STUDIA UNIVERSITATIS MOLDAVIAE, 2015, nr.2(82)
Seria "Ştiinţe exacte şi economice" ISSN 1857-2073 ISSN online 2345-1033 p.31-40

ALGORITM PENTRU DETERMINAREA STRATEGIILOR OPTIME STAȚIONARE ÎN PROBLEMELE STOCASTICE DE CONTROL OPTIMAL DISCRET PE REȚELE DECIZIONALE CU MULTIPLE CLASE RECURENTE

Maria CAPCELEA, Titu CAPCELEA

Universitatea de Stat din Moldova

Este elaborat şi argumentat teoretic un algoritm eficient pentru determinarea strategiilor optime staționare în problemele stocastice de control optimal discret cu perioada de dirijare infinită, definite pe rețele decizionale cu multiple clase recurente, în care este aplicat criteriul de optimizare a combinației convexe a costurilor medii în clasele recurente. Sunt examinate probleme în care costurile de tranziție între stările sistemului dinamic şi probabilitățile de tranziție, definite în stările necontrolabile, sunt constante independente de timp. Algoritmul elaborat este bazat pe modelul de programare liniară pentru determinarea strategiilor optime în problemele de control definite pe rețele decizionale perfecte [3,4].

Cuvinte-cheie: *procese discrete, problemă stocastică de control optimal discret, rețele decizionale cu multiple clase recurente, strategii staționare, metoda programării liniare, algoritm polinomial.*

AN ALGORITHM FOR DETERMINING STATIONARY OPTIMAL STRATEGIES FOR STOCHASTIC DISCRETE OPTIMAL CONTROL PROBLEMS DEFINED ON NETWORKS WITH MULTIPLE RECURRENT CLASSES

An efficient algorithm for determining optimal stationary strategies for the stochastic discrete optimal control problems with infinite time horizon is developed and theoretically justified. The problems are defined on decision networks with multiple recurrent classes. The average costs convex combination optimization criterion is applied. We examine problems in which the costs of transitions between the states of the dynamic system and transition probabilities, defined on the uncontrollable states, are constants independent on time. The algorithm is based on the linear programming model developed for determining optimal strategies in control problems defined on perfect decision networks [3,4].

Keywords: *discrete processes, stochastic discrete optimal control problem, multichain networks, stationary strategies, linear programming approach, polynomial time algorithm.*

1. Formularea problemei. Fie sistemul dinamic L cu mulțimea finită de stări X, $X=\{x_1,...,x_n\}$. La fiecare moment $t=0, 1, 2, ...$, starea sistemului este $x(t)\in X$. Mulțimea X este divizată în două submulțimi disjuncte X_C şi X_N $(X=X_C\cup X_N,\ X_C\cap X_N=\varnothing)$, unde X_C este mulțimea de stări $x\in X$, în care tranziția sistemului în următoarea stare y este fixată de către decident la fiecare moment t, iar X_N este mulțimea de stări $x\in X$ din care sistemul trece în următoarea stare y în mod aleator, conform unei repartiții de probabilitate date. Astfel, pentru fiecare $x\in X_N$ pe mulțimea admisibilă de tranziții (x,y) între stări sunt date probabilitățile $p_{x,y}$ de trecere a sistemului din starea x în y, care satisfac relațiile:

$$p_{x,y}\geq 0,\ \sum_{y\in X(x)} p_{x,y}=1,\ \forall x\in X_N,\ X(x)=\{y\in X\ |\ \textit{există trecere din } x \textit{ în } y\}.$$

Dacă între stările $x\in X_N$ şi $y\in X$ nu există trecere directă, atunci se consideră $p_{x,y}=0$.

Dinamica sistemului L este descrisă folosind graful orientat al tranzițiilor de stare $G=(X, E)$, în care mulțimea de noduri X este mulțimea stărilor sistemului dinamic $X=\{x_1,...,x_n\}$, iar $E=\{(x,y)\ |\ x\in X, y\in X(x)\}$ este mulțimea arcelor ce corespunde tranzițiilor sistemului. Arcul arbitrar $(x,y)\in E$ exprimă posibilitatea de trecere a sistemului L din starea $x=x(t)$ în starea $y=x(t+1)$ la fiecare moment discret de timp t şi corespunde unei dirijări staționare a sistemului dinamic în starea $x\in X$. Fiecărui arc $(x,y)\in E$ i se asociază valoarea $c_{x,y} \in \mathbb{R}$, care exprimă costul de trecere a sistemului L din starea $x = x(t)$ în $y = x(t+1)$ pentru fiecare $t = 0, 1, 2, \ldots$.

Să considerăm rețeaua decizională $R = (G, X_C, X_N, c, p)$, determinată de graful tranzițiilor de stare G, mulțimile de stări X_C și X_N ($X_C \neq \varnothing, X_N \neq \varnothing$), funcția cost $c : E \to \mathbb{R}$, repartiția de probabilitate $p : E_N \to [0,1]$, definită pe mulțimea de arce $E_N = \{(x, y) \in E \mid x \in X_N, y \in X\}$ și starea inițială arbitrar fixată $x_{i_0} \in X$ a sistemului dinamic.

Evoluția sistemului dinamic stocastic L pe rețeaua decizională R este modelată folosind aparatul lanțurilor Markov în timp discret. Strategia staționară în rețeaua R se definește ca o aplicație $s : X_C \to X$, care asociază fiecărui vârf $x \in X_C$ un singur vârf $y = s(x) \in X(x)$. Aplicația s induce mulțimea valorilor booleene $s_{x,y}$, $x \in X_C, y \in X(x)$ ce satisfac relația $s_{x,y} = \begin{cases} 1 \ \textit{dacă} \ y = s(x) \\ 0 \ \textit{dacă} \ y \neq s(x) \end{cases}$. Atunci când este fixată strategia s în rețeaua R, pentru fiecare stare $x \in X_C$ tranziția $(x, s(x))$ din x în $y = s(x) \in X$ este realizată cu probabilitatea $p_{x,s(x)} = 1$; prin urmare, valorile $s_{x,y}$ $(x \in X_C, y \in X(x))$ le vom considera ca și probabilități de tranziție a sistemului dinamic din stările controlabile $x \in X_C$ în $y \in X(x)$.

Fixând arbitrar o strategie staționară s în rețeaua R, obținem lanțul Markov $\{X_k^s \mid k \geq 0\}$ al stărilor sistemului dinamic, căruia îi corespunde matricea probabilităților de tranziție $P^s = (p_{x,y}^s)$ cu elementele definite astfel:

$$p_{x,y}^s = \begin{cases} p_{x,y} \ \textit{dacă} \ x \in X_N \ \textit{si} \ y \in X(x) \\ 0 \ \textit{dacă} \ x \in X_N \ \textit{si} \ y \notin X(x) \\ 1 \ \textit{dacă} \ x \in X_C \ \textit{si} \ y = s(x) \\ 0 \ \textit{dacă} \ x \in X_C \ \textit{si} \ y \neq s(x) \end{cases}.$$

În graful G lanțului Markov menționat îi corespunde graful parțial $G_s = (X, E_s \cup E_N)$, unde $E_s = \{(x, y) \in E \mid x \in X_C, y = s(x)\}$.

Vom considera că orice strategie staționară s fixată în rețeaua R induce, prin intermediul matricei probabilităților de tranziție P^s, un lanț Markov cu multiple clase recurente, ceea ce înseamnă că graful asociat G_s conține un număr $r' \geq 2$ de componente tare conexe izolate $R^{s,j} = (X^{s,j}, E^{s,j})$, $j = 1, \ldots, r'$ ce corespund mulțimilor ireductibile de stări $X^{s,1}, \ldots, X^{s,r'}$.

Atunci când în rețeaua R este fixată arbitrar starea inițială $x_{i_0} \in X$ și strategia s, vom considera costul mediu per tranziție a sistemului dinamic ca fiind o variabilă aleatoare, indusă de repartițiile de probabilitate pe mulțimile admisibile de tranziții în stările necontrolabile și de dirijările fixate în stările controlabile. Pentru lanțul Markov $\{X_k^s \mid k \geq 0\}$ se poate defini un cost mediu (ce ia o valoare finită) de realizare a acestuia în modul următor:

Având fixată strategia s în rețeaua R, în baza formulei $\mu_x^s = \sum_{y \in X(x)} c_{x,y} s_{x,y}$ se determină costurile de tranziție a sistemului dinamic prin stările controlabile $x \in X_C$, precum și costurile imediate $\mu_z^s = \sum_{y \in X(z)} c_{z,y} p_{z,y}$

de tranziție prin stările necontrolabile $z \in X_N$. Astfel, se obține vectorul costurilor imediate $\mu^s = \left(\mu_i^s\right)_{i=1}^n = \left(\sum_{y \in X(x_i)} c_{x_i,y} p_{x_i,y}^s\right)_{i=1}^n$.

Pentru strategia fixată s se determină vectorul corespunzător $\omega^s = \left(\omega_i^s\right)_{i=1}^n$ al costurilor medii per tranziție a sistemului dinamic (cost mediu de realizare a lanțului Markov cu starea inițială fixată x_i) conform formulei $\omega^s = Q^s \mu^s$ [3,4], unde Q^s este matricea probabilităților limită asociată lanțului Markov, corespunzătoare strategiei staționare fixate s.

Pentru lanțul Markov ireductibil (stările mulțimii X în ansamblu constituie o singură clasă recurentă) rândurile matricei limită asociate Q^s coincide; prin urmare, coincid şi componentele vectorului ω^s. Pe o rețea decizională perfectă (rețeaua în care orice strategie fixată s în R induce un lanț Markov ireductibil) [3,4], pentru orice strategie fixată s şi starea inițială arbitrară a lanțului Markov, costurile medii per tranziție coincid. De acest lucru s-a ținut cont atunci când a fost elaborat modelul de programare liniară pentru determinarea strategiilor optime staționare în problema stocastică de control optimal discret, definită pe rețea decizională perfectă [3,4]. Însă, pentru lanțul Markov cu multiple clase recurente în matricea limită Q^s există rânduri distincte (sunt distincte rândurile cu indicii ce corespund stărilor care aparţin diferitelor clase recurente; la fel, sunt distincte rândurile ce corespund stărilor tranziente). În acest caz, vom avea costuri medii per tranziție diferite în dependență de starea inițială a sistemului. Ținând cont de aceasta, pentru problemele de control cu multiple clase recurente vom defini costul mediu per tranziție în rețeaua decizională, folosind conceptul de distribuție staționară a lanțului Markov. Distribuția staționară este vectorul probabilist $\pi^s \in \Pi := \left\{\pi^s = \left(\pi_1^s, \ldots, \pi_n^s\right) \middle| \pi_j^s \geq 0,\ j = 1, \ldots, n,\ \sum_{j=1}^n \pi_j^s = 1\right\}$, care este soluție a sistemului omogen de ecuații algebrice liniare $\pi = \pi P^s$, unde P^s este matricea probabilităților de tranziție ce corespunde strategiei fixate s [3,4]. Componenta π_j^s a distribuției staționare a lanțului Markov cu starea inițială $x_{i_0} \in X$ este probabilitatea ca după un număr mare de tranziții sistemul dinamic să se afle în starea x_j şi poate fi interpretată ca raportul numărului preconizat de vizitări ale stării x_j la numărul total de tranziții k ale sistemului, atunci când $k \to \infty$ [2].

Atunci când mulțimea de stări X este finită, $|X| = n$, există distribuție staționară a lanțului Markov dacă acesta conține clase recurente. Pentru lanțurile Markov ireductibile distribuția staționară este unică (pentru orice stare inițială $x_{i_0} \in X$ şi fiecare strategie fixată s sistemul de ecuații liniare $\pi = \pi P^s$ posedă soluție unică) [2,6]. În cazul lanțurilor Markov cu multiple clase recurente avem un număr infinit de distribuții staționare $\left\{\pi^s \in \Pi \middle| \pi^s = \pi^s P^s\right\}$, iar mulțimea distribuțiilor staționare reprezintă o mulțime poliedrală în $\mathbb{R}^n$ [2,5], vârfurile căreia corespund distribuțiilor staționare ale sublanțurilor Markov ireductibile (adică distribuțiilor în clasele recurente ale lanțului inițial).

Fie $\pi^{s,x_{i_0}}$ o distribuție staționară a lanțului Markov $\left\{X_k^s \middle| X_0^s = x_{i_0},\ k \geq 0\right\}$ (cu starea inițială dată $x_{i_0} \in X$) ce corespunde strategiei staționare s arbitrar fixate în R. Atunci, distribuția $\pi^{s,x_{i_0}}$ defineşte mulțimea stărilor recurente şi mulțimea stărilor tranziente ale lanțului Markov considerat.

Dacă $\mu^s = \left(\mu_i^s\right)_{i=1}^n$ este vectorul costurilor imediate (pentru strategia fixată s), $\mu_i^s = \sum_{y \in X(x_i)} c_{x_i,y} p_{x_i,y}^s$, atunci, pentru fiecare stare $x_i \in X$, cantitatea $\mu_i^s \pi_i^{s,x_{i_0}}$ reprezintă un cost mediu pentru tranzitul stării x_i dacă sistemul începe tranziţiile din x_{i_0}, iar cantitatea $\psi_{x_{i_0}}(s)=\mu^s\left(\pi^{s,x_{i_0}}\right)^T=\sum_{i=1}^{n}\mu_i^s\pi_i^{s,x_{i_0}}$ – costul mediu per tranziţie a stărilor sistemului dinamic cu starea iniţială x_{i_0} când perioada de dirijare a sistemului este infinită [6, p.172]. Valoarea costului $\psi_{x_{i_0}}(s)$ poate varia în dependenţă de starea iniţială x_{i_0} a sistemului.

Pentru fiecare clasă recurentă $X^{s,j},\ j=1,...,k'$ a lanţului Markov $\left\{X_k^s \,\middle|\, k\geq 0\right\}$ vom considera distribuţia staţionară $\pi_{X^j}^s\in\mathbb{R}^n$ (unică) în $X^{s,j}$, ale cărei componente nenule se determină ca soluţie unică $\tilde{\pi}_{X^j}^s=\left(\left(\tilde{\pi}_k^s\right)_{X^j}\right)_{k=1}^m\in\mathbb{R}^m$ a sistemului de ecuaţii liniare:

$$\begin{cases}\tilde{\pi}_{X^j}^s=\tilde{\pi}_{X^j}^s P_{X^j}^s\\ \sum_{k=1}^{m}\left(\tilde{\pi}_k^s\right)_{X^j}=1\end{cases},$$

în care $P_{X^j}^s$ este submatricea matricei P^s de dimensiune $m\times m$, elementele căreia sunt elemente ale matricei P^s de la intersecţia rândurilor şi coloanelor ce corespund indicilor stărilor clasei recurente $X^{s,j}$. Componentele vectorului $\pi_{X^j}^s\in\mathbb{R}^n$, $\left(\pi_j^s\right)_{X^j}\geq 0,\ j=1,...,n$, ce corespund stărilor din exteriorul clasei recurente, se consideră egale cu zero.

Mulţimea de distribuţii staţionare ale lanţului Markov cu multiple clase recurente se determină ca soluţie a sistemului liniar omogen:

$$\begin{cases}\pi^s=\pi^s P^s\\ \sum_{j=1}^{n}\pi_j^s=1\end{cases}, \tag{1}$$

în care $\pi^s=\left(\pi_j^s\right)_{j=1}^n$, $\pi_j^s\geq 0,\ j=1,...,n$. Soluţia generală a sistemului (1) reprezintă, de fapt, combinaţia liniară convexă a distribuţiilor staţionare în clasele recurente $X^{s,1},...,X^{s,k'}$ [2,5]:

$$\pi^s=\sum_{j=1}^{k'}\alpha_j\pi_{X^j}^s,\ \alpha_j\geq 0,\ j=1,...,k',\ \sum_{j=1}^{k'}\alpha_j=1. \tag{2}$$

Dacă π^s este mulţimea de distribuţii staţionare ale lanţului Markov, atunci totalitatea de costuri medii asociate lanţului poate fi exprimată astfel:

$$\psi_{\pi^s}(s)=\mu^s\left(\pi^s\right)^T=\sum_{j=1}^{k'}\alpha_j\mu^s\left(\pi_{X^j}^s\right)^T=\sum_{j=1}^{k'}\alpha_j\psi_{X^j}(s),$$

ceea ce înseamnă că mulţimea de costuri medii $\psi_{\pi^s}(s)$ în reţeaua cu multiple clase recurente este combinaţia convexă a costurilor medii în clasele recurente. Coeficienţii α_j ai combinaţiei convexe pot fi definiţi dacă este definită starea iniţială $x_{i_0}\in X$. Costul de valoare minimă pe mulţimea $\psi_{\pi^s}(s)$ se atinge pe una din clasele recurente $X^{s,j}$, în dependenţă de starea iniţială $x_{i_0}\in X$ (din starea iniţială $x_{i_0}\in X$ lanţul Markov va fi absorbit în careva clasă recurentă).

Se va ţine cont că orice strategie s în reţeaua decizională R induce un lanţ Markov cu un număr finit de clase recurente şi stări tranziente, lanţ care poate fi interpretat ca reuniune a sublanţurilor Markov cu o singură clasă recurentă şi, eventual, stări tranziente. Astfel, strategia s este reuniunea substrategiilor în reţelele cu o singură clasă recurentă (şi stări tranziente). În problema stocastică de control optimal discret, definită pe reţeaua decizională $R=\left(G,X_C,X_N,c,p\right)$ cu multiple clase recurente, se vor afla strategiile staţionare s^*

STUDIA UNIVERSITATIS MOLDAVIAE, 2015, nr.2(82)
Seria "Științe exacte și economice" ISSN 1857-2073 ISSN online 2345-1033 p.31-40

în rețeaua R, astfel încât, pentru orice stare inițială $x_{i_0} \in X$, combinația convexă a costurilor medii în subrețelele cu o singură clasă recurentă, ce corespund strategiei s^*, să ia valoare minimă.

În problema stocastică de control optimal discret, definită pe rețea decizională perfectă, toate stările lanțului Markov sunt recurent pozitive şi orice strategie fixată în rețeaua decizională induce un lanț Markov cu o singură clasă recurentă [3,4]. În acest caz, în stările controlabile se fixează dirijări, astfel încât pentru strategia fixată în rețea costul mediu per tranziție de realizare a lanțului Markov să ia valoare minimă. În problema de control, în care strategia staționară în rețeaua decizională induce un lanț Markov cu o singură clasă recurentă şi stări tranziente, se află mai întâi stările recurent pozitive şi se fixează strategiile de cost mediu minim în clasa recurentă corespunzătoare, după care, dacă starea inițială a sistemului este o stare tranzientă, atunci, pornind de la aceasta, se fixează arbitrar o dirijare în fiecare stare controlabilă din exteriorul clasei recurente, astfel încât dirijările fixate să definească un drum până la clasa recurentă. Reuniunea dirijărilor fixate în stările controlabile (ce aparțin clasei recurente şi din exteriorul acesteia) definesc strategiile optime în problema de control optimal definită pe rețeaua decizională cu o singură clasă recurentă şi stări tranziente. Costul mediu (inclusiv cel minim) pentru orice stare inițială $x_{i_0} \in X$ este egal cu costul mediu de tranziție a sistemului prin stările clasei recurente, deoarece pentru stările tranziente se satisface proprietatea că după un număr finit de paşi lanțul Markov al stărilor este absorbit în clasa recurentă. În [3,4] a fost elaborat şi argumentat un model de programare liniară, soluțiile optime de bază ale căruia permit fixarea strategiilor optime staționare în problemele de control optimal discret, definite pe rețele decizionale cu o singură clasă recurentă.

În rețelele decizionale în care strategia fixată induce un lanț Markov cu multiple clase recurente şi, eventual, stări tranziente, atunci când numărul de tranziții ale sistemului $k \to \infty$, lanțul Markov $\left\{X_k^s \,\middle|\, k \geq 0\right\}$ va fi absorbit în una din clasele recurente şi va rămâne acolo pentru totdeauna. De fapt, pentru o strategie fixată s în rețea sunt posibile următoarele situații:

1. Starea inițială $x_{i_0} \in X$ a sistemului dinamic este stare recurent pozitivă ce aparține unei clase recurente $X^{s,j}$ a lanțului Markov. Atunci, lanțul Markov va vizita doar stări ale clasei $X^{s,j}$ (deoarece nu există dirijări admisibile spre exteriorul acesteia) şi urmează să se fixeze dirijări în stările controlabile ale clasei $X^{s,j}$, astfel încât costul mediu per tranziție în această clasă (care este acelaşi, independent de starea inițială $x_{i_0} \in X^{s,j}$) să fie minim.

2. Starea inițială $x_{i_0} \in X$ este tranzientă şi de la ea există cel puțin un drum până la careva componentă tare conexă izolată ce corespunde unei clase recurente. Este posibil să existe drumuri către mai multe clase recurente. Pentru diferite clase recurente costurile medii pot fi distincte, de aceea este important, în vederea stabilirii strategiilor de cost mediu minim, în care clasă recurentă va fi absorbit lanțul Markov. Dacă decidentul are posibilitate să dirijeze traiectoria până la fiecare clasă recurentă, se va alege un drum până la clasa de cost mai mic. În situațiile de incertitudine, când în dependență de repartiția de probabilitate pot fi accesate mai multe clase recurente, se vor menționa toate variantele posibile.

Reieşind din cele expuse anterior, în vederea determinării strategiilor optime staționare se va proceda astfel:

a) Dintre clasele recurente ale lanțului Markov se determină o clasă $X^{s,j}$ $\left(s \in S,\ j \in \{1,...,k'\}\right)$ în care poate fi fixată strategia de cost mediu minim, după care acestei clase i se asociază o mulțime de stări tranziente din care aceasta poate fi accesată. Pentru stările tranziente controlabile se vor fixa dirijări, care asigură existența unui drum până la clasa recurentă (componentă tare conexă izolată în graful G al rețelei decizionale R);
b) Din graful G sunt eliminate stările şi arcele clasei recurente determinate şi ale mulțimii de stări tranziente asociate, iar în rețeaua deczională astfel formată se determină o nouă clasă recurentă de cost mediu minim şi mulțimea de stări tranziente asociate acesteia.

Procedeul continuă până atunci când toate stările mulțimii X sunt incluse în careva clasă recurentă sau mulțime de stări tranziente.

2. Modelul de programare liniară pentru determinarea strategiilor optime pe mulțimea stărilor recurente. În această secțiune este arătat că modelul de programare liniară elaborat în [3,4] poate fi aplicat pentru a fixa strategiile optime staționare în problema stocastică de control optimal discret, definită pe rețea decizională cu multiple clase recurente.

Conform celor menționate anterior, pentru strategia fixată s şi mulțimea de distribuții staționare π^s ale lanțului Markov, valorile $\psi\left(\pi^s\right)=\sum_{x\in X_C}\mu_x^s\pi_x^s+\sum_{z\in X_N}\mu_z^s\pi_z^s$ exprimă costurile medii per tranziție de realizare a lanțului Markov (pentru o stare inițială arbitrară) când este fixată strategia s.

Ținând cont de faptul că pentru soluția generală a sistemului (1) are loc reprezentarea de forma (2), problema de programare liniară

$$\psi\left(\pi^s\right)\to\min$$
$$\begin{cases}\pi^s=\pi^sP^s\\ \sum_{j=1}^{n}\pi_j^s=1\end{cases}\tag{3}$$
$$\pi_j^s\ge 0,\ j=1,\dots,n$$

posedă soluții admisibile de bază. Mulțimea de soluții admisibile π^s ale problemei (3) este de forma (2). Orice soluție optimă de bază a problemei (3) reprezintă un vârf al mulțimii poliedrale, definite de relația (2), care este, de fapt, distribuția staționară $\pi_{X^j}^s$ pe una din clasele recurente $X^{s,j},\ j\in\{1,\dots,k'\}$. Astfel, pentru s fixat costul mediu minim se atinge pe una din clasele recurente ale lanțului Markov. Dacă pentru strategia fixată s lanțul Markov conține mai multe clase recurente de cost mediu minim (cel mult k' clase recurente), atunci avem acelaşi număr de soluții optime de bază în problema (3), fiecare din acestea fiind o distribuție staționară pe o clasă recurentă $X^{s,j},\ j\in\{1,\dots,k'\}$.

Pentru fiecare strategie s din mulțimea de strategii staționare admisibile S în rețeaua R se găsesc soluțiile optime de bază în problema asociată de programare liniară (3) şi se fixează acele soluții pentru care costul este minim, ceea ce este echivalent cu rezolvarea problemei de programare liniară:

$$\psi\left(s,\pi^s\right)\to\min$$
$$\begin{cases}\pi^s=\pi^sP^s\\ \sum_{j=1}^{n}\pi_j^s=1\end{cases}\ .\tag{4}$$
$$s\in S,\ \pi_j^s\ge 0,\ j=1,\dots,n$$

Soluțiile optime de bază ale problemei (4) vor defini clasele recurente $X^{s,j}$ de cost mediu minim în rețeaua decizională R.

Ținând cont de semnificația variabilelor booleene $s_{x,y},\ x\in X_C, y\in X\left(x\right)$, pentru strategia arbitrar fixată s sistemul de ecuații $\pi^s=\pi^sP^s,\ \sum_{j=1}^{n}\pi_j^s=1$ se scrie sub forma:

$$\begin{cases}\sum_{x\in X_C}s_{x,y}\pi_x^s+\sum_{z\in X_N}p_{z,y}\pi_z^s=\pi_y^s, & \forall y\in X\\ \sum_{x\in X_C}\pi_x^s+\sum_{z\in X_N}\pi_z^s=1\end{cases},$$

iar funcția obiectiv în problema (4) se scrie astfel:

$$\psi\left(s,\pi^s\right)=\sum_{x\in X_C}\sum_{y\in X(x)}c_{x,y}s_{x,y}\pi_x^s+\sum_{z\in X_N}\mu_z\pi_z^s\ .$$

Astfel, pentru problema stocastică de control optimal discret definită pe rețeaua decizională $R=\left(G,X_C,X_N,c,p\right)$, în care strategia staționară induce un lanț Markov cu multiple clase recurente, are loc următorul rezultat.

Lema 1. *Orice soluție optimă de bază a problemei mixte de programare liniară în numere întregi*

$$\psi\left(s,\pi^s\right)=\sum_{x\in X_C}\sum_{y\in X(x)} c_{x,y}s_{x,y}\pi_x^s+\sum_{z\in X_N}\mu_z\pi_z^s\to\min, \tag{5}$$

$$\begin{cases}\sum\limits_{x\in X_C} s_{x,y}\pi_x+\sum\limits_{z\in X_N} p_{z,y}\pi_z=\pi_y, \quad \forall y\in X\\ \sum\limits_{x\in X_C}\pi_x+\sum\limits_{z\in X_N}\pi_z=1\\ \sum\limits_{y\in X(x)} s_{x,y}=1, \quad \forall x\in X_C\\ s_{x,y}\in\{0,1\}, \quad \forall x\in X_C, \ \ y\in X; \ \ \pi_x\geq 0, \ \forall x\in X\end{cases} \tag{6}$$

definește subrețeaua rețelei $R=\left(G,X_C,X_N,c,p\right)$, *ce corespunde unui lanț Markov cu o singură clasă recurentă , în care poate fi fixată strategia de cost mediu per tranziție minim.*

În continuare, pentru orice stare $y\in X$ vom utiliza notațiile $X_C^-\left(y\right)$, $X^-\left(y\right)$ pentru mulțimile:

$$X_C^-\left(y\right)=\left\{x\in X_C \mid y\in X, y-fixat, \left(x,y\right)\in E\right\}, \ X^-\left(y\right)=\left\{x\in X \mid y\in X, y-fixat, \left(x,y\right)\in E\right\}.$$

În baza lemei 1, analog ca şi în [3,4], se poate demonstra următorul rezultat.

Teorema 1. *Fie* $\alpha_{x,y}^*$ $\left(x\in X_C, y\in X\right)$, π_x^* $\left(x\in X\right)$ *o soluție optimă de bază a problemei de programare liniară:*

$$\psi(\bar{\alpha},\bar{\pi})=\sum_{x\in X_C}\sum_{y\in X(x)} c_{x,y}\alpha_{x,y}+\sum_{z\in X_N}\mu_z\pi_z\to\min, \tag{7}$$

$$\begin{cases}\sum\limits_{x\in X_C^-(y)}\alpha_{x,y}+\sum\limits_{z\in X_N} p_{z,y}\pi_z=\pi_y, \quad \forall y\in X,\\ \sum\limits_{x\in X_C}\pi_x+\sum\limits_{z\in X_N}\pi_z=1,\\ \sum\limits_{y\in X(x)}\alpha_{x,y}=\pi_x, \quad \forall x\in X_C,\\ \alpha_{x,y}\geq 0, \ \forall x\in X_C, \ y\in X\left(x\right); \ \pi_x\geq 0, \ \forall x\in X,\end{cases} \tag{8}$$

unde $\mu_z=\sum\limits_{y\in X(z)} p_{z,y}c_{z,y}, \ \forall z\in X_N$. *Atunci, componentele* $\pi_{x_j}^*>0$ $(x_j\in X)$ *ale soluției optime de bază definesc clasa recurentă* X^* *a lanțului Markov, ce corespunde strategiei staționare* s^* *de cost mediu minim în rețeaua* $R=\left(G,X_C,X_N,c,p\right)$. *Strategia optimă* s^* *în rețeaua indusă de stările clasei recurente* X^* *se determină în baza relației* $s_{x,y}^*=\begin{cases}1 \ \ dacă \ \alpha_{x,y}^*>0\\ 0 \ \ dacă \ \alpha_{x,y}^*=0\end{cases}$, *unde* $x\in X_C$, $y\in X(x)$. *Costul mediu per tranziție de realizare a strategiei* s^* *coincide cu valoarea funcției obiectiv* $\psi\left(\bar{\alpha}^*,\bar{\pi}^*\right)$.

3. Algoritmul pentru determinarea strategiilor optime staționare. În baza teoremei 1, propunem următorul algoritm pentru fixarea strategiilor optime staționare în problema stocastică de control optimal discret, definită pe rețeaua R cu multiple clase recurente:

STUDIA UNIVERSITATIS MOLDAVIAE, 2015, nr.2(82)
Seria "Științe exacte și economice" *ISSN 1857-2073* *ISSN online 2345-1033* *p.31-40*

Date inițiale:

$X = \{x_1, ..., x_n\}$ – mulțimea finită de stări ale sistemului dinamic;

E – mulțimea de arce ale grafului asociat ce reprezintă tranzițiile admisibile între stări;

X_C, X_N – mulțimile de stări controlabile şi, respectiv, necontrolabile;

c – funcția discretă ce defineşte costurile de tranziție între stări;

p – funcția discretă ce defineşte repartiția de probabilitate corespunzătoare tranzițiilor din stările necontrolabile;

Date de ieşire:

$\bar{X}^i, i = \overline{1, j-1}$ – clasele recurente ordonate în sensul creşterii costurilor medii;

s^i – strategia staționară optimă în clasa $\bar{X}^i, i = \overline{1, j-1}$;

ψ^i – costul mediu per tranziție minim în clasa $\bar{X}^i, i = \overline{1, j-1}$.

$X1^i_{tr}$ – mulțimea de stări tranziente asociate clasei $\bar{X}^i, i = \overline{1, j-1}$

E^i_{tr} – mulțimea de arce asociate stărilor tranziente din $X1^i_{tr}$, $i = \overline{1, j-1}$

$j = 1;\ X^j = X;\ E^j = E$

$G^j = (X^j, E^j)$ – Graful ce corespunde tranzițiilor de stare ale sistemului;

$X_t = \varnothing;\ X^j_C = X_C;\ X^j_N = X_N;\ c^j = c;\ \ p^j = p$

$X1^j_{tr} = \varnothing;\ E1_{tr} = \varnothing;\ criteriu = 1$

Se scrie PPL (7), (8) asociată rețelei $R^j = (G^j, X^j_C, X^j_N, c^j, p^j)$

Cât timp $criteriu = 1$ execută

{

Se află o soluție optimă de bază $\alpha^*_{x,y}\ (x \in X_C, y \in X(x)),\ \pi^*_x\ (x \in X)$ a PPL (7), (8)

Se calculează costul mediu minim $\psi^i = \psi(\bar{\alpha}^*, \bar{\pi}^*)$

Se precizează mulțimea de stări $\bar{X}^j$ ale clasei recurente de cost mediu minim în rețeaua R^j: $\bar{X}^j = \{x_j \in X \mid \pi^*_{x_j} > 0\}$

Se fixează strategia optimă staționară s^j în clasa recurentă $\bar{X}^j$:

dacă $\alpha^*_{x,y} > 0$ pentru $x \in \bar{X}^j \cap X_C,\ y \in X(x)$, atunci $s^j_{x,y} = 1$;

în caz contrar, se consideră $s^j_{x,y} = 0$.

Se precizează mulțimea de arce $\bar{E}^j$ ale grafului $(\bar{X}^j, \bar{E}^j)$ ce corespunde clasei recurente determinate $\bar{X}^j$:

Pentru fiecare $x \in \bar{X}^j$ procedăm astfel:

dacă $x \in \bar{X}^j \cap X^j_N$, atunci în $\bar{E}^j$ se includ toate arcele incidente spre exterior vârfului x;

dacă $x \in \bar{X}^j \cap X^j_C$, atunci în $\bar{E}^j$ se include arcul $(x, s^j(x))$ pe care a fost fixată dirijarea optimă în x, iar restul arcelor incidente spre exterior nodului x sunt eliminate din mulțimea E^j: $E^j = E^j \setminus V1,\ \ V1 = \{(x,y) \in E^j \mid x \in \bar{X}^j \cap X^j_C, y \neq s^j(x)\}$;

$X^j_{tr} = \varnothing;\ E^j_{tr} = \varnothing;\ Y^j = \bar{X}^j;\ X1^j_{tr} = \varnothing$

Cât timp există noduri $x \in X^j \setminus Y^j$ pentru care arcele incidente spre exterior (din mulțimea E^j) au extremitatea finală în Y^j

{

Se adaugă la mulțimea X_{tr}^{j} stările $x \in X^{j} \setminus Y^{j}$ din care există arc incident spre exterior cu extremitatea finală în Y^{j}

Pentru fiecare element x al mulțimii X_{tr}^{j} se efectuează următoarele:

{

a) dacă starea $x \in X_C$, atunci se adaugă la mulțimea E_{tr}^{j} unul din arcele $(x, y) \in E^{j}$, $y \in Y^{j}$ şi se fixează $s_{x,y}^{*} = 1$. Celelalte arce $e \in E^{j}$, cu extremitatea inițială în starea examinată $x \in X_C$ sunt eliminate din mulțimea E^{j};

b) dacă starea $x \in X_N$, atunci se adaugă la mulțimea E_{tr}^{j} arcele $(x, y) \in E^{j}$, $y \in Y^{j}$;

c) se adaugă la mulțimea E_{tr}^{j} arcele mulțimii $V2 = \left\{(x, y) \in E^{j} \,\middle|\, x \in \left(X1_{tr}^{j} \cup X_{tr}^{j}\right) \cap X_N,\ y \in X_{tr}^{j}\right\}$.

}

$Y^{j} = Y^{j} \cup X_{tr}^{j};\ X1_{tr}^{j} = X1_{tr}^{j} \cup X_{tr}^{j};\ X_{tr}^{j} = \varnothing$

}

La mulțimea $E1_{tr}$ sunt adăugate arcele mulțimii $V3 = \left\{(x, y) \in E^{j} \,\middle|\, x \in X1_{tr}^{j} \cap X_N,\ y \notin \left(X1_{tr}^{j} \cup \bar{X}^{j}\right)\right\}$:

$$E1_{tr} = E1_{tr} \cup V3$$

$j = j + 1;\ X_t = X_t \cup Y^{j-1}$

Dacă $X \setminus X_t \neq \varnothing$, atunci execută

{

Se construieşte o nouă rețea decizională $R^{j} = \left(G^{j}, X_C^{j}, X_N^{j}, c^{j}, p^{j}\right)$ ale cărei componente se determină în modul următor:

Graful parțial $G^{j} = \left(X^{j}, E^{j}\right)$ se construieşte astfel: $X^{j} = X \setminus X_t;\ E^{j} = E^{j-1} \setminus \left(E_{tr}^{j-1} \cup \bar{E}^{j-1} \cup E1_{tr}\right)$;

$X_C^{j} = X_C \cap X^{j};\ X_N^{j} = X_N \cap X^{j}$;

Matricele c^{j} şi p^{j} se obțin din matricele c^{j-1} şi p^{j-1} prin eliminarea tuturor rândurilor şi coloanelor ce corespund indicilor stărilor din mulțimea $X \setminus X^{j}$

Se scrie PPL (7), (8) asociată rețelei R^{j}

}

Altfel

{

$criteriu = 0$

Se afişează mulțimile $\bar{X}^{i}, s^{i}, \psi^{i}, X1_{tr}^{i}, E_{tr}^{i},\ i = \overline{1, j-1}$

}

}

Concluzii

Abordarea considerată privind soluționarea problemei stocastice de control optimal discret, definită pe rețele decizionale cu multiple clase recurente, permite formularea unor algoritmi cu estimație polinomială a complexității pentru determinarea soluției optime a problemei.

Bibliografie:

1. HOWARD, R. *Dynamic programming and Markov processes*. Technology Press of Massachusetts Institute of Technology, 1960. 148 p.
2. LAWLER, G. *Introduction to stochastic processes*. Second edition. Chapman & Hall/CRC, 2006.
3. LOZOVANU, D., CAPCELEA, M. Algorithms for solving stochastic discrete optimal control problems on networks. În: *Buletinul Academiei de Ştiințe a Moldovei. Matematica*, 2014, nr.3(76), p.80-88.

4. LOZOVANU, D., PICKL, S. *Optimization of stochastic discrete systems and control on complex networks. Computational networks*. Springer, Advances in Computational Management Science, 2015, vol.12.
5. PINSKY, M., KARLIN, S. *An introduction to stochastic modeling*. Fourth edition. Elsevier, 2011.
6. PUTERMAN, M. *Markov decision processes. Discrete stochastic dynamic programming*. Wiley-Interscience, 2005.

Notă*:* Cercetările asupra rezultatelor prezentate în acest articol au fost parțial finanțate din cadrul Proiectului ştiinţific 14.819.02.14F „*Metode numerice de soluționare a problemelor de optimizare stochastice*".

Prezentat la 03.06.2015