ADAM PŁOCKI

Convergence d'un vecteur invariante d'une matrice stochastique $Q(\varepsilon) = (1-\varepsilon)Q_1+\varepsilon Q_2$, ou $Q_1^T = Q_1$.

Introduction. Soit (E, S, P) un system probabilistique et ξ_n une variable aléatoire déffinie sur E pour $n \in N_0$.

Soit $\xi_n(E) = \{a_1, a_2, \dots, a_s\}$ pour $n \in N_0$. Cette travail conçernent d'une suite des variables aléatoires, qui est un chaîne de Markow de s états.

Posons $P(\xi_n = a_i) = P(\{e \in E\colon \xi_n(e) = a_i\} = p_i^n$ et pour $n>m$, $P(\xi_n = a_j|\xi_m = a_i) = P(\{e \in E\colon \xi_n(e) = a_j\} | \{e \in E\colon \xi_m(e) = a_i\}) = p_{ij}(m, n)$.

Les probabilité conditionelles $p_{ij}(m, n)$ forment en cet case une matrice carrée d'ordre s, nous la marquons par $Q(m, n)$. Le vecteur $\overline{m}_n = [p_1^n, p_2^n, \dots, p_s^n]$ définit uniquement une loi d'une variable aléatoire ξ_n pour $n \in N_0$. Nous l'appellons le vecteur de la distribution en moment n.

Soit (ξ_n) un chaîne homogene, c'est dire que $Q(m, n) = Q(n-m)$. Le vecteur $\overline{m}_0 = [p_1^0, p_2^0, \dots, p_s^0]$ nous appellons le vecteur **invariante** si $\overline{m}_0 Q(1) = \overline{m}_0$, où $Q(1)$ est une matrice de passer dans un pas pour le chaîne homogene (ξ_n).

Un stationnaire chaîne de Markow (ξ_n) définit uniquement la matrice $Q(1)$ et aussi le vecteur invariante $\overline{m}_0$.

On peut démontrer, que pour chaque homogene chaîne de Markow (ξ_n) $n \in N_0$, on peut trouver un vecteur initial $\overline{m}_0$ tel, pour recevoir un chaîne stationnaire, alors pour chaque matrice $Q(1)$ on peut trouver un vecteur invariante $\overline{m}_0$.

Soient $Q_1 = [q_{ij}^1]$ et $Q_2 = [q_{ij}^2]$ deux matrices stochastiques d'ordre s. Soit $Q_1 \geqslant 0$ et $Q_2 > 0$.

Alors

$$\sum_{i=1}^{s} q_{ij}^1 = \sum_{i=1}^{s} q_{ij}^2 = 1 .$$

On démontre, que pour chaque matrice non négatif Q_1 il existe au moins un vecteur invariante $\overline{m}_0^1$, et pour la matrice positif il existe exactement un vecteur invariante $\overline{m}_0^2$.

Parce que pour $j = 1, 2, \dots, s$

$$\sum_{i=1}^{s} q_{ij}^1(1-\varepsilon)+\sum_{i=1}^{s} q_{ij}^2\varepsilon = 1 ,$$

alors pour ε suffisamment proche de zéro la matrice $Q(\varepsilon) = (1-\varepsilon)Q_1+\varepsilon Q_2$ est aussi une matrice stochastique d'ordre s.
Puisque $Q(\varepsilon)>0$, alors pour la matrice $Q(\varepsilon)$ il existe exactement un vecteur invariante $\overline{m}_0(\varepsilon)$.

Si $\varepsilon\to 0+$, la matrice $Q(\varepsilon)$ converge vers la matrice Q_1. Il y a alors un problème, que fait le vecteur $\overline{m}_0(\varepsilon)$? A quelle des vecteurs invariants converge $\overline{m}_0(\varepsilon)$, si $\varepsilon\to 0+$. Marqueons en suit $\overline{m}_0(\varepsilon) = \overline{m}(\varepsilon)$.

Dans cet travaill le problème est tranché dans le cas, où la matrice Q_1 es symétrique, c'est-à-dire quand $Q_1^T = Q_1$.

Avant, nous formulons quelques théorèmes conçernant le problème posé.

THÉORÈME 1. *S'il existe* $\lim\limits_{\varepsilon\to 0+} \overline{m}(\varepsilon) = \overline{m}$, *il est que* $\overline{m}$ *est un vecteur invariante correspondant à la matrice* Q_1, *c'est-à dire* $\overline{m}Q_1 = \overline{m}$.

Démonstration. D'après la définition $Q(\varepsilon) = Q_1+[Q_2-Q_1]\varepsilon$ et

$$\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon)[Q_1+(Q_2-Q_1)\varepsilon],$$

alors

$$\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon)Q_1+\overline{m}(\varepsilon)[(Q_2-Q_1)\varepsilon].$$

Mais,

$$\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon), \qquad \text{d'où}$$

$$\overline{m}(\varepsilon) = \overline{m}(\varepsilon)Q_1+\overline{m}(\varepsilon)[(Q_2-Q_1)\varepsilon].$$

Si $\varepsilon\to 0+$, nous avons que $(Q_2-Q_1)\varepsilon\to[0]$, où $[0]$ est une matrice nulle. Parce que les coordinates de vecteur $\overline{m}(\varepsilon)$ sont positif et sa somme est egale 1, alors $\overline{m}(\varepsilon)[(Q_2-Q_1)\varepsilon]$ converge vers un vecteur nulle si $\varepsilon\to 0+$.
Par hypothèse $\lim\limits_{\varepsilon\to 0+} \overline{m}(\varepsilon) = \overline{m}$, d'où $\lim\limits_{\varepsilon\to 0+} \overline{m}(\varepsilon) = \lim\limits_{\varepsilon\to 0+} \overline{m}(\varepsilon)Q_1$, c'est-à-dire $\overline{m} = \overline{m}Q_1$, ce qu'il fallait démontrer.

Il faut remarquer, que si généralment $Q(\varepsilon)$ et Q sout deux matrices stochastiques d'ordre s et $\lim\limits_{\varepsilon\to 0+} Q(\varepsilon) = Q$, le vecteur invariante $\overline{m}(\varepsilon)$ correspondant à la matrice $Q(\varepsilon)$ ne doit pas être convergent. Par exemple:

$$Q(\varepsilon) = \begin{cases} \begin{bmatrix} 1-\varepsilon & \varepsilon \\ 2\varepsilon & 1-2\varepsilon \end{bmatrix} & \text{où } \varepsilon \text{ rationnel} \\ \begin{bmatrix} 1-\varepsilon & \varepsilon \\ 3\varepsilon & 1-3\varepsilon \end{bmatrix} & \text{où } \varepsilon \text{ irrationnel, où } 0<\varepsilon<\tfrac{1}{3} \end{cases}$$

Il est

$$\lim_{\varepsilon\to 0+} Q(\varepsilon) = Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

par contre

$$\overline{m}(\varepsilon) = \begin{cases} [\tfrac{2}{3}, \tfrac{1}{3}] & \text{où } \varepsilon \text{ rationnel,} \\ [\tfrac{3}{4}, \tfrac{1}{4}] & \text{où } \varepsilon \text{ irrationnel.} \end{cases}$$

Si $\varepsilon\to 0+$, $\overline{m}(\varepsilon)$ n'est pas convergent.

THÉORÈME 2. *Si* $\lim_{\varepsilon\to 0+} Q(\varepsilon) = Q$ *et le vecteur invariante* $\overline{m}$ *correspondant à la matrice* Q *est unique, le vecteur* $\overline{m}(\varepsilon)$ *correspondant à la matrice* $Q(\varepsilon)$ *est convergent vers le vecteur* $\overline{m}$, *si* $\varepsilon\to 0+$.

Démonstration: Supposons, que $\lim_{\varepsilon\to 0+} \overline{m}(\varepsilon)$ n'existe pas. Alors, ou peut trouve deux suites (ε_n^1) et (ε_n^2) conveıgentes vers 0 du côté droit, tels, que $\lim_{n\to\infty} \overline{m}(\varepsilon_n^1) = \overline{m}^1$, $\lim_{n\to\infty} \overline{m}(\varepsilon_n^2) = \overline{m}^2$ et $\overline{m}^1 \neq \overline{m}^2$, où $(\overline{m}(\varepsilon_n^1))$ et $(\overline{m}(\varepsilon_n^2))$ sont deux suites des vecteurs invariantes correspondants tour à tour à la matrice $Q(\varepsilon_n^1)$ et $Q(\varepsilon_n^2)$.

Par hypothèse pour tout ε, il est: $\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon)$, il en résulte

$$\overline{m}(\varepsilon_n^1)Q(\varepsilon_n^1) = \overline{m}(\varepsilon_n^1) \quad \text{et} \quad \overline{m}(\varepsilon_n^2)Q(\varepsilon_n^2) = \overline{m}(\varepsilon_n^2)\,.$$

Par hypothèse

$$\lim_{n\to\infty} Q(\varepsilon_n^1) = \lim_{n\to\infty} Q(\varepsilon_n^2) = Q\,,$$

alors

$$\overline{m}^1 Q = \overline{m}^1 \quad \text{et} \quad \overline{m}^2 Q = \overline{m}^2\,.$$

Les vecteurs $\overline{m}^1$ et $\overline{m}^2$ sont invariantes pour la matrice Q, et puisque Q posedent le vecteur invariante unique, il est $\overline{m}^1 = \overline{m}^2$, ce qui est absurde.

La deuxième partie de la thése résulte immediatement d'après théorème 1.

THÉORÈME 3. *Si le vecteur* $\overline{m}$ *est un vecteur invariant pour la matrice* Q_1, *et* $\overline{m}^2$ *est l'unique vecteur invariant correspondant à la matrice* Q_2, *il est* $\overline{m}^2 = \overline{m}(\varepsilon)$, *où* $\overline{m}(\varepsilon)$ *est un vecteur correspondant à la matrice* $Q(\varepsilon) = (1-\varepsilon)Q_1 + \varepsilon Q_2$.

Demonstration. Par hypothèse $\overline{m}^2 \cdot Q_2 = \overline{m}^2$.

Puisque $Q(\varepsilon) = Q_1 - Q_1\varepsilon + Q_2\varepsilon$, alors

(i) $$\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon)Q_1 - \overline{m}(\varepsilon)Q_1\varepsilon + \overline{m}(\varepsilon)Q_2\varepsilon.$$

Mais $\overline{m}(\varepsilon)Q(\varepsilon) = \overline{m}(\varepsilon)$ et $\overline{m}(\varepsilon)Q_1 = \overline{m}(\varepsilon)$, car $\overline{m}(\varepsilon)$ est aussi un vecteur invariant correspondant à la matrice Q_1, alors (i) est sous la forme

$$\overline{m}(\varepsilon) = \overline{m}(\varepsilon) - \overline{m}(\varepsilon)\varepsilon + \overline{m}(\varepsilon)Q_2\varepsilon$$

d'où

$$\overline{m}(\varepsilon)Q_2 = \overline{m}(\varepsilon)\,.$$

$\overline{m}(\varepsilon)$ est alors le vecteur invariante correspondant à la matrice Q_2 et d'après l'hypothèse $\overline{m}(\varepsilon) = \overline{m}^2$, alors la théorème est démontré.

THÉORÈME 4. *Soient* $\overline{m}^k \cdot Q = \overline{m}^k$ *et* $a_k > 0$ *pour* $k = 1, 2, \ldots, r$. *Le vecteur*

$$\overline{m} = \frac{1}{a_1 + a_2 + \ldots + a_r} \sum_{k=1}^{r} a_k \overline{m}^k$$

est aussi le vecteur invariant correspondant à la matrice Q, *c'est-à-dire* $\overline{m}Q = \overline{m}$.

Démonstration.

$$\overline{m}Q = \left[\frac{1}{a_1+a_2+\ldots+a_r}\sum_{k=1}^{r} a_k\overline{m}^k\right]Q = \frac{1}{a_1+a_2+\ldots+a_r}\sum_{k=1}^{r}(a_k\overline{m}^k Q) =$$

$$= \frac{1}{a_1+a_2+\ldots+a_r}\sum_{k=1}^{r} a_k(\overline{m}^k Q) =$$

$$= \frac{1}{a_1+a_2+\ldots+a_r}\sum_{k=1}^{r} a_k\overline{m}^k = \overline{m}, \text{ ce qu'il fallait démontrer.}$$

Soit Q_1 est une matrice stochastique positif d'ordre s. Les recherches des vecteurs invariants de la matrice stochastique donné se reduit de deux etapes:

1^0 Le recherche des vecteurs propres de ce matrice, qui correspondent à le valeur caractéristique 1,

2^0 La réglementation des ces vecteurs, la somme des coordonneés d'un vecteur invariant doit être égal 1.

Il est facilement démontrer, que etap première se reduit à la résolution d'un système de s l'équations linéaires homogènes, non-où leur determinant est égal zèro. Dans [3] est démontré, que touts non-nulles coordonnées d'un vecteur, qui est la solution non-nulle de cet système ont le même signe, d'où la possibilite de la réglementation de chaque de ces solution.

La famille W des vecteurs propres de la matrice stochastique est alors une espace vectoriel.

Soit $\{\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_r\}$ est une base orthogonale. Considerons $Q(\varepsilon) = (1-\varepsilon)Q_1 + \varepsilon Q_2$. Si $\overline{m}(\varepsilon)$ posséde la limite, quand $\varepsilon \to 0+$, elle est — que nous démontrons-un vecteur invariant de la matrice Q_1. Il existe alors dans la famille W un vecteur $\overline{w}$ réglementation, où $\overline{m} = \lim_{\varepsilon\to 0+} \overline{m}(\varepsilon)$.

Le problème, qui est posé en avant, se reduit de trouver dans la famille W tel vecteur $\overline{w}$, quelle aprés une réglementation donne le vecteur $\overline{m}$.

Nous tranchons ce problème dans le cas, où $Q_1^T = Q_1$. Soit $\overline{w}(\varepsilon)$ est un vecteur propre de la matrice $Q(\varepsilon)$, correspondant à la valuer caractéristique 1, et qui aprés la réglementation donne le vecteur $\overline{m}(\varepsilon)$. Lest corrdonnées de le vecteur $\overline{w}(\varepsilon)$ sont les functions de ε, de la classe C^∞. Posons par $\overline{w}^{(k)}(\varepsilon)$ le vecteur, les coordonnées de qui, sont les dérivées d'ordre k des coordonnées de $\overline{w}(\varepsilon)$. Soit $\overline{w}^{(k)} = \overline{w}^{(k)}(0)$ pour $k \in N$.

Il est alors

$$\text{(ii)} \qquad \overline{w}(\varepsilon) = \overline{w} + \varepsilon\overline{w}^{(1)} + \frac{\varepsilon^2}{2!}\overline{w}^{(2)} + \ldots$$

Parce que $\bar{m}$ est le vecteur invariant de la matrice Q_1, donc $\bar{w}$ est le vecteur propre de cette matrice correspondant à la valeur caractéristique 1. Le vecteur $\bar{w}$ est alors certain combinaison linéaire des vecteurs de la base orthogonale, c'est-à-dire que

$$\text{(iii)} \qquad \bar{w} = a_1 \bar{w}_1 + a_2 \bar{w}_2 + \ldots + a_r \bar{w}_r .$$

Il est $\bar{w}(\varepsilon)[Q_1 + (Q_2 - Q_1)\varepsilon] = \bar{w}$ et d'après (iii) nous avons

$$\text{(iv)} \qquad \left(\bar{w} + \varepsilon \bar{w}^{(1)} + \frac{\varepsilon^2}{2!}\bar{w}^{(2)} + \ldots\right)[Q_1 + (Q_2 - Q_1)\varepsilon] = \bar{w} + \varepsilon \bar{w}^{(1)} + \frac{\varepsilon^2}{2!}\bar{w}^{(2)} + \ldots$$

En comparant les coefficients prés de ε nous avons

$$\text{(v)} \qquad \bar{w}(Q_2 - Q_1) + \bar{w}^{(1)} Q_1 = \bar{w}^{(1)}$$

On multiple scalairement (v) par le vecteur de la base $\bar{w}_k$ nous avons

$$\text{(vi)} \qquad [\bar{w}(Q_2 - Q_1]) \circ \bar{w}_k + (\bar{w}^{(1)} Q_1) \circ \bar{w}_k = \bar{w}^{(1)} \circ \bar{w}_k .$$

Mais la matrice Q_1 est symétrique, alors d'après l'hypothèse il est

$$(\bar{w}^{(1)} Q_1) \circ \bar{w}_k = \bar{w}^{(1)} \circ [\bar{w}_k Q_1] = \bar{w}^{(1)} \circ \bar{w}_k .$$

L'égalité (vi) est ainsi sous la forme

$$[\bar{w}(Q_2 - Q_1)] \circ \bar{w}_k = 0 .$$

En profitant (iii) nous avons:

$$[(a_1 \bar{w}_1 + a_2 \bar{w}_2 + \ldots + a_r \bar{w}_r)(Q_2 - Q_1)] \circ \bar{w}_k = 0$$

d'où

$$\text{(vii)} \qquad \sum_{j=1}^{r} a_j [\bar{w}_j (Q_2 - Q_1)] \circ \bar{w}_k = 0 \quad \text{où} \quad k = 1, 2, \ldots, r .$$

La formule (vii) permet trouver les coefficients a_j dans (iii), et qu'est la même chose, trouver le vecteur $\bar{w}$. Le vecteur $\bar{m}$ nous trouverons après la réglementation du vecteur $\bar{w}$. Le problème, qui est tranché ici, dans le cas particulier se lie avec la convergente des stationnaires chaînes de Markow, qui sont définié par la matrice de passer dans un pas $Q(\varepsilon)$ et le vecteur initiale (invariante) $\bar{m}_0(\varepsilon) = \bar{m}(\varepsilon)$, si $\varepsilon \to 0+$.

Ce problème se lie aussi avec des perturbations aléatoires de transmission d'informations par le canal. La mesure de cette perturbations est ε.

Le problème de trouver la limite $\lim_{\varepsilon \to 0+} \bar{m}(\varepsilon)$, où Q_1 est une arbitraire, non-symetrique matrice stochastique, est ouvert.

Travaux citès

[1] И. Гельфанд: *Лекции по линейной алгебре*. Издательство НАУКА, Москва 1971.

[2] A. Papoulis: *Prawdopodobieństwo, zmienne losowe i procesy stochastyczne*. WNT, Warszawa 1972.

[3] В. И. Романовский: *Дискретные цепи Маркова*. ГОСТЕХИЗДАТ 1949.