Nihonkai Math. J.
Vol.9 (1998), 193–204

Entropic decision criteria for random behaviour systems

Rodica Brânzei

Abstract

In this paper we study random behaviour systems with a finite number of states from the point of view of optimal transition from a state of the system to another one. We use entropic decision criteria like the Shannon entropy and the entropy of a Markov chain with a finite number of states. Our results are obtained by using the principle of maximum entropy according to Charnes [3], Gerchak [6], Guiaşu and Shenitzer [7] and generalize some results on linearly constrained entropy optimization problems from Gerchak [6], Erlander [4]. We determine an optimal strategy of transition of a finite number states system from a state to another one by a dual approach on a linearly constrained optimization problem following the line from Ben-Tal [1], [2], Fang, Tsao [5], [9], our objective function being defined as a weighted combination of the expected transition utility and the entropy of a Markov chain. Several pairs of dual problems with entropic criteria will be considered in order to determine the optimal transition probabilities from a state of the system to another one.

AMS Subject Classification: 49K45, 60J99, 94A17

Keywords: **Shannon entropy, entropy of a Markov chain, transition probabilities, entropy optimization problem, duality.**

1 Introduction

In many fields of the operational research as well as in pattern recognition we can identify random behaviour systems with a finite number of states for which it is important to determine an optimal way of transition from a state to another one.

Let $Z = \{z_1, ..., z_n\}$ be the set of the states of a random behaviour system. We note $N = \{1, ..., n\}$. Let us suppose that the probabilities π_i, $i \in N$, of the states, the utilities u_i, $i \in N$, of the states and the transition utilities from the state z_i of the system to the state z_j, denoted by u_{ij}, $i, j \in N$, are apriori known.

The evolution of the system is described by the transition probabilities p_{ij}, $i, j \in N$, where p_{ij} represent the probability of transition from the state z_i to the state z_j, $i, j \in N$, satisfying the following conditions:

$$(1.1) \qquad p_{ij} > 0,\ i, j \in N,\ \sum_{j=1}^{n} p_{ij} = 1,\ i \in N;\ \sum_{i=1}^{n} \pi_i p_{ij} = \pi_j,\ j \in N.$$

Certain observations made usually on such a system provide a set of explicit constraints on the values of the transition probabilities p_{ij}, $i, j \in N$, expressed by equalities and inequalities. In this paper we shall suppose that the transition probabilities satisfy the following set of linear constraints

$$(1.2) \qquad \sum_{i=1}^{n} \sum_{j=1}^{n} w_{k,ij} p_{ij} = d_k,\ k = 1, ..., m,$$

where $w_{k,ij}, d_k, i, j \in N,\ k = 1, ..., m$ are given real constants.

We suppose, without loss of generality, that the constraints (1.2) imply (1.1).

Let $p = (p_{11}, ..., p_{1n}, ..., p_{nn})^t$ be the vector of the transition probabilities. We also use the notations $d = (d_1, ..., d_m)^t$, $\mathcal{P} = \{p = (p_{ij})^t,\ i, j \in N \mid p_{ij} > 0,\ i, j \in N,\ p \text{ satisfying } (1.2)\}$.

The expected utility of the transition of the considered system from a state to another is given by

$$(1.3) \qquad U(p) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij} p_{ij},\ p \in \mathcal{P}.$$

Let $\{X_t \mid t = 0, 1, ...\}$ be a stationary Markov chain with the states Z having the probabilities π_i and the utilities u_i, $i \in N$. The entropy of such a Markov chain is defined by

$$(1.4) \qquad H(p) = -\sum_{i=1}^{n} \sum_{j=1}^{n} u_i \pi_i p_{ij} \ln p_{ij}.$$

Each vector $p \in \mathcal{P}$ defines a strategy of transition of the system from a state to another one. An optimal transition probability vector will be, for us, one maximizing the following weighted combination of $U(p)$ and $H(p)$

$$F(p) = \lambda_1 U(p) + \lambda_2 H(p), \ \lambda_1, \lambda_2 > 0, \ p \in \mathcal{P}. \tag{1.5}$$

We are interested in finding a Markov chain, among all admissible ones, attaining the maximum of the function $F(p)$ given by (1.5).

In section 2 we shall give some results on pairs of dual problems with the Shannon entropy in the expression of the objective function of primal problems. In section 3, we formulate explicitly our entropy optimization problem and determine the optimal transition probabilities from a state of the considered system to another one and section 4 contains conclusions.

2 Preliminary results

Let $x = (x_1, ..., x_n)^t$ be a probability distribution defined on the set of states Z, satisfying the following linear constraints

$$W \cdot x = d, \text{ with } x > 0, \tag{2.1}$$

where $W = (w_{k,i}), w_{k,i} \in \mathbb{R}, k = 1, ..., m, i \in N, d = (d_1, ..., d_m)^t$. We assume, without loss of generality, that the constraints (2.1) include $\sum_{i=1}^{n} x_i = 1$.

Let X be the set of probability distributions $x > 0$ satisfying (2.1). For $\lambda_1, \lambda_2 > 0$ and $u_i \in \mathbb{R}_+$, $i \in N$, let us consider the optimization problem

$$\max_{x \in X} F(x) = \max_{x \in X} \left(\lambda_1 \sum_{i=1}^{n} u_i x_i - \lambda_2 \sum_{i=1}^{n} x_i \ln x_i \right). \tag{P}$$

Theorem 2.1 *The problem* (P) *is a convex optimization with linear restrictions one and it admits an optimal unique solution for every pair* $\lambda_1, \lambda_2 > 0$.

Proof. We shall show that the Hessian matrix of $F(x)$, denoted by $\nabla^2(F)$, is negatively defined. We have

$$\frac{\partial F(x)}{\partial x_i} = \lambda_1 u_i - \lambda_2 (1 + \ln x_i), \qquad i \in N,$$

$$\frac{\partial F(x)}{\partial x_i \partial x_j} = \begin{cases} -\lambda_2 \dfrac{1}{x_i}, & \text{if } j = 1, \\ 0, & \text{if } j \neq i. \end{cases} \qquad i, j \in N.$$

Let $v = (v_1, \ldots, v_n)^t$ be a vector with $v_i \in \mathbb{R}$, $i \in N$. We have

$$v^t \cdot \nabla^2(F) \cdot v = \sum_{i=1}^{n} v_i^2 (\nabla^2(F))_{ii} = -\sum_{i=1}^{n} v_i^2 \cdot \lambda_2 \cdot \frac{1}{x_i} < 0$$

for $v \neq 0$. It follows that the objective function of problem (P) is strictly concave. Since its domain is compact and convex, we obtain that (P) admits a unique optimal solution. □

Lemma 2.1 *For any $x \in X$, $\lambda_1, \lambda_2 > 0$ and for any $u_i \in \mathbb{R}_+$, $i \in N$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ the following inequality holds*

$$\begin{aligned} &\lambda_1 \sum_{i=1}^{n} u_i x_i - \lambda_2 \sum_{i=1}^{n} x_i \ln x_i \leq \\ &\leq \lambda_2 \sum_{i=1}^{n} \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki}\right) - 1\right) - \sum_{k=1}^{m} d_k y_k. \end{aligned} \tag{2.2}$$

Proof. Let us apply the inequality $\ln z \leq z - 1$, valid for every $z > 0$ with equality if and only if $z = 1$, to the positive numbers a_i, $i \in N$, defined by

$$a_i = \frac{1}{x_i} \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki}\right) - 1\right).$$

It follows that

$$x_i \left(\sum_{k=1}^{m} y_k w_{ki} + \lambda_1 u_i\right) - \lambda_2 \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki}\right) - 1\right) \leq \lambda_2 x_i \ln x_i.$$

Summing by i, $i \in N$, one obtains

$$\begin{aligned} &\sum_{i=1}^{n} x_i \left(\sum_{k=1}^{m} y_k w_{ki}\right) + \lambda_1 \sum_{i=1}^{n} u_i x_i - \\ &-\lambda_2 \sum_{i=1}^{n} \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki}\right) - 1\right) \leq \lambda_2 \sum_{i=1}^{n} x_i \ln x_i. \end{aligned} \tag{2.3}$$

The inequality (2.3) is equivalent to (2.2) because since $W \cdot x = d$, we can write

$$\sum_{i=1}^{n} x_i \left(\sum_{k=1}^{m} y_k w_{ki}\right) = \sum_{k=1}^{m} \left(\sum_{i=1}^{n} w_{ki} x_i\right) y_k = \sum_{k=1}^{m} d_k y_k. \ \square$$

Using Lemma 2.1 we can associate to problem (P) with fixed λ_1, λ_2 the following unconstrained dual problem

$$\text{(D)} \qquad \begin{aligned} &\min_{y \in \mathbb{R}^m} G(y) = \\ &= \min_{y \in \mathbb{R}^m} \left\{ \lambda_2 \sum_{i=1}^{n} \exp\left(\lambda_2^{-1} \left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki} \right) - 1 \right) - \sum_{k=1}^{m} d_k y_k \right\}. \end{aligned}$$

Hypothesis H1. The matrix W has full row–rank.
Hypothesis H2. Int $X \neq \emptyset$.

Theorem 2.2 *In the hypotheses* H1 *and* H2, *the problem* (D) *admits only one optimal solution* $y^0 = (y_1^0, \ldots, y_m^0) \in \mathbb{R}^m$ *for any fixed pair* $\lambda_1, \lambda_2 > 0$. *Moreover, its corresponding primal problem* (P) *admits the optimal solution* $x^0 = (x_1^0, \ldots, x_n^0)^t$, *where* $x_i^0 = \exp\left(\lambda_2^{-1} \left(\lambda_1 u_i + \sum_{k=1}^{m} y_k^0 w_{ki} \right) - 1 \right)$, $i \in N$, *and the optimal values of the two dual problems coincide.*

Proof. We show that in the hypothesis H1 the Hessian matrix of $G(y)$, denoted by $\nabla^2(G)$, is positively defined. We have

$$\frac{\partial G(y)}{\partial y_j} = -d_j + \sum_{i=1}^{n} w_{ji} \exp\left(\lambda_2^{-1} \left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki} \right) - 1 \right);$$

$$\frac{\partial^2 G(y)}{\partial y_j \partial y_\ell} = \frac{1}{\lambda_2} \sum_{i=1}^{n} w_{ji} w_{\ell i} \exp\left(\lambda_2^{-1} \left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki} \right) - 1 \right).$$

For $v = (v_1, \ldots, v_m)^t \in \mathbb{R}^m$, we have

$$v^t \nabla^2(G) v = \frac{1}{\lambda_2} \sum_{i=1}^{n} \exp\left(\lambda_2^{-1} \left(\lambda_1 u_i + \sum_{k=1}^{m} y_k w_{ki} \right) - 1 \right) \cdot \left(\sum_{k=1}^{m} v_k w_{ki} \right) > 0,$$

for $v \neq 0$, if we take into account Hypothesis H1.

Since the Hessian matrix of $G(y)$ is positively defined according to a characterization theorem of a first degree differentiable function on an open convex domain one obtains that $G(y)$ is strictly convex, therefore problem (D) has no more than an optimal solution.

¿From Hypothesis H2 and from the Fenchel's theorem (Rockafellar [8]), we have that between the two dual problems (P) and (D) there is no duality gap. According to Theorem 2.1, the problem (P) has always a finite optimum and since $G(y)$ is strictly convex it follows that the problem (D) admits only one optimal solution $y^0 = (y_1^0, ..., y_m^0)$. Since $G(y)$ is of the class $C^2(\mathbb{R}^m)$ we deduce that the first order optimality conditions hold at y^0, that is

$$\sum_{i=1}^{n} w_{ki} \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k^0 w_{ki}\right) - 1\right) = d_k, \ k = 1, ..., m. \tag{2.4}$$

If we use the notation $x_i^0 = \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k^0 w_{ki}\right) - 1\right) = d_k, \ i \in N,$ we note that the equality (2.4) means $W \cdot x^0 = d$, so $x^0 \in X$. We apply now the Fenchel's duality theorem. □

Theorem 2.3 *For any fixed* $\lambda_1, \lambda_2 > 0$ *and any* $u_i \in \mathbb{R}_+$, $i \in N$,

(i) (*weak duality*) $\max_{x \in X} F(x) \leq \min_{y \in \mathbb{R}^m} G(y)$;

(ii) (*strong duality*) *In Hypotheses* H1 *and* H2 *the problem* (D) *has a unique optimal solution* $y^0 \in \mathbb{R}^m$ *and* $x^0 \in X$ *given by*

$$x_i^0 = \exp\left(\lambda_2^{-1}\left(\lambda_1 u_i + \sum_{k=1}^{m} y_k^0 w_{ki}\right) - 1\right), \ i \in N,$$

is the optimal solution of Problem (P)*; moreover,* $F(x^0) = G(y^0)$.

Proof. The conclusion of (i) follows from Lemma 2.1 and the form of the dual problem (D).

The conclusion of (ii) follows from Theorem 2.2. □

Now we define a bijection $B : N \times N \longrightarrow N$; any previous simple sum may be written then as a double sum based on the following relation

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = \sum_{\ell=1}^{n} x_{B^{-1}(\ell)}.$$

We can associate to Problems (P) and (D), on the base of the definition of the bijection B, the following pair of dual problems:

$$\text{(P)}' \qquad \max_{x\in X'} F'(x) = \max_{x\in X'} \left\{ \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} u_{ij}x_{ij} - \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln x_{ij} \right\},$$

where $X' = \left\{ x = (x_{ij})^t_{i,j\in N} \mid x_{ij} > 0, i,j\in N, \sum_{i=1}^{n}\sum_{j=1}^{n} w_{k,ij}x_{ij} = d_k,\ k=1,\dots,m \right\};$

$$\text{(D)}' \qquad \begin{aligned} &\min_{y\in \mathbb{R}^m} G'(y) = \\ &= \min_{y\in \mathbb{R}^m} \left\{ \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} \exp\left(\lambda_2^{-1} \left(\lambda_1 u_{ij} + \sum_{k=1}^{m} y_k w_{k,ij} \right) - 1 \right) - \sum_{k=1}^{m} d_k y_k \right\}. \end{aligned}$$

Theorem 2.4 *For any fixed* $\lambda_1, \lambda_2 > 0$ *and* $u_{ij} \in \mathbb{R}_+,\ i,j\in N,$

(i) (*weak duality*) $\max_{x\in X'} F'(x) \le \min_{y\in\mathbb{R}^m} G'(y)$;

(ii) (*strong duality*) *If the matrix* W *has full row-rank and* int $X' \neq \emptyset$, *then the problem* (D)$'$ *has a unique optimal solution* $y^0 \in \mathbb{R}^m$ *and* $x^0 = (x^0_{ij})^t_{i,j\in N} \in X'$ *given by*

$$x^0_{ij} = \exp\left(\lambda_2^{-1} \left(\lambda_1 u_{ij} + \sum_{k=1}^{m} y^0_k w_{k,ij} \right) - 1 \right),\quad i,j\in N$$

is the optimal solution of Problem (P)$'$*; moreover,* $F'(x^0) = G'(y^0)$.

Proof. The conclusion follows immediately from Theorem 2.3. □

3 Optimal transition probabilities

The optimal transition probabilities from a state of the system to another one appear as the optimal solution of the optimization problem having its objective function defined by (1.5) with $U(p)$ and $H(p)$ given by (1.3), (1.4)

$$\text{(P1)} \qquad \max_{p\in\mathcal{P}} F(p) = \max_{p\in\mathcal{P}} \left\{ \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} u_{ij}p_{ij} - \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} u_i\pi_i p_{ij}\ln p_{ij} \right\}.$$

We shall determine the optimal solution of this problem by a dual approach using the results obtained in the previous section regarding the problems (P)$'$ and (D)$'$ expressed by Theorem 2.4.

We apply to Problem (P1) the transformation defined by

$$x_{ij} = \frac{u_i\pi_i}{c} p_{ij}, \; i,j \in N, \text{ where } c = \sum_{i=1}^{n} u_i\pi_i. \tag{3.1}$$

We use the following notations:

$$\begin{gathered} x = (x_{ij})^t_{i,j\in N}, \; \widetilde{H}(x) = -\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln x_{ij}, \; v_i = \frac{u_i\pi_i}{c}, \; i \in N, \\ v = (v_1, ..., v_n) \text{ and } H(v) = -\sum_{i=1}^{n} v_i \ln v_i. \end{gathered} \tag{3.2}$$

Lemma 3.1 *For any $p \in \mathcal{P}$ and x defined by* (3.1) *the equality*

$$H(p) = c\widetilde{H}(x) - cH(v) \tag{3.3}$$

holds, where the notations (3.2) *are used.*

Proof. Indeed,

$$\begin{gathered} H(p) = -\sum_{i=1}^{n}\sum_{j=1}^{n} u_i\pi_i p_{ij}\ln p_{ij} = -\sum_{i=1}^{n}\sum_{j=1}^{n} u_i\pi_i p_{ij} \ln \frac{u_i\pi_i p_{ij}}{c} + \\ +\sum_{i=1}^{n}\sum_{j=1}^{n} u_i\pi_i p_{ij}\ln\frac{u_i\pi_i}{c} = -c\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln x_{ij} + c\sum_{i=1}^{n}\frac{u_i\pi_i}{c}\ln\left(\frac{u_i\pi_i}{c}\right) = \\ = c\widetilde{H}(x) - cH(v). \; \square \end{gathered}$$

Lemma 3.2 *The transformation induced by v_i, $i \in N$, by the formulae*

$$\bar{w}_{k,ij} = \frac{1}{v_i} w_{k,ij}; \; \bar{u}_{ij} = \frac{1}{v_i} u_{ij}, \; i,j \in N, \; k = 1, ..., m \tag{3.4}$$

has the following properties:

$$\begin{cases} \bar{w}_{k,ij} x_{ij} = w_{k,ij} p_{ij}, & i,j \in N, \; k = 1, ..., m; \\ \bar{u}_{ij} x_{ij} = u_{ij} p_{ij}, & i,j \in N. \end{cases} \tag{3.5}$$

Proof. Indeed, for $k = 1, ..., m$

$$(\widetilde{W}x)_k = \sum_{i=1}^{n}\sum_{j=1}^{n} \widetilde{w}_{k,ij} x_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{c}{u_i\pi_i} w_{k,ij}\right) \cdot \frac{u_i\pi_i p_{ij}}{c} = (Wp)_k.$$

Therefore, $\widetilde{W}x = Wp$. Similarly, we prove the second relation (3.4). □

¿From Lemma 3.2 it follows that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \widetilde{w}_{k,ij} x_{ij} = d_k, \;\; k = 1, ..., m \tag{3.6}$$

that implies $\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = 1$. According to Lemmas 3.1 and 3.2, one obtains

$$\begin{aligned} &\lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} u_{ij} p_{ij} - \lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} u_i \pi_i p_{ij} \ln p_{ij} = \\ &= \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{u}_{ij} x_{ij} - c\lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \ln x_{ij} - c\lambda_2 H(v). \end{aligned} \tag{3.7}$$

We remark that $c\lambda_2 H(v)$ from (3.7) is a constant. Let us note $\bar{\lambda}_2 = c\lambda_2$ and $\widetilde{X} = \{x = (x_{ij})^t_{i,j\in N} \mid x_{ij} > 0, \text{ with } x \text{ satisfying (3.6)}\}$.

Let us consider the optimization problem in the variable $x = (x_{ij})^t_{i,j\in N}$

$$\max_{x\in\widetilde{X}} \left\{ \lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{u}_{ij} x_{ij} - \bar{\lambda}_2 \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \ln x_{ij} \right\}. \tag{$\bar{\text{P}}1$}$$

The transformation defined by (3.1) and (3.4) assures, according to (3.7), the equivalence of the problems (P1) and ($\bar{\text{P}}$1).

It is obvious that p is an admissible solution for the problem (P1) if and only if x defined by (3.1) is an admissible solution for the problem ($\bar{\text{P}}$1).

Now, taking into account the existent analogy between the problems (P)′ and ($\bar{\text{P}}$1), the dual of the problem ($\bar{\text{P}}$1) is

$$\min_{y\in\mathbb{R}^m} \left\{ \bar{\lambda}_2 \sum_{i=1}^{n}\sum_{j=1}^{n} \exp\left(\bar{\lambda}_2^{-1}\left(\lambda_1 \bar{u}_{ij} + \sum_{k=1}^{m} y_k \widetilde{w}_{k,ij}\right) - 1\right) - \sum_{k=1}^{m} d_k y_k \right\}. \tag{$\bar{\text{D}}1$}$$

According to Theorem 2.4, if the objective function of problem $(\tilde{\mathrm{D}}1)$ gets its minimum in $y^0 \in \mathbb{R}^m$ and if x^0 given by

$$x_{ij}^0 = \exp\left(\tilde{\lambda}_2^{-1}\left(\lambda_1 \tilde{u}_{ij} + \sum_{k=1}^{m} y_k^0 \tilde{w}_{k,ij}\right) - 1\right), \quad i,j \in N,$$

belongs to $\widetilde{X}$ then $x^0 = (x_{ij}^0)_{i,j\in N}^t$ is an optimal solution for the problem $(\tilde{\mathrm{P}}1)$ and the optimal values of the objective functions of the problems $(\tilde{\mathrm{P}}1)$ and $(\tilde{\mathrm{D}}1)$ coincide.

¿From this duality and the relations (3.1), (3.3), (3.4), (3.5) it follows that the dual of the problem (P1) is

$$\min_{y\in\mathbb{R}^m} G(y) =$$

$$\text{(D1)} \qquad = \min_{y\in\mathbb{R}^m}\left\{c\lambda_2 \sum_{i=1}^{n}\sum_{j=1}^{n} \exp\left((u_i\pi_i\lambda_2)^{-1}\left(\lambda_1 u_{ij} + \sum_{k=1}^{m} y_k w_{k,ij}\right) - 1\right) - \right.$$
$$\left. - \sum_{k=1}^{m} d_k y_k + c\lambda_2 \sum_{i=1}^{n} \frac{u_i\pi_i}{c} \ln \frac{u_i\pi_i}{c}\right\}.$$

The duality of the problems (P1) and (D1) is given by

Theorem 3.1 *If the matrix* $W = (w_{k,ij})_{i,j\in N}$, $k = 1, ..., m$ *has full row–rank and* int $\mathcal{P} \neq \emptyset$ *then*

(i) (weak duality) *If* $p = (p_{ij})_{i,j\in N}^t$, $y = (y_1, ..., y_m) \in \mathbb{R}^m$ *are admissible solutions for the problems* (P1), *respectively* (D1), *then* $F(p) \leq G(y)$*;*

(ii) (strong duality) *If* $\min_{y\in\mathbb{R}^m} G(y)$ *is reached in* $y^0 \in \mathbb{R}^m$ *and* p^0 *given by*

$$\text{(3.8)} \qquad p_{ij}^0 = \frac{c}{u_i\pi_i} \exp\left((u_i\pi_i\lambda_2)^{-1}\left(\lambda_1 u_{ij} + \sum_{k=1}^{m} y_k^0 w_{k,ij}\right) - 1\right), \quad i,j \in N,$$

belongs to int $\mathcal{P}$, *then* p^0 *is an optimal solution for the problem* (P1) *and* $F(p^0) = G(y^0)$.

Remark 3.1 The formula (3.8) from Theorem 3.1 defines the optimal transition way of a random behaviour system from any of its states to another one.

4 Conclusion

In this paper we have studied the problem of the optimal transition from a state to another one into a random behaviour system having a finite number of states by means of entropic criteria. We obtained the optimal transition probabilities in the hypothesis that these probabilities satisfy a set of explicit linear constraints of the equality type. The case without explicit constraints is a particular one, the corresponding optimal transition probabilities being obtained from (3.8) by taking for $i, j \in N$, $w_{k,ij} = 1$, $d_k = 1$ for $k = 1, ..., n$ and $w_{k,ij} = \pi_i$, $d_k = \pi_j$ for $k = n + 1, ..., 2n$.

The pairs of dual problems considered here differ in their form and number from those used in Gerchak [6], Erlander [4], ours being extensions of them. Our results can be applied in every real life field where a random behaviour system with a finite number of states is identified for which it is important to maximize simultanously the expected utility and the entropy associated to transition from a state to another one during the time.

References

[1] A. Ben–Tal, A. Charnes, *A Dual Optimization Framework for Some Problems of Information Theory and Statistics*, Problems of Control and Information Theory 8(1979), 387–401.

[2] A. Ben–Tal, M. Teboulle, A. Charnes, *The Role of Duality in Optimization Problems Involving Entropy Functionals with Applications to Information Theory*, Journal of Optimization Theory and Applications, 58/2(1988), 209–223.

[3] A. Charnes, W.W. Cooper, L. Seiford, *Extremal Principles and Optimization Qualities for Khinchin–Kullback–Leibler Estimation*, Statistics (Zentralinstitut für Mathematik und Mechanik), 9(1978), 21–29.

[4] S. Erlander, *Entropy in Linear Programs*, Mathematical Programming 21(1981), 137–151.

[5] S.C. Fang, H.S.J. Tsao, *Linear Programming with Entropic Perturbation*, ZOR – Methods and Models of Operations Research 37 (1993), 171–186.

[6] Y. Gerchak, *Maximal Entropy of Markov Chains with Common Steady–State Probabilities*, Journal of the Operational Research Society, vol. 32 (1991), 233–234.

[7] S. Guiaşu, A. Shenitzer, *The Principle of Maximum Entropy*, Mathematical Intelligencer, 8 (1985), 42–48.

[8] R.T. Rockafellar, *Convex Analysis,* Princeton University Press, New Jersey, 1970.

[9] H.S. Tsao, S.C. Fang, D.N. Lee, *On the Optimal Entropy Analysis,* European Journal of Operational Research 59 (1992), 324–329.

Rodica Brânzei, Al.I.Cuza University, Faculty of Computer Science, Bd.Copou 11, 6600 Iaşi, România, e-mail: branzeir@thor.infoiasi.ro

Received February 23, 1998