modelu. Důsledky teorie řízení Markovových řetězců pro deterministické modely jsou zřejmě nezávislé na tomto speciálním předpokladu o jejich struktuře. Podstatné je pouze, že veličiny $p(i, j;\, z)$ představují poměrné části jedinců typu i, kteří za jedno období přejdou v jedince typu j při ovlivnění z.

Asymptotické chování (9) pro $N \to \infty$ je určeno chováním mocnin matice $\boldsymbol{M}_\omega$. V tomto odstavci, využívajíce zejména obdoby s § 7, odvodíme metodu k nalezení takového ω, pro něž je růst populace nejrychlejší (případně zánik nejpomalejší). Nejprve připomeneme jednu důležitou vlastnost nezáporných matic.

Perronova věta. *Budiž $\boldsymbol{M} = \|m_{ij}\|_{i,j=1}^r$ nezáporná matice. Nechť existuje celé kladné v tak, že matice $\boldsymbol{M}^v$ má všechny prvky kladné. Potom má matice $\boldsymbol{M}$ jednoduché kladné charakteristické číslo $\varrho(\boldsymbol{M})$, které je v absolutní hodnotě větší než ostatní charakteristická čísla a platí*

$$\lim_{n\to\infty} m_{ij}^{(n)}\, \varrho(\boldsymbol{M})^{-n} = \mu_i v_j\,, \quad i, j \in I\,, \tag{10}$$

kde $\mu_1, \ldots, \mu_r,\ v_1, \ldots, v_r$ jsou kladná čísla, $\|m_{ij}^{(n)}\|_{i,j=1}^r = \boldsymbol{M}^n$.

Budeme předpokládat, že soustava (4) nezáporných matic splňuje následující podmínku: Budiž

$$\overline{m}_{ij} = \min_z m(i, j;\, z)\,, \quad \overline{\boldsymbol{M}} = \|\overline{m}_{ij}\|_{i,j=1}^r\,.$$

Existuje celé kladné v tak, že všechny prvky matice $\overline{\boldsymbol{M}}^v$ jsou kladné. Tato podmínka má za následek, že pro každé $\omega \sim z(j)$ matice $\boldsymbol{M}_\omega$ definovaná v (7) vyhovuje Perronově větě. Růst mocnin matice $\boldsymbol{M}_\omega$ je tedy určen maximálním charakteristickým číslem $\varrho(\boldsymbol{M}_\omega) = \varrho(\omega)$. Ze (8) a (10) dostáváme

$$\lim_{n\to\infty} S_j^n\, \varrho(\omega)^{-n} = v_j \sum_i S_i^0 \mu_i\,, \quad j \in I\,.$$

Označíme-li tedy

$$\hat{\varrho} = \max_\omega \varrho(\omega)\,,$$

potom $\hat{\varrho}$ udává maximální možný růst populace. Naše úloha spočívá proto v nalezení $\hat{\varrho}$ a odpovídajícího mu řízení.

Pomocná věta 1. *Buďtež $e_j \geqq 0$, $j \in I$, $\beta > 0$ reálná čísla, $e_1 + \ldots + e_r > 0$ a budiž $\omega \sim z(j)$ homogenní řízení. Je-li $\beta\, \varrho(\omega) < 1$, potom soustava rovnic*

$$v_j = e_j + \beta \sum_k m(j, k; z(j))\, v_k\,, \quad j \in I\,, \tag{11}$$

má kladné řešení. Je-li $\beta\, \varrho(\omega) \geqq 1$, pak neexistuje nezáporné řešení (11).

Důkaz. Použijeme označení (7). Je-li $\beta\, \varrho(\omega) < 1$, potom dle (10) řada

$$v_j = \sum_{n=0}^{\infty} \beta^n \sum_k m_{jk}^{(n)} e_k\,, \quad j \in I\,, \tag{12}$$

konverguje a je $v_j > 0$. Snadno se ověří, že (12) dává řešení (11).

Předpokládejme $\beta\, \varrho(\omega) \geqq 1$ a nechť $v_1 \geqq 0, \ldots, v_r \geqq 0$, splňují (11). Z (11) odvodíme postupně

$$v_j \geqq e_j\,, \quad v_j \geqq e_j + \beta \sum_k m_{jk} e_k, \ldots\,,$$

$$v_j \geqq \sum_{n=0}^{N} \beta^n \sum_k m_{jk}^{(n)} e_k\,, \quad j \in I\,.$$

Z (10) vyplývá, že poslední výraz konverguje k $+\infty$ pro $N \to \infty$, neboť $\beta\, \varrho(\omega) \geqq 1$. Vidíme, že předpoklad $v_j \geqq 0$, $j \in I$, vede ke sporu. □

Věta 1. *Buďtež $e_j \geqq 0$, $j \in I$, $\beta > 0$ reálná čísla, $e_1 + \ldots + e_r > 0$. Je-li $\beta\hat{\varrho} < 1$, potom rovnice*

$$\hat{v}_j = e_j + \max_z \beta \sum_k m(j, k; z)\, \hat{v}_k\,, \quad j \in I\,, \tag{13}$$

má kladné řešení. Je-li $\beta\hat{\varrho} \geqq 1$, pak neexistuje nezáporné řešení (13).

Důkaz. Nechť $\beta\hat{\varrho} < 1$. Zvolme libovolné $\omega^0 \sim z^0(j)$ a definujme postupně $\omega^1 \sim z^1(j)$, $\omega^2 \sim z^2(j), \ldots$ následujícím způsobem: Je-li $\omega^n \sim z^n(j)$ již definováno, najdeme řešení soustavy

$$v_j^n = e_j + \beta \sum_k m(j, k; z^n(j))\, v_k^n\,; \quad j \in I\,. \tag{14}$$

Řešení existuje a je kladné dle pomocné věty 1, neboť je $\beta\, \varrho(\omega^n) < 1$. Když

$$\sum_k m(j, k; z^n(j))\, v_k^n = \max_z \sum_k m(j, k; z)\, v_k^n\,, \quad j \in I\,, \tag{15}$$

postup se přeruší, protože potom $\hat{v}_j = v_j^n$, $j \in I$, je řešením (13), jak plyne snadno ze (14), (15). Neplatí-li (15), zvolíme $\omega^{n+1} \sim z^{n+1}(j)$ tak, aby bylo

$$\sum_k m(j, k; z^{n+1}(j))\, v_k^n = \max_z \sum_k m(j, k; z)\, v_k^n\,, \quad j \in I\,.$$

Ukážeme, že je

$$(16) \qquad v_j^{n+1} \succ v_j^n\,, \quad j \in I\,,$$

když (15) neplatí. Pišme

$$(17) \qquad v_j^{n+1} - v_j^n = \beta \sum_k [m(j, k; z^{n+1}(j)) - m(j, k; z^n(j))]\, v_k^n +$$
$$+ \beta \sum_k m(j, k; z^{n+1}(j))\, (v_k^{n+1} - v_k^n)\,, \quad j \in I\,.$$

Lze použíti pomocné věty 1, jelikož $\beta\, \varrho(\omega^{n+1}) < 1$. První výraz na pravé straně (17) je nezáporný a není roven nule pro všechna j, není-li splněno (15). Tedy pomocná věta 1 implikuje (16). Navzájem různých vektorů $(v_1^n, \ldots, v_r^n)$, $n = 0, 1, \ldots$, je konečně mnoho. Odtud plyne, že pro nějaké n (16) neplatí. Pak je splněno (15) a tedy (13) má kladné řešení.

Nechť $\beta\hat{\varrho} \geqq 1$. Předpokládejme, že (13) má nezáporné řešení. Budiž $\omega \sim z(j)$ takové, že $\varrho(\omega) = \hat{\varrho}$. Použijeme označení (7). Z (13) dostáváme

$$\hat{v}_j \geqq e_j + \beta \sum_k m_{jk} \hat{v}_k \geqq e_j$$

a stejně jako v důkazu pomocné věty 1

$$\hat{v}_j \geqq \sum_{n=0}^{N} \beta^n \sum_k m_{jk}^{(n)} e_k\,, \quad N = 0, 1, 2, \ldots,$$

pomocí úplné indukce. Pro $N \to \infty$ docházíme ke sporu, protože řada napravo diverguje vzhledem k (10). □

Důsledek. Budiž $1 \leqq \beta\hat{\varrho}$. Definujme postupně $\omega^0, \omega^1, \ldots$ a $(v_1^0, \ldots, v_r^0), (v_1^1, \ldots$ $\ldots, v_r^1), \ldots$ stejně jako v důkaze věty 1. Potom pro nějaké n řešení soustavy (14) buď neexistuje nebo má aspoň jednu složku zápornou.

Důkaz. Kdyby tvrzení důsledku nebylo správné, pak by popsaná rekurentní konstrukce vedla k nezápornému řešení rovnice (13), což je ve sporu s větou 1 pro $1 \leqq \beta\hat{\varrho}$. □

Věta 2. *Existuje $\delta > 0$ s touto vlastností: Kdykoliv pro $\omega \sim z(j)$ a $\beta \in (\hat{\varrho}^{-1} - \delta, \hat{\varrho}^{-1} + \delta)$ buď (11) nemá kladné řešení nebo*

$$(18) \qquad \sum_k m(j, k; z(j))\, v_k = \max_z \sum_k m(j, k; z)\, v_k\,, \quad j \in I\,,$$

kde $(v_1, \ldots, v_r)$ je kladné řešení (11), potom $\varrho(\omega) = \hat{\varrho}$.

Důkaz. Označme

$$\varrho' = \max \{\varrho(\omega') : \varrho(\omega') \neq \hat{\varrho}\} .$$

Důkaz je nutný pouze v případě, je-li množina na pravé straně neprázdná. Potom $\hat{\varrho}^{-1} < \varrho'^{-1}$. Položme $\varrho'^{-1} = \hat{\varrho} + \delta_0$ a budiž $\beta \in [\hat{\varrho}^{-1}, \hat{\varrho}^{-1} + \delta_0)$. Vzhledem k pomocné větě 1 soustava (11) nemá kladné řešení pouze tehdy, je-li $\varrho(\omega) = \hat{\varrho}$, neboť v opačném případě $\beta\,\varrho(\omega) < 1$. Dále (18) a (11) mají za následek, že (13) je splněno při $(\bar{v}_1, \ldots, \bar{v}_r) = (v_1, \ldots, v_r)$ což je ve sporu s větou 1. (18) proto nemůže nastat. Vidíme, že všechna $\beta \in [\hat{\varrho}^{-1}, \hat{\varrho}^{-1} + \delta_0)$ splňují podmínky věty 2.

Z (12) plyne, že pro libovolné $\omega \sim z(j)$, řešení $v_j = v_j(\beta, \omega)$, $j \in I$, soustavy (11) je rostoucí funkcí β pro $\beta\,\varrho(\omega) < 1$. Označíme-li tedy

$$\bar{v}_j = \max \{v_j(\hat{\varrho}^{-1}, \omega') : \omega' \sim z'(j), \varrho(\omega') \neq \hat{\varrho}\} , \quad j \in I ,$$

platí

(19) $$v_j(\beta, \omega') \leqq \bar{v}_j \quad \text{pro} \quad \beta \leqq \hat{\varrho}^{-1} , \quad \varrho(\omega') \neq \hat{\varrho} , \quad j \in I .$$

Když naopak $\varrho(\omega') = \hat{\varrho}$, potom

$$\lim_{\beta \to \hat{\varrho}^{-1}} v_j(\beta, \omega') = \infty , \quad j \in I .$$

V důsledku toho platí pro vhodné δ_1

(20) $$v_j(\beta, \omega') > \bar{v}_j \quad \text{pro} \quad \beta \in (\hat{\varrho}^{-1} - \delta_1, \hat{\varrho}^{-1}) , \quad \varrho(\omega') = \hat{\varrho} , \quad j \in I .$$

Budiž $\beta \in (\hat{\varrho}^{-1} - \delta_1, \hat{\varrho}^{-1})$. Dle pomocné věty 1 první z alternativ věty 2 nemůže nastat, neboť je $\beta\,\varrho(\omega') < 1$ pro každé ω'. Nechť $\omega \sim z(j)$ je takové, že platí (18). Mějme libovolné $\omega' \sim z'(j)$. Z (11) a (18)

(21) $$v_j = e_j + \beta \sum_k m(j, k; z(j))\, v_k \geqq e_j + \beta \sum_k m(j, k; z'(j))\, v_k , \quad j \in I .$$

Dále $(v'_1, \ldots, v'_r) = (v_1(\beta, \omega'), \ldots, v_r(\beta, \omega'))$ splňuje

(22) $$v'_j = e_j + \beta \sum_k m(j, k; z'(j))\, v'_k , \quad j \in I .$$

Odečteme-li (22) od (21), dostáváme

$$v_j - v'_j = f_j + \beta \sum_k m(j, k; z'(j))\, (v_k - v'_k) , \quad j \in I ,$$

kde $f_j \geqq 0$. Je-li $f_j = 0$, $j \in I$, pak je $v_j = v'_j$, $j \in I$. Jinak $v_j > v'_j$, $j \in I$, dle pomocné věty 1, protože je $\beta\,\varrho(\omega') < 1$. Tedy

(23) $$v_j = \max_{\omega'} v_j(\beta, \omega') , \quad j \in I .$$

Z (19) a (20) usoudíme snadno, že (23) je možné pouze tehdy, je-li $\varrho(\omega) = \hat{\varrho}$. Vidíme, že $\delta = \min \{\delta_0, \delta_1\}$ vyhovuje větě 2. □

Iterační postup. Nechť jsou dána nezáporná čísla $(e_1, \ldots, e_r)$, $e_1 + \ldots + e_r > 0$. Postup začíná volbou počáteční hodnoty $\beta_0 > 0$, počátečního kroku h, $0 < 2h \leqq \beta_0$ a počátečního řízení $\omega^{00} \sim z^{00}(j)$. Čísla β_0, h musí splňovat

$$\hat{\varrho}^{-1} \in [\beta_0 - 2h, \beta_0 + 2h]. \tag{24}$$

Sestrojují se dvě posloupnosti $\{\beta_m, m = 0, 1, \ldots\}$ a $\{\omega^{m0}, m = 0, 1, \ldots\}$. Platí $\beta_m \to \hat{\varrho}^{-1}$ a $\varrho(\omega^{m0}) = \hat{\varrho}$ pro všechna m dostatečně velká.

Abychom postup popsali, předpokládejme, že β_m a $\omega^{m0} \sim z^{m0}(j)$ byly již definovány. Když soustava

$$v_j^{m0} = e_j + \beta_m \sum_k m(j, k; z^{m0}(j))\, v_k^{m0}, \quad j \in I, \tag{25}$$

má kladné řešení, definuje se pomocná posloupnost $\omega^{m1} \sim z^{m1}(j)$, $\omega^{m2} \sim z^{m2}(j), \ldots$ stejným způsobem jako v důkaze věty 1, tj. takto: Má-li soustava

$$v_j^{mn} = e_j + \beta_m \sum_k m(j, k; z^{mn}(j))\, v_k^{mn}, \quad j \in I, \tag{26}$$

kladné řešení, volí se $\omega^{mn+1} \sim z^{mn+1}(j)$ tak, že platí

$$\sum_k m(j, k; z^{mn+1}(j))\, v_k^{mn} = \max_z \sum_k m(j, k; z)\, v_k^{mn}, \quad j \in I. \tag{27}$$

Mohou nastat dva případy. Buď pro nějaké n (26) nemá kladné řešení nebo pro nějaké n (27) je splněno při $\omega^{mn+1} = \omega^{mn}$. V obou případech klademe

$$\omega^{m+1\,0} = \omega^{mn}. \tag{28}$$

V prvém případě je β_{m+1} definováno vztahem

$$\beta_{m+1} = \beta_m - h \,.\, 2^{-m}, \tag{29}$$

ve druhém případě vztahem

$$\beta_{m+1} = \beta_m + h \,.\, 2^{-m}. \tag{30}$$

(28), (29) zůstávají v platnosti i tehdy, když (25) nemá kladné řešení ($n = 0$).

Z věty 1 a jejího důsledku plyne, že (29) používáme, kdykoliv je $\beta_m \geqq \hat{\varrho}^{-1}$, (30) používáme, kdykoliv je $\beta_m < \hat{\varrho}^{-1}$. Pomocí matematické indukce, vycházejíce z (24), snadno dokážeme

$$\hat{\varrho}^{-1} \in [\beta_m - h \,.\, 2^{-m+1}, \beta_m + h \,.\, 2^{-m+1}], \quad m = 0, 1, \ldots$$

Tedy $\beta_m \to \hat{\varrho}^{-1}$ a z věty 2 vyplývá, že je $\varrho(\omega^{m0}) = \hat{\varrho}$ pro všechna dostatečně velká m.

13. ŘÍZENÍ NA ZÁKLADĚ NEÚPLNÉ INFORMACE

V předchozích odstavcích jsme předpokládali, že po každém kroku známe stav, ve kterém se soustava nachází, tj. v čase n známe posloupnost $X_0, X_1, \ldots, X_n$. V řadě úloh stav soustavy nelze pozorovat přímo. Zmíníme se velmi stručně o nejjednodušším takovém modelu. Ukážeme, jak se přejde k řízenému řetězci s úplnou informací o stavu soustavy, charakterizovaném aposteriorním rozložením pravděpodobností.

Příklad 6. Vraťme se k příkladu 3. Učiňme o vzniku zmetků tuto hypotézu: Každý ze strojů bloku se může nacházet ve dvou stavech: dobrém (0) a poruchovém (1). Stroj nacházející se ve stavu 0 vyrobí vadný výrobek s pravděpodobností p_0, dobrý výrobek s pravděpodobností $q_0 = 1 - p_0$. Přitom s pravděpodobností p přejde do stavu 1, s pravděpodobností q ve stavu 0 setrvá. Ve stavu 1 stroj setrvává. Pravděpodobnost vyrobení zmetku je p_1, dobrého výrobku $q_1 = 1 - p_1$. Seřízením je stroj převeden do stavu 0.

Stav bloku je tedy určen dvojicí čísel (l, m), kde $l = 0, 1$ je stav stroje 1, $m = 0, 1$ stav stroje 2. Pravděpodobnosti přechodu v případě, že blok neseřizujeme, $p(i, j; 4)$, jsou uvedeny v tabulce 9.

Tabulka 9.

i \ j	0 0	0 1	1 0	1 1
0 0	q^2	pq	pq	p^2
0 1	0	q	0	p
1 0	0	0	q	p
1 1	0	0	0	1

Za předpokladu, že stav každého ze strojů bloku známe, liší se nová formulace jen velmi málo od příkladu 3. Připusťme však, že stav stroje nemůžeme zjistit, dokud nepřikročíme k jeho seřízení. To znamená, že v obdobích mezi seřizováním stroje je veškerá naše informace o jeho stavu obsažena ve znalosti výskytu dobrých výrobků a zmetků. Úlohou je nalézt optimální způsob seřizování bloku na základě této informace.

Budiž $\{X_n, n = 0, 1, \ldots\}$ řízený Markovův řetězec s množinou stavů $I = \{1, 2, \ldots, r\}$ a soustavou matic přechodu

$$\|p(i, j; z)\|_{i,j=1}^{r}, \quad z \in J = \{1, 2, \ldots, s\}. \tag{1}$$

(I když se obsah pojmu řízený řetězec v tomto odstavci neshoduje přesně s dříve zavedeným, není nebezpečí nedorozumění.) Předpokládejme, že realizaci řetězce $\{X_n, n = 0, 1, \ldots\}$ nemůžeme bezprostředně pozorovat. Pozorovat lze pouze posloupnost náhodných veličin $\{Y_n, n = 1, 2, \ldots\}$, nabývajících hodnot $\{1, 2, \ldots, h\}$, která souvisí s řetězcem následujícím způsobem: Je dána soustava rozložení pravděpodobností

$$[\pi(1; j, z), \pi(2; j, z), \ldots, \pi(h; j, z)], \quad j \in I, \quad z \in J. \tag{2}$$

$\pi(k; j, z)$ je pravděpodobnost, že $Y_n = k$ za podmínky $X_n = j$ při parametru řízení z. Přitom jakékoliv další podmínky týkající se veličin $\{X_m, m \neq n\}$ a $\{Y_m, m \neq n\}$

tuto pravděpodobnost nemění. Y_n představuje vlastně pozorování X_n zatížené chybou.

Výnos z řetězce je definován funkcí čtyř proměnných

$$c(i, j, k; z), \quad i, j \in I, \quad k = 1, \ldots, h, \quad z \in J.$$

$c(i, j, k; z)$ je výnos z přechodu ze stavu i do stavu j za podmínky, že byla pozorována hodnota k a parametr řízení byl roven z. *Řízením* rozumíme posloupnost funkcí $\omega = \{z_n(k_1, \ldots, k_n), n = 0, 1, \ldots\}$, kde z_0 je konstanta udávající počáteční hodnotu parametru řízení a $z_n(k_1, \ldots, k_n)$ je hodnota parametru řízení, kterou volíme za předpokladu, že výsledkem pozorování veličin $Y_1, \ldots, Y_n$ je posloupnost $k_1, \ldots, k_n$. Celkový výnos z N prvních kroků řetězce je tedy roven

$$\sum_{n=1}^{N} c(X_{n-1}, X_n, Y_n; z_{n-1}(Y_1, \ldots, Y_{n-1})).$$

Jelikož řetězec $\{X_n, n = 0, 1, \ldots\}$ nelze pozorovat, nemůžeme také předpokládat, že známe jeho počáteční hodnotu. Nechť však je dáno počáteční rozložení pravděpodobností

$$p_j = P(X_0 = j), \quad j \in I.$$

Při daném řízení $\omega = \{z_n(k_1, \ldots, k_n), n = 0, 1, \ldots\}$ a počátečním rozložení $\mathbf{p} = (p_1, \ldots, p_r)$ soustavy (1), (2) definují rozložení pravděpodobností posloupností $\{X_n, n = 0, 1, \ldots, Y_n, n = 1, 2, \ldots\}$. Z předchozího vyplývá, že toto rozložení je určeno podmíněnými pravděpodobnostmi

$$(3) \quad P(X_{n+1} = j, Y_{n+1} = k \mid X_0 = i_0, X_1 = i_1, Y_1 = k_1, \ldots, X_n = i, Y_n = k_n) = $$
$$= p(i, j; z_n(k_1, \ldots, k_n))\, \pi(k; j, z_n(k_1, \ldots, k_n)), \quad n = 0, 1, \ldots$$

Střední hodnotu a pravděpodobnosti při tomto rozložení budeme označovat $E_{\mathbf{p}}^{\omega}$, $P_{\mathbf{p}}^{\omega}(.)$.

Budeme se zabývat celkovým očekávaným výnosem do předem zvolené doby

$$(4) \qquad u^N(\mathbf{p}; \omega) = E_{\mathbf{p}}^{\omega} \sum_{n=1}^{N} c(X_{n-1}, X_n, Y_n; z_{n-1}(Y_1, \ldots, Y_{n-1})).$$

Zaveďme *aposteriorní rozložení pravděpodobností* veličiny X_n, vyplývající z pozorování $Y_1, \ldots, Y_n$,

$$(5) \qquad q_j^n = q_j^n(k_1, \ldots, k_n; \omega) = P_{\mathbf{p}}^{\omega}(X_n = j \mid Y_1 = k_1, \ldots, Y_n = k_n),$$
$$j \in I, \quad n = 1, 2, \ldots, \quad q_j^0 = p_j, \quad j \in I.$$

r-tici $(q_1^n, \ldots, q_r^n)$ budeme značit $\mathbf{q}^n = \mathbf{q}^n(k_1, \ldots, k_n; \omega)$. Ukážeme, že existuje řízení maximalizující (4), které závisí pouze na aposteriorních pravděpodobnostech. K obdobnému výsledku by bylo možno dospět i při ostatních definicích výnosu.

Nejprve vypíšeme Bayesovu formuli pro výpočet aposteriorního rozložení pravděpodobností. Máme dle (3)

$$P_{\mathbf{p}}^{\omega}(X_{n+1} = j, Y_{n+1} = k \mid Y_1 = k_1, \dots, Y_n = k_n) =$$
$$= \sum_i P_{\mathbf{p}}^{\omega}(X_n = i, X_{n+1} = j, Y_{n+1} = k \mid Y_1 = k_1, \dots, Y_n = k_n) =$$
$$= \sum_i P_{\mathbf{p}}^{\omega}(X_n = i \mid Y_1 = k_1, \dots, Y_n = k_n)\, P_{\mathbf{p}}^{\omega}(X_{n+1} = j, Y_{n+1} =$$
$$= k \mid Y_1 = k_1, \dots, Y_n = k_n, X_n = i) =$$
$$= \sum_i q_i^n(k_1, \dots, k_n; \omega)\, p(i, j; z_n(k_1, \dots, k_n))\, \pi(k; j, z_n(k_1, \dots, k_n)) .$$

Odtud odvodíme, že je jednak

$$(6) \qquad P_{\mathbf{p}}^{\omega}(Y_{n+1} = k \mid Y_1 = k_1, \dots, Y_n = k_n) = \sum_i \sum_j q_n^i\, p(i, j; z_n)\, \pi(k; j, z_n) ,$$

jednak při $Y_{n+1} = k$

$$(7) \qquad q_j^{n+1} = \left(\sum_i q_i^n\, p(i, j; z_n)\, \pi(k; j, z_n)\right) \left(\sum_i \sum_j q_i^n\, p(i, j; z_n)\, \pi(k; j, z_n)\right)^{-1} .$$

Všimněme si vztahů (6), (7). Předpokládejme, že pozorujeme posloupnost $\{Y_n, n = 1, 2, \dots\}$ a napočítáváme postupně aposteriorní rozložení. Takto získáme posloupnost r-rozměrných náhodných veličin $\{\mathbf{Q}^n, n = 0, 1, \dots\}$, kde

$$\mathbf{Q}^0 = \mathbf{p} , \quad \mathbf{Q}^n = (Q_1^n, \dots, Q_r^n) = \mathbf{q}^n(Y_1, \dots, Y_n; \omega) , \quad n = 1, 2, \dots$$

Je-li $\mathbf{Q}^n = \mathbf{q} = (q_1, \dots, q_r)$ a hodnota parametru řízení z, potom Q_j^{n+1} může dle (7) nabývat pouze hodnot

$$(8) \qquad \left(\sum_j q_i\, p(i, j; z)\, \pi(k; j, z)\right) \left(\sum_i \sum_j q_i\, p(i, j; z)\, \pi(k; j, z)\right)^{-1} =$$
$$= \varphi_j^k(\mathbf{q}; z) , \quad k = 1, \dots, h ,$$

každou dle (6) s pravděpodobností

$$(9) \qquad \sum_i \sum_j q_i\, p(i, j; z)\, \pi(k; j, z) ,$$

nezávisle na $\mathbf{Q}^0, \mathbf{Q}^1, \dots, \mathbf{Q}^{n-1}$.

Vidíme, že posloupnost $\{\mathbf{Q}^n, n = 0, 1, \dots\}$ tvoří řízený Markovův řetězec (s úplnou informací o stavu soustavy) s množinou stavů R, jejímiž prvky jsou r-tice nezáporných čísel o součtu 1 a s pravděpodobnostmi přechodu určenými vztahy (8), (9). Dokážeme, že lze nalézt řízení

$$(10) \qquad \omega = \{z_n(\mathbf{q}), n = 0, 1, \dots\} ,$$

pro něž (4) nabývá maximální hodnoty. V (10) je $z_n(\mathbf{q})$ hodnota parametru řízení za předpokladu, že aposteriorní rozložení na základě $Y_1, \ldots, Y_n$ je rovno $\boldsymbol{q}$. To znamená, že informace o průběhu posloupnosti $\{Y_n, n = 1, 2, \ldots\}$ postačující k maximalizaci (4) je obsažena v aposteriorním rozložení pravděpodobností.

Nejprve určíme maximum (4) při řízeních splňujících (10). Označme

$$r(\boldsymbol{q}, z) = \sum_i \sum_j \sum_k q_i \, p(i, j; z) \, \pi(k; j, z) \, c(i, j, k; z), \quad \boldsymbol{q} \in R,$$

očekávaný výnos za jeden přechod při hodnotě parametru řízení z, je-li počáteční rozložení pravděpodobností $\boldsymbol{q}$. Očekávaný výnos při platnosti (10) splňuje vztah

$$u^M(\boldsymbol{q}; \omega) = r(\boldsymbol{q}, z_0(\boldsymbol{q})) + \sum_i \sum_j \sum_k q_i \, p(i, j; z_0(\boldsymbol{q})) \, \pi(k; j, z_0(\boldsymbol{q})) \, . \tag{11}$$
$$. \, u^{M-1}(\Phi^k(\boldsymbol{q}; z_0(\boldsymbol{q})); \omega'), \ \boldsymbol{q} \in R,$$

kde

$$\Phi^k(\boldsymbol{q}; z) = (\varphi_1^k(\boldsymbol{q}; z), \ldots, \varphi_r^k(\boldsymbol{q}; z)), \quad k = 1, \ldots, h,$$
$$\omega' = \{z_{n+1}(\boldsymbol{q}), n = 0, 1, \ldots\}.$$

Na definici $\Phi^k(\boldsymbol{q}; z)$ v případě, kdy (8) nemá smysl, celkem nezáleží. (11) je bezprostředním důsledkem (8), (9). Optimální výnos stanovíme z Bellmanovy rovnice, která je obdobnou (4 § 6). Definujme postupně

$$\hat{u}^1(\boldsymbol{q}) = \max_z [r(\boldsymbol{q}, z)],$$
$$\hat{u}^{M+1}(\boldsymbol{q}) = \max_z [r(\boldsymbol{q}, z) + \tag{12}$$
$$+ \sum_i \sum_j \sum_k q_i \, p(i, j; z) \, \pi(k; j, z) \, \hat{u}^M(\Phi^k(\boldsymbol{q}; z))], \quad \boldsymbol{q} \in R.$$

Je-li $\hat{z}^M(\boldsymbol{q})$ ta hodnota parametru řízení, pro níž výraz v hranaté závorce ve (12) nabývá maxima, potom dle (11) řízení $\hat{\omega}_N = \{\hat{z}^{N-n-1}(\boldsymbol{q}), n = 0, 1, \ldots, N-1\}$ splňuje

$$u^N(\boldsymbol{q}; \hat{\omega}_N) = \hat{u}^N(\boldsymbol{q}), \ \boldsymbol{q} \in R.$$

Dokážeme, že pro libovolné $\omega = \{z_n(k_1, \ldots, k_n), n = 0, 1, \ldots\}$ a libovolné celé kladné N platí

$$\hat{u}^N(\boldsymbol{q}) \geqq u^N(\boldsymbol{q}; \omega), \quad \boldsymbol{q} \in R. \tag{13}$$

Budiž $\boldsymbol{q} \in R$. Označme

$$u_m^N(k_1, \ldots, k_m; \omega) = E_{\boldsymbol{q}}^{\omega}\{\sum_{n=m+1}^{N} c(X_{n-1}, X_n, Y_n; z_{n-1}) \mid Y_1 = k_1, \ldots, Y_m = k_m\},$$
$$m = 1, \ldots, N-1, \quad u_0^N(\omega) = u^N(\boldsymbol{q}; \omega).$$

Obdobou (2 § 6) je tento vztah plynoucí ze (3)

$$(14)\qquad u_{m-1}^N(k_1, \dots, k_{m-1}; \omega) = E_{\boldsymbol{q}}^{\omega}\{c(X_{m-1}, X_m, Y_m; z_{m-1}) \mid Y_1 =$$

$$= k_1, \dots, Y_{m-1} = k_{m-1}\} + E_{\boldsymbol{q}}^{\omega}\{\sum_{n=m+1}^{N} c(X_{n-1}, X_n, Y_n; z_{n-1}) \mid Y_1 =$$

$$= k_1, \dots, Y_{m-1} = k_{m-1}\} = r(\boldsymbol{q}^{m-1}(k_1, \dots, k_{m-1}; \omega), z_{m-1}(k_1, \dots, k_{m-1})) +$$

$$+ \sum_i \sum_j \sum_k q_i^{m-1}(k_1, \dots, k_{m-1}; \omega)\, p(i, j; z_{m-1}(k_1, \dots, k_{m-1}))\,.$$

$$.\, \pi(k; j, z_{m-1}(k_1, \dots, k_{m-1}))\, u_m^N(k_1, \dots, k_{m-1}, k; \omega)\,.$$

Postupně pro $m = N - 1, N - 2, \dots, 0$ odvodíme nerovnost

$$(15)\qquad u_m^N(k_1, \dots, k_m; \omega) \leqq \hat{u}^{N-m}(\boldsymbol{q}^m(k_1, \dots, k_m; \omega))\,.$$

Při $m = N - 1$ máme

$$u_{N-1}^N(k_1, \dots, k_{N-1}; \omega) = r(\boldsymbol{q}^{N-1}(k_1, \dots, k_{N-1}; \omega), z_{N-1}(k_1, \dots, k_{N-1})) \leqq$$

$$\leqq \max_z r(\boldsymbol{q}^{N-1}(k_1, \dots, k_{N-1}; \omega), z) = \hat{u}^1(\boldsymbol{q}^{N-1}(k_1, \dots, k_{N-1}; \omega))\,.$$

Platí-li (15), dostáváme ze (14)

$$u_{m-1}^N(k_1, \dots, k_{m-1}; \omega) \leqq \max_z \left[r(\boldsymbol{q}^{m-1}(k_1, \dots, k_{m-1}; \omega), z) +\right.$$

$$+ \sum_i \sum_j \sum_k q_i^{m-1}(k_1, \dots, k_{m-1}; \omega)\, p(i, j; z)\,.$$

$$.\, \pi(k; j, z)\, \hat{u}^{N-m}\, (\Phi^k(\boldsymbol{q}^{m-1}(k_1, \dots, k_{m-1}; \omega); z))\left.\right] = \hat{u}^{N-m+1}(\boldsymbol{q}^{m-1}(k_1, \dots, k_{m-1}; \omega))\,.$$

Tedy nerovnost (15) platí také pro $m = 0$. To však je nerovnost (13), kterou jsme chtěli dokázat.

Řízení řetězce s neúplnou informací o stavu soustavy bylo převedeno na řízení řetězce s úplnou informací a množinou stavů R. I když základní principy řízení, vyložené v §§ 6–10 pro konečné řetězce se vztahují na nový případ, nacházení optimálních řízení z Bellmanových rovnic se stává poměrně obtížným. Stejně tak odůvodnění limitních přechodů v §§ 9, 10 vyžaduje složitějšího matematického aparátu.

LITERATURA

Knihy

R. Bellman: Dynamic programming, Princeton 1957.

R. Bellman, S. Dreyfus: Applied dynamic programming, Princeton 1962.

W. Feller: An introduction to probability theory and its applications. New York, London 1957.

Ф. Р. Гантмахер: Теория матриц. Москва 1966.

T. E. Harris: The theory of branching processes. Berlin, Göttingen, Heidelberg 1963.

R. A. Howard: Dynamic programming and Markov processes. New York, London 1960.

Kai Lai Chung: Markov chains with stationary transition probabilities. Berlin, Göttingen, Heidelberg 1960.

J. G. Kemeny, J. L. Snell: Finite Markov chains. Princeton 1960.

Т. А. Сарымсаков: Основы теории процессов Маркова. Москва 1954.

Práce v časopisech a sbornících

(AMS — Annals of Mathematical Statistics,
ТВП — Теория вероятностей и ее применения).

K. J. Åström: Optimal control of Markov processes with incomplete state information. J. of Math. Analysis and Appl. *10* (1965), 174—204.

R. Bellman: A Markovian decision process. J. of Math. and Mech. *6* (1957), 679—684.

D. Blackwell: Discrete dynamic programming. AMS *33* (1962), 719—726.

Memoryless strategies in finite stage dynamic programming. AMS *35* (1964), 863—865.

Discounted dynamic programming. AMS *36* (1965), 226—235.

C. Derman: Stable sequential control rules and Markov chains. J. of Math. Analysis and Appl. *6* (1963), 257—265.

Markovian sequential decision processes. Proc. of Symp. on Applied Math. XVI. Providence 1964, 281—289.

On sequential control processes. AMS *35* (1964), 341—349.

Markovian sequential control processes with denumerable state space. J. of Math. Analysis and Appl. *10* (1965), 295—301.

Denumerable state Markovian decision processes. Average cost criterion. AMS *37* (1966), 1545—1553.

C. Derman, G. J. Lieberman: A Markovian decision model for a joint replacement and stocking problem. Management Sc. *13* (1967), 609—617.

C. Derman, R. E. Strauch: A note on memoryless rules for controlling sequential control processes. AMS *37* (1966), 276—278.

Е. Б. Дынкин: Управляемые случайные последовательности. ТВП *X* (1965), 3—18.

J. H. Eaton, L. A. Zadeh: Optimal pursuit strategies in discrete state probabilistic systems. Transactions of ASME Ser. D, J. of Basic Engineering *84* (1961), 23—29.

H. Hatori: On Markov chains with rewards. Kodai math. seminar reports *18* (1966), 184—192. On continuous-time Markov processes with rewards I. Tamtéž 212—218.

H. Hatori, T. Mori: On continuous-time Markov processes with rewards II. Tamtéž 353—356.

R. A. Howard: Research in semi-Markovian decision structures. J. of Operations Res. Soc. of Japan 6 (1964), 163—199.

System analysis of semi-Markov processes. IEEE Transactions, Milit. Electron. *8* (1964), 114—124.

Dynamic inference. Operations Res. *13* (1965), 712—733.

M. Iosifescu, P. Mandl: Application des systèmes à liaisons complètes à un problème de réglage. Rev. roum. Math. pures et appl. *XI* (1966), 533—539.

W. S. Jewell: Markov renewal programming I, II. Operations Res. *11* (1963), 938—971.

Z. Koutský: O jedné úloze optimálního rozhodování v markovských procesech. Ekon. matem. obzor *1* (1965), 370—382.

Н. В. Крылов: Построение оптимальной стратегии для конечной управляемой цепи. ТВП *X* (1965), 51—60.

A. Maitra: Dynamic programming for countable state systems. Sankhyā *A 27* (1965), 241—248.

P. Mandl: An iterative method for maximizing the characteristic root of positive matrices. Rev. roum. Math. pures et appl. *XII* (1967), 1305—1310.

И. В. Романовский: Существование оптимального стационарного управления в марковском процессе решения. ТВП *X* (1965), 130—133.

В. В. Рыков: Управляемые марковские процессы с конечными пространствами состояний и управлений. ТВП *XI* (1966), 343—351.

R. E. Strauch: Negative dynamic programming. AMS *37* (1966), 871—890.

А. Н. Ширяев: К теории решающих функций и управлению процессом наблюдения по неполным данным. Transactions of the Third Prague Conference on Information Theory etc., Prague 1964, 657—682.

Последовательный анализ и управляемые случайные процессы (дискретное время). Кибернетика (1965), 1—24.

Некоторые новые результаты в теории управляемых случайных процессов. Transactions of the Fourth Prague Conference on Information Theory etc., Prague 1967, 131—203.

О. В. Висков, А. Н. Ширяев: Об управлениях приводящих к оптимальным стационарным режимам. Труды матем. инст. им. В. А. Стеклова LXXI, Москва 1964, 35—45.

P. Wolfe, G. B. Dantzig: Linear programming in a Markov chain. Operations Res. *10* (1962), 702—710.