Bull. Math. Soc. Sci. Math. Roumanie
Tome 50(98) No. 4, 2007, 295–304

On maxentropic reconstruction of multiple Markov chains

by
Costel Bălcău and Vasile Preda

Abstract

We use the maxentropic reconstruction method for obtaining some homogeneous and stationary multiple Markov chains.

Key Words: Homogeneous and stationary multiple Markov chain, maximum entropy.
2000 Mathematics Subject Classification: Primary 94A17, Secondary: 90C25.

1 Introduction

The problems of reconstruction of a countable probability distribution or of a homogeneous and stationary simple or multiple Markov chain with discrete time and countable state space, when only a partial information is know, present a remarkable interest in many applications from various sciences (see Fang, Rajasekera and Tsao [2], Guiaşu [3], Iosifescu [7], Iosifescu and Grigorescu [8], for example). Jaynes [10, 11] and Kullback [12] proposed a variational method for solving such problems, called the *method of standard maximum entropy* (*SME*). This method states that one should choose the probability distribution that maximizes the Shannon entropy [16]. In the case of multiple Markov chains, we can use the Iosifescu-Theodorescu entropy [9] (see Preda and Bălcău [14]).

In this article we recast the proposed problem as a linear inverse problem and we solve it by using the *method of maximum entropy in the mean* (*MEM*). This method, originally proposed by Rietch [15], was used by Gzyl [4] for finite probability distributions, Gzyl and Velásquez [6] for homogeneous simple Markov chains with finite state space, Preda and Bălcău [13] for countable simple Markov chains with common steady-state probabilities and for matrix scaling problems.

In Section 2 we present the general frame of the MEM method and we apply this method for maxentropic reconstruction of some nonnegative multidimensional matrices. In Section 3 we derive a method to obtain countable multiple Markov chains with fixed joint probabilities. We give an example for our approach.

2 MEM method

We consider the following linear inverse problem

$$\begin{aligned} &\mathrm{Ax} = \mathrm{b}, \\ &\mathrm{x} \in \mathcal{C}, \end{aligned} \tag{1}$$

where

$$\mathcal{C} = \left\{ \mathrm{x} = (x_i)_{i\in I} \ / \ x_i \geq 0, \ \forall i \in I, \ \sum_{i\in I} x_i < \infty \right\}, \tag{2}$$

$\mathrm{A} : \mathcal{C} \to \mathbb{R}^K$ is a bounded linear operator, $\mathrm{b} \in \mathbb{R}^K$, I and K being two countable (finite or infinite) sets of indices.

Let $(\Omega, \mathcal{B}, \nu)$ be a probability space, and let $\mathrm{X} : (\Omega, \mathcal{B}) \to (\mathcal{C}, \mathcal{B}(\mathcal{C}))$ be a random variable. Usually, $\Omega = \mathcal{C}$ and we suppose that

$$\overline{\mathrm{co}}\,(\operatorname{supp}\nu) = \mathcal{C},$$

where $\operatorname{supp}\nu$ is the support of ν and $\overline{\mathrm{co}}\,(\operatorname{supp}\nu)$ is the closed convex hull generated by $\operatorname{supp}\nu$.

If the set K is infinite, then we suppose also that

$$\Omega \text{ is countable, b is bounded, } \mathrm{A} \geq 0 \text{ and } \mathrm{b} \geq 0. \tag{3}$$

Definition 2.1. *Let $\mathcal{M}_0$ be the set of all probability measures on $(\Omega, \mathcal{B})$, and let*

$$\mathcal{M}(\nu, \mathrm{A}, \mathrm{b}) = \{\mu \in \mathcal{M}_0 \ / \ \mu \prec \nu, \ \mathrm{A}E_\mu[\mathrm{X}] = \mathrm{b}\}.$$

For all $\mu \in \mathcal{M}(\nu, \mathrm{A}, \mathrm{b})$, let

$$H(\mu; \nu) = \begin{cases} \displaystyle\int_\Omega \frac{d\mu}{d\nu} \ln \frac{d\mu}{d\nu}\, d\nu, & \textit{if } E_\mu\left[\left|\ln \frac{d\mu}{d\nu}\right|\right] < \infty, \\ +\infty, & \textit{otherwise} \end{cases}$$

(the relative entropy of μ with respect to ν; the cross-entropy of ν with respect to μ; the Kullback-Leibler number).

We consider the following entropy optimization problem, according to SME method:

$$\text{Program } (P1) : \left| \begin{array}{l} \max -H(\mu; \nu) \quad \text{s.t.} \\ \mu \in \mathcal{M}(\nu, \mathrm{A}, \mathrm{b}). \end{array} \right.$$

MEM method is based on the following result (see [5] or [1], Theorem 1):

Theorem 2.1. *If μ is a feasible solution for program $(P1)$, then $\mathrm{x} = E_\mu[\mathrm{X}]$ is a solution of linear inverse problem (1).*

Next, we derive a convex dual program of program $(P1)$.

Definition 2.2. *For all $\lambda \in \mathbb{R}^K$, let the measure $\mu(\lambda)$ defined by*

$$d\mu(\lambda) = \frac{e^{-\langle \lambda, \mathrm{AX} \rangle}}{Z(\lambda)}\, d\nu, \ \textit{where } Z(\lambda) = \int_\Omega e^{-\langle \lambda, \mathrm{AX} \rangle} d\nu.$$

Definition 2.3. *Let the set*

$$\Lambda(\nu, \mathrm{A}, \mathrm{b}) = \begin{cases} \{\lambda \in \mathbb{R}^K \ / \ Z(\lambda) < \infty\}, & \textit{if } K \textit{ is finite,} \\ \{\lambda \in \mathbb{R}_+^K \ / \ Z(\lambda) < \infty, \ \sum\limits_{k \in K} \lambda_k < \infty\}, & \textit{if } K \textit{ is infinite,} \end{cases}$$

and let the function $L : \Lambda(\nu, \mathrm{A}, \mathrm{b}) \to \mathbb{R}$ *defined by*

$$L(\lambda) = \ln Z(\lambda) + \langle \lambda, \mathrm{b} \rangle, \ \forall \lambda \in \Lambda(\nu, \mathrm{A}, \mathrm{b}).$$

Remark 2.1. *For all* $\lambda \in \Lambda(\nu, \mathrm{A}, \mathrm{b})$*,* $\mu(\lambda)$ *is a probability measure on* $(\Omega, \mathcal{B})$ *(i.e.* $\mu(\lambda) \in \mathcal{M}_0$*) and* $\mu(\lambda) \prec \nu$*.*

We can define the following geometric dual problem for program $(P1)$:

$$\text{Program } (D1) : \left| \begin{array}{l} \min L(\lambda) \quad \text{s.t.} \\ \lambda \in \Lambda(\nu, \mathrm{A}, \mathrm{b}). \end{array} \right.$$

We have the following duality theorem (see [13], Theorems 1 and 2):

Theorem 2.2. *(i)* **(weak duality)** *If* μ *is a primal feasible solution of program* $(P1)$ *and* λ *is a dual feasible solution of program* $(D1)$*, then*

$$-H(\mu; \nu) \leq L(\lambda).$$

Moreover, the equality holds if and only if

$$d\mu = \frac{e^{-\langle \lambda, \mathrm{AX} \rangle}}{Z(\lambda)} \, d\nu.$$

(ii) **(strong duality)** *Assume that*

$$\int_\Omega e^{s \|\mathrm{X}\|} d\nu < \infty, \ \forall s \in \mathbb{R}. \tag{4}$$

If $\lambda^* \in \operatorname{Int} \Lambda(\nu, \mathrm{A}, \mathrm{b})$ *is a dual optimal solution of program* $(D1)$*, then* $\mu(\lambda^*)$ *is a primal optimal solution of program* $(P1)$ *and the duality gap vanishes, i.e.*

$$-H(\mu(\lambda^*); \nu) = L(\lambda^*).$$

Remark 2.2. *The assumptions (3),* $\lambda \geq 0$*,* $\sum\limits_{k \in K} \lambda_k < \infty$ *and (4) are imposed for proving the theorem in the countable infinite case (see [13]).*

Next we apply the MEM method to obtain d-dimensional matrices

$$\mathrm{T} = (T_{i_1, \ldots, i_d})_{i_1, \ldots, i_d \in J}$$

verifying the constraints similar to (1). Particularly, in Section 3 we apply the procedure to obtain countable Markov chains of order r with fixed joint probabilities of the states at r consecutive times and with r-step transition probabilities verifying some given linear constraints.

Let $d \in \mathbb{N}^*$, J be a countable (finite or infinite) set of indices, and let

$$I = J^d \ , \ \Omega = \prod_{(i_1,\ldots,i_d)\in I} \Omega_{i_1,\ldots,i_d},$$

$$\mathcal{B} = \bigotimes_{(i_1,\ldots,i_d)\in I} \mathcal{B}(\Omega_{i_1,\ldots,i_d}), \ \nu = \bigotimes_{(i_1,\ldots,i_d)\in I} \nu_{i_1,\ldots,i_d},$$

where, for all $(i_1, \ldots, i_d) \in I$, $(\Omega_{i_1,\ldots,i_d}, \mathcal{B}(\Omega_{i_1,\ldots,i_d}), \nu_{i_1,\ldots,i_d})$ is a probability space.

Taking $\mu = \bigotimes_{(i_1,\ldots,i_d)\in I} \mu_{i_1,\ldots,i_d}$, where, for all $(i_1, \ldots, i_d) \in I$, $\mu_{i_1,\ldots,i_d}$ is a probability measure on the space $(\Omega_{i_1,\ldots,i_d}, \mathcal{B}(\Omega_{i_1,\ldots,i_d}))$, the primal problem $(P1)$ has now the following form:

$$\text{Program } (P2): \left| \begin{array}{l} \max -H(\mu;\nu) \quad \text{s.t.} \\ \mu_{i_1,\ldots,i_d} \prec \nu_{i_1,\ldots,i_d}, \ \forall (i_1,\ldots,i_d) \in I, \\ \displaystyle\sum_{(i_1,\ldots,i_d)\in I} A^{(k)}_{i_1,\ldots,i_d} E_{\mu_{i_1,\ldots,i_d}}[X_{i_1,\ldots,i_d}] = b_k, \ \forall k \in K. \end{array} \right.$$

According to the above assumptions, K is a countable set of indices, $\mathrm{X} = (X_{i_1,\ldots,i_d})_{(i_1,\ldots,i_d)\in I}$ is a random variable on $(\Omega, \mathcal{B})$ with values in the space $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$ defined by

$$\mathcal{C} = \left\{ \mathrm{x} = (x_{i_1,\ldots,i_d})_{(i_1,\ldots,i_d)\in I} \ / \ \mathrm{x} \geq 0, \sum_{(i_1,\ldots,i_d)\in I} x_{i_1,\ldots,i_d} < \infty \right\},$$

$\mathrm{A} : \mathcal{C} \to \mathbb{R}^K$ is a bounded linear operator, $\mathrm{b} \in \mathbb{R}^K$, and the conditions (3) are also satisfied.

The maxentropic reconstruction of some d-dimensional matrices

$$\mathrm{T} = (T_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \in J}$$

that verify the following linear system

$$\begin{array}{l} \mathrm{AT} = \mathrm{b}, \\ \mathrm{T} \in \mathcal{C}, \end{array} \tag{5}$$

is based on the following direct consequence of Theorem 2.1.

Corollary 2.1. *If μ is a feasible solution for program $(P2)$, then the matrix $\mathrm{T} = (T_{i_1,\ldots,i_d})_{(i_1,\ldots,i_d)\in I}$ given by*

$$T_{i_1,\ldots,i_d} = E_{\mu_{i_1,\ldots,i_d}}[X_{i_1,\ldots,i_d}], \ \forall (i_1,\ldots,i_d) \in I$$

is a d-dimensional matrix which verifies (5).

Next, we derive a geometric dual program of program $(P2)$. We have

$$Z(\lambda) = \int_{\Omega} e^{-\langle \lambda, \mathrm{AX} \rangle} d\nu = \prod_{(i_1,\dots,i_d)\in I} \zeta_{i_1,\dots,i_d}((\mathrm{A}^*\lambda)_{i_1,\dots,i_d}),$$

where

$$\zeta_{i_1,\dots,i_d}(y) = \int_{\Omega_{i_1,\dots,i_d}} e^{-y x_{i_1,\dots,i_d}}\, d\nu_{i_1,\dots,i_d}(x_{i_1,\dots,i_d}).$$

The dual problem for program $(P2)$ has the following form:

$$\text{Program } (D2): \left| \begin{array}{l} \min L(\lambda) = \displaystyle\sum_{(i_1,\dots,i_d)\in I} \ln \zeta_{i_1,\dots,i_d}((\mathrm{A}^*\lambda)_{i_1,\dots,i_d}) + \langle \lambda, \mathrm{b} \rangle \quad \text{s.t.} \\ \lambda \in \Lambda(\nu, \mathrm{A}, \mathrm{b}), \end{array} \right.$$

where $\Lambda(\nu, \mathrm{A}, \mathrm{b})$ is given by Definition 2.3.

As a direct consequence of Theorem 2.2, we have the next duality result.

Corollary 2.2. *ssume that* $\displaystyle\sum_{(i_1,\dots,i_d)\in I} \int_{\Omega_{i_1,\dots,i_d}} X_{i_1,\dots,i_d} d\nu_{i_1,\dots,i_d} < \infty.$

(i) **(weak duality)** *If μ is a primal feasible solution of program $(P2)$ and λ is a dual feasible solution of program $(D2)$, then*

$$-H(\mu;\nu) \leq L(\lambda).$$

Moreover, the equality holds if and only if

$$d\mu_{i_1,\dots,i_d} = \frac{e^{-(\mathrm{A}^*\lambda)_{i_1,\dots,i_d} X_{i_1,\dots,i_d}}}{\zeta_{i_1,\dots,i_d}((\mathrm{A}^*\lambda)_{i_1,\dots,i_d})}\, d\nu_{i_1,\dots,i_d},\ \forall (i_1,\dots,i_d) \in I.$$

(ii) **(stron duality)** *If $\lambda^* \in \mathrm{Int}\,\Lambda(\nu, \mathrm{A}, \mathrm{b})$ is a dual optimal solution of program $(D2)$, then $\mu(\lambda^*) = \displaystyle\bigotimes_{(i_1,\dots,i_d)\in I} \mu_{i_1,\dots,i_d}(\lambda^*)$ given by*

$$d\mu_{i_1,\dots,i_d}(\lambda^*) = \frac{e^{-(\mathrm{A}^*\lambda^*)_{i_1,\dots,i_d} X_{i_1,\dots,i_d}}}{\zeta_{i_1,\dots,i_d}((\mathrm{A}^*\lambda^*)_{i_1,\dots,i_d})}\, d\nu_{i_1,\dots,i_d},\ \forall (i_1,\dots,i_d) \in I$$

is a primal optimal solution of program $(P2)$ and the duality gap vanishes, i.e.

$$-H(\mu(\lambda^*);\nu) = L(\lambda^*).$$

3 Reconstruction of some countable multiple Markov chains

In this section we apply the above results to obtain homogeneous and stationary Markov chains of order r ($r \in \mathbb{N}^*$) $\{X(t)\ /\ t \in \mathbb{N}\}$ with the countable (finite or infinite) state space J and with r-step transition probability array $\mathrm{P}^{(r)} =$

$\left(P^{(r)}_{i_1,\ldots,i_r;j}\right)_{i_1,\ldots,i_r,j\in J}$ verifying the constraints (5) for $d = r+1$. We recall that for any $j, i_1, \ldots, i_r \in I$

$$P^{(r)}_{i_1,\ldots,i_r;j} = P(X(t+r) = j / X(t) = i_1, \ldots, X(t+r-1) = i_r),\ \forall t \in \mathbb{N}.$$

Particularly, we obtain homogeneous and stationary Markov chains of order r without supplementary constraints, i.e. the r-step transition probability array verifying only the imposed constraints

$$\left\{\begin{array}{rl} P^{(r)}_{i_1,\ldots,i_r;j} & \geq 0,\ \forall j, i_1, \ldots, i_r \in J, \\ \sum\limits_{j\in J} P^{(r)}_{i_1,\ldots,i_r;j} & = 1,\ \forall i_1, \ldots, i_r \in J, \\ \sum\limits_{i_1,\ldots,i_r\in J} \pi^{(r)}_{i_1,\ldots,i_r} P^{(r)}_{i_1,\ldots,i_r;j} & = \sum\limits_{i_1,\ldots,i_{r-1}\in J} \pi^{(r)}_{i_1,\ldots,i_{r-1},j},\ \forall j \in J, \end{array}\right. \tag{6}$$

where $\pi^{(\mathrm{r})} = \left(\pi^{(r)}_{i_1,\ldots,i_r}\right)_{i_1,\ldots,i_r\in J}$ is the given joint probability of the states at r consecutive times, i.e. for any $i_1, \ldots, i_r \in I$

$$\pi^{(r)}_{i_1,\ldots,i_r} = P(X(t) = i_1, \ldots, X(t+r-1) = i_r),\ \forall t \in \mathbb{N}.$$

Remark 3.1. *Obviously, the chain $\{X(t) \ /\ t \in \mathbb{N}\}$ is completely characterized by the distribution $\pi^{(\mathrm{r})}$ and the transition probability array $\mathrm{P}^{(r)}$ verifying the constraints (6) and the following constraints*

$$\pi^{(r)}_{i_1,\ldots,i_r} > 0,\ \forall i_1, \ldots, i_r \in I, \tag{7}$$

$$\sum_{i_1,\ldots,i_r\in I} \pi^{(r)}_{i_1,\ldots,i_r} = 1, \tag{8}$$

$$\sum_{i_1,\ldots,i_k\in I} \pi^{(r)}_{i_1,\ldots,i_k,i_{k+1},\ldots,i_r} = \sum_{i_1,\ldots,i_k\in I} \pi^{(r)}_{i_{k+1},\ldots,i_r,i_1,\ldots,i_k},\ \forall 1 \leq k \leq r-1. \tag{9}$$

In this particular case $\mathrm{P}^{(r)}$ is a solution of (5) by taking

$$\begin{aligned} d &= r+1, \\ K &= J^r \cup (-J), \\ A^{(k_1,\ldots,k_r)}_{i_1,\ldots,i_r,j} &= \left\{\begin{array}{ll} 1, & \text{if } (i_1,\ldots,i_r) = (k_1,\ldots,k_r), \\ 0, & \text{if } (i_1,\ldots,i_r) \neq (k_1,\ldots,k_r), \end{array}\right. \quad \forall i_1,\ldots,i_r,j,k_1,\ldots,k_r \in J, \\ A^{(-k)}_{i_1,\ldots,i_r,j} &= \left\{\begin{array}{ll} \pi^{(r)}_{i_1,\ldots,i_r}, & \text{if } j = k, \\ 0, & \text{if } j \neq k, \end{array}\right. \quad \forall i_1,\ldots,i_r,j,k \in J, \\ b_{k_1,\ldots,k_r} &= 1,\ \forall k_1,\ldots,k_r \in J, \\ b_{-k} &= \pi_k,\ \forall k \in J, \end{aligned}$$

where

$$\pi_k = \sum_{i_1,\ldots,i_{r-1}\in J} \pi^{(r)}_{i_1,\ldots,i_{r-1},k},\ \forall k \in J.$$

The primal problem $(P2)$ has now the following form:

$$\text{Program } (P3): \left| \begin{array}{l} \max -H(\mu;\nu) \quad \text{s.t.} \\ \mu_{i_1,\ldots,i_r,j} \prec \nu_{i_1,\ldots,i_r,j},\ \forall i_1,\ldots,i_r,j \in J, \\ \displaystyle\sum_{j\in J} E_{\mu_{i_1,\ldots,i_r,j}}[X_{i_1,\ldots,i_r,j}] = 1,\ \forall i_1,\ldots,i_r \in J, \\ \displaystyle\sum_{i_1,\ldots,i_r\in J} \pi^{(r)}_{i_1,\ldots,i_r} E_{\mu_{i_1,\ldots,i_r,j}}[X_{i_1,\ldots,i_r,j}] = \pi_j,\ \forall j \in J. \end{array} \right.$$

The maxentropic reconstruction of r-step transition probability array $\mathrm{P}^{(r)} = \left(P^{(r)}_{i_1,\ldots,i_r;j}\right)_{i_1,\ldots,i_r,j\in J}$ is based on the following direct consequence of Corollary 2.1.

Corollary 3.1. *If μ is a feasible solution of $(P3)$, then the matrix $\mathrm{P}^{(r)} = \left(P^{(r)}_{i_1,\ldots,i_r;j}\right)_{i_1,\ldots,i_r,j\in J}$ given by*

$$P^{(r)}_{i_1,\ldots,i_r;j} = E_{\mu_{i_1,\ldots,i_r,j}}[X_{i_1,\ldots,i_r,j}],\ \forall i_1,\ldots,i_r,j \in J$$

is a solution for system (6).

Obviously, we have

$$\langle \lambda, \mathrm{b}\rangle = \sum_{k_1,\ldots,k_r\in J} \lambda_{k_1,\ldots,k_r} + \sum_{k\in J} \pi_k \lambda_{-k},$$

$$(\mathrm{A}^*\lambda)_{i_1,\ldots,i_r,j} = \lambda_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda_{-j},\ \forall i_1,\ldots,i_r,j \in J,$$

and hence we obtain that the geometric dual problem of $(P3)$ has the following form:

$$\text{Program } (D3): \left| \begin{array}{l} \min L(\lambda) = \displaystyle\sum_{i_1,\ldots,i_r\in J} \Big[\sum_{j\in J} \ln \zeta_{i_1,\ldots,i_r,j}(\lambda_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda_{-j}) \\ \qquad\qquad + \lambda_{i_1,\ldots,i_r}\Big] + \displaystyle\sum_{i\in J} \pi_i \lambda_{-i} \quad \text{s.t.} \\ \lambda \in \Lambda_1(\nu, \pi^{(r)}), \end{array} \right.$$

where

$$\Lambda_1(\nu,\pi^{(r)}) = \begin{cases} \{\lambda \in \mathbb{R}^K \ /\ L(\lambda) < \infty\}, & \text{if } J \text{ is finite,} \\ \{\lambda \in \mathbb{R}^K_+ \ /\ L(\lambda) < \infty,\ \sum_{k\in K} \lambda_k < \infty\}, & \text{if } J \text{ is infinite.} \end{cases}$$

ccording to Corollary 2.2 we have the next duality result.

Corollary 3.2. *Assume that* $\sum_{(i_1,\ldots,i_r,j)\in I} \int_{\Omega_{i_1,\ldots,i_r,j}} X_{i_1,\ldots,i_r,j} d\nu_{i_1,\ldots,i_r,j} < \infty.$

(i) **(weak duality)** *If μ is a primal feasible solution of program $(P3)$ and λ is a dual feasible solution of program $(D3)$, then*

$$-H(\mu;\nu) \leq L(\lambda).$$

Moreover, the equality holds if and only if

$$d\mu_{i_1,\ldots,i_r,j} = \frac{e^{-[\lambda_{i_1,\ldots,i_r}+\pi^{(r)}_{i_1,\ldots,i_r}\lambda_{-j}]X_{i_1,\ldots,i_r,j}}}{\zeta_{i_1,\ldots,i_r,j}(\lambda_{i_1,\ldots,i_r}+\pi^{(r)}_{i_1,\ldots,i_r}\lambda_{-j})}\, d\nu_{i_1,\ldots,i_r,j},\ \forall (i_1,\ldots,i_r,j)\in I.$$

(ii) **(strong duality)** *If $\lambda^* \in \operatorname{Int}\Lambda_1(\nu,\pi^{(r)})$ is a dual optimal solution of program $(D3)$, then $\mu(\lambda^*) = \bigotimes_{(i_1,\ldots,i_r,j)\in I} \mu_{i_1,\ldots,i_r,j}(\lambda^*)$ given by*

$$d\mu_{i_1,\ldots,i_r,j}(\lambda^*) = \frac{e^{-[\lambda^*_{i_1,\ldots,i_r}+\pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j}]X_{i_1,\ldots,i_r,j}}}{\zeta_{i_1,\ldots,i_r,j}(\lambda^*_{i_1,\ldots,i_r}+\pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j})}\, d\nu_{i_1,\ldots,i_r,j},\ \forall (i_1,\ldots,i_r,j)\in I$$

is a primal optimal solution of program $(P3)$ and

$$-H(\mu(\lambda^*);\nu) = L(\lambda^*).$$

Remark 3.2. *Taking $r=1$, we obtain the maxentropic reconstruction of simple Markov chains with given common steady-state probabilities, with supplementary constraints of type (5) (see [13]) or without supplementary constraints (see [6]).*

Example 3.1. *Let $a \in \mathbb{R}_+$. For any $(i_1,\ldots,i_r,j)\in I$, let*

$$\Omega_{i_1,\ldots,i_r,j} = \{0,a\},\ \nu_{i_1,\ldots,i_r,j} = (1-\theta_{i_1,\ldots,i_r,j})\varepsilon_0 + \theta_{i_1,\ldots,i_r,j}\varepsilon_a,$$

where $\theta_{i_1,\ldots,i_r,j}\in[0,1]$. Assume that

$$\sum_{(i_1,\ldots,i_r,j)\in I} \theta_{i_1,\ldots,i_r,j} < \infty.$$

We have

$$\zeta_{i_1,\ldots,i_r,j}(y) = 1-\theta_{i_1,\ldots,i_r,j}+\theta_{i_1,\ldots,i_r j}e^{-y\cdot a}.$$

The dual problem $(D3)$ has the following form:

$$\left|\begin{array}{l} \min L(\lambda) = \displaystyle\sum_{i_1,\ldots,i_r\in J}\left\{\sum_{j\in J}\ln\left[1-\theta_{i_1,\ldots,i_r,j}+\theta_{i_1,\ldots,i_r j}e^{-a(\lambda_{i_1,\ldots,i_r}+\pi^{(r)}_{i_1,\ldots,i_r}\lambda_{-j})}\right]\right. \\ \qquad\qquad\left. +\lambda_{i_1,\ldots,i_r}\right\} + \displaystyle\sum_{i\in J}\pi_i\lambda_{-i} \quad s.t. \\ \lambda\in\Lambda_1(\nu,\pi^{(r)}). \end{array}\right.$$

Let λ^* be an interior optimal solution of this problem. Then the primal problem $(P3)$ has the optimal solution $\mu(\lambda^*) = \bigotimes_{(i_1,\ldots,i_r,j)\in I} \mu_{i_1,\ldots,i_r,j}(\lambda^*)$ given by

$$\mu_{i_1,\ldots,i_r,j}(\lambda^*) = \frac{(1-\theta_{i_1,\ldots,i_r,j})\varepsilon_0 + \theta_{i_1,\ldots,i_r,j} e^{-a[\lambda^*_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j}]}\varepsilon_a}{1-\theta_{i_1,\ldots,i_r,j} + \theta_{i_1,\ldots,i_r,j} e^{-a[\lambda^*_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j}]}},$$

$\forall(i_1,\ldots,i_r,j) \in I$.

Hence the $r+1$-dimensional matrix $\mathrm{P}^{(r)} = \left(P^{(r)}_{i_1,\ldots,i_r;j}\right)_{i_1,\ldots,i_r,j\in J}$ given by

$$\begin{aligned} P^{(r)}_{i_1,\ldots,i_r;j} &= E_{\mu_{i_1,\ldots,i_r,j}(\lambda^*)}[X_{i_1,\ldots,i_r,j}] \\ &= \frac{a\theta_{i_1,\ldots,i_r,j} e^{-a[\lambda^*_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j}]}}{1-\theta_{i_1,\ldots,i_r,j} + \theta_{i_1,\ldots,i_r,j} e^{-a[\lambda^*_{i_1,\ldots,i_r} + \pi^{(r)}_{i_1,\ldots,i_r}\lambda^*_{-j}]}}\ ,\ \forall i_1,\ldots,i_r,j \in J, \end{aligned}$$

is a r-step transition probability array of a homogeneous and stationary Markov chains of order r that verifies the imposed constraints (6), where

$$\pi^{(\mathrm{r})} = \left(\pi^{(r)}_{i_1,\ldots,i_r}\right)_{i_1,\ldots,i_r\in J}$$

is the given joint probability of the states at r consecutive times.

Remark 3.3. *If the given joint probability $\pi^{(\mathrm{r})}$ is unknown, then it can be also reconstructed using the MEM method, from the imposed constraints (7), (8) and (9).*

References

[1] C. Bălcău, Some examples for maxentropic reconstruction of countable probability distributions, *Bul. Sti. Univ. Pitesti ser. Mat.-Inf.*, **11** (2005), 13-22.

[2] S.-C. Fang, J.R. Rajasekera, H.-S.J. Tsao, *Entropy Optimization and Mathematical Programming*, Kluwer Academic, Dordrecht, 1997.

[3] S. Guiaşu, *Quantum Mechanics*, Nova Science Publ., Huntington, New York, 2001.

[4] H. Gzyl, Maxentropic reconstruction of some probability distributions, *Studies in Appl. Math.*, **105** (2000), 235-243.

[5] H. Gzyl, Ill-posed linear inverse problems and maximum entropy in the mean, *Acta Cientif. Venez.*, **53** (2002), 74-93.

[6] H. Gzyl, Y. VelÁsquez, Reconstruction of transition probabilities by maximum entropy in the mean, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering, R. Fry (ed.), American Institute of Physics*, **617** (2002), 192-203.

[7] M. Iosifescu, *Finite Markov Processes and their Applications*, Wiley Series in Probability and Mathematical Statistics, Chichester, 1980.

[8] M. Iosifescu, S. Grigorescu, *Dependence with Complete Connections and its Applications*, Cambridge University Press, Cambridge, 1990.

[9] M. Iosifescu, R. Theodorescu, On the entropy of chains with complete connections, *Comun. Acad. R.P.R.*, **11** (1961), 821-824.

[10] E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.*, **106** (1957), 620-630.

[11] E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.*, **108** (1957), 171-190.

[12] S. Kullbeck, *Information Theory and Statistics*, Wiley, New York, 1959.

[13] V. Preda, C. Bălcău, On maxentropic reconstruction of countable Markov chains and matrix scaling problems, *Studies in Appl. Math.*, **111** (2003), 85-100.

[14] V. Preda, C. Bălcău, Homogeneous and stationary multiple Markov chains with maximum Iosifescu-Theodorescu entropy, *Rev. Roumaine de Math. Pures et Appl.*, to appear.

[15] E. Rietch, A maximum entropy approach to inverse problems, *J. Geophys.*, **42** (1977), 489-506.

[16] C.E. Shannon, A mathematical theory of comunication, Part III, *Bell System Technical Journal*, **27** (1948), 623-656.

Received: 27.08.2007.

University of Piteşti,
Faculty of Mathematics and Informatics
E-mail: cbalcau@linux.math.upit.ro

University of Bucureşti,
Faculty of Mathematics and Informatics;
E-mail: preda@fmi.unibuc.ro