JOURNAL OF STATISTICAL THEORY AND PRACTICE
2017, VOL. 11, NO. 3, 407–417
http://dx.doi.org/10.1080/15598608.2017.1282895

On a reward rate estimation for the finite irreducible continuous-time Markov chain

Alexander Andronov

Mathematical Methods and Modeling, Transport and Telecommunication Institute, Riga, Latvia

ABSTRACT

A continuous-time homogeneous irreducible Markov chain $\{X(t)\},\ t \in [0,\infty)$, taking values on $N = \{1,\ldots,k\}, k < \infty$, is considered. Matrix $\lambda = (\lambda_{ij})$ of the intensity of transition λ_{ij} from state i to state j is known. A unit of the sojourn time in state i gives reward β_i so the total reward during time t is $Y(t) = \int_0^t \beta_{X(s)} ds$. The reward rates $\{\beta_i\}$ are not known and it is necessary to estimate them. For that purpose the following statistical data on r observations are at our disposal: (1) t, observation time; (2) i, initial state $X(0)$; (3) j, final state $X(t)$; and (4) y, acquired reward $Y(t)$. Two methods are used for the estimation: the weighted least-squares method and the saddle-point method for the Laplace transformation of the reward. Simulation study illustrates the suggested approaches.

ARTICLE HISTORY
Received 29 April 2016
Accepted 12 January 2017

KEYWORDS
Methods of point estimation; Markov chain; simulation

AMS SUBJECT CLASSIFICATION
62M05; 62F10

1. Introduction

We consider a continuous-time homogeneous irreducible Markov chain $\{X(t)\}, t \in [0,\infty)$, taking values on $N = \{1,\ldots,k\}, k < \infty$; see, for example, Kijima (1997) and Pacheco, Tang, and Prabhu (2009). Let $\lambda_{i,j}$ be the intensity of a transition from state i to state j $(\lambda_{i,i} = 0), \lambda = (\lambda_{i,j})$ be the $k \times k$ matrix, $\Lambda_i = \sum_{j=1}^{k} \lambda_{i,j}$ be the intensity of transition from state i, and $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. The $k \times k$ matrix $A = (A_{i,j}) = (\lambda_{i,j} - \delta_{i,j}\Lambda_i))$ is named the generator. Further, a unit of the sojourn time in state i gives reward β_i so the total reward during time t is $Y(t) = \int_0^t \beta_{X(s)} ds$. It is supposed that $\beta_i \neq \beta_j$ for $i \neq j$.

We suppose that transition intensities $\{\lambda_{i,j}\}$ are known but reward rates $\{\beta_i\}$ are not known. We need to estimate them, using the following statistical data on r observations with numbers $\eta = 1,\ldots,r$: (1) t_η, observation time, (2), i_η, initial state $X(0)$,(3) j_η, final state $X(t_\eta)$, and (4) y_η, acquired reward $Y(t_\eta)$.

First we consider reward $Y(t)$ as a response of the linear regression:

$$Y(t) = \sum_{i=1}^{k} \beta_i T_i(t) = \sum_{i=1}^{k} \beta_i E(T_i(t)) + U(t,\beta) \tag{1}$$

CONTACT Alexander Andronov lora@mailbox.riga.lv Mathematical Methods and Modeling, Transport and Telecommunication Institute, LV-1019, Lomonosova 1, Riga, Latvia.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujsp.

where $T_i(t)$ is a sojourn time in the state i of a Markov chain during time t, $U(t, \beta)$ is a random variable with zero mean depending on t, and $\beta = (\beta_1, ..., \beta_k)$.

A more general setting of this problem has been considered by Andronov (1992; 2014). Elaborated methods are used for the specific case under study.

Usually the maximum likelihood method is used for the estimation. This method is based on the distribution function of the reward *Y(t)*. The calculation of this distribution function is a rather complicated problem, although the corresponding Laplace transformation is known and is not complicated; see, for example, Bladt, Meini, Neuts, and Sericola (2002) and Sericola (2000). Methods known for inverting transformation of probability distributions (Abate and Whitt 1992; Grawford, Minin, and Suchard 2014) are inadmissible in our case. Initially we used the following alternative approach: At the beginning the moments of reward $\{EY(t)^r\}$ were calculated, and then the required distribution function was approximated on the basis of the acquired moments. Unfortunately, this approach did not give good results and another approach was chosen.

Let $J_j(t)$ be an indicator function of the event $\{X(t) = j\}$, which equals 1 if the event takes place and equals 0 otherwise. We denote the Laplace transformation of the reward *Y(t)* and $J_j(t)$ by $\Pi_{i,j}(s, t, \beta) = E(exp(-sY(t)) \times J_j(t)|X(0) = i), i, j = 1, ..., k$. It is tempting to use this transformation immediately for the estimation of parameters β. For that the method of the saddle point is used. This is a point $(s*, \beta^*)$ giving a maximum of criterion under interest with respect to s and minimum with respect to β.

The article is organized as follows. The considered continuous-time Markov chain is described in the second section. The method of weighted least squares is presented in the third section. Sections 4–6 are devoted to the method of the saddle point and its modifications. Results of the simulation study are presented in section 7. Conclusion remarks end the article.

2. Preliminaries

Let $P_{i,j}(t) = P\{X(t) = j|X(0) = i\}$ be the transition probability of the Markov chain *X(t)*, and $P(t) = (P_{i,j}(t))_{k\times k}$ be the corresponding matrix. We suppose that all eigenvalues of the generator matrix A are different. In this case probabilities $P(t) = (P_{i,j}(t))_{k\times k}$ can be represented in a simple way. Let ξ_i and Z_i, $i = 1, ..., k$, be the eigenvalue and the corresponding eigenvector of A, $\xi = (\xi_1, ..., \xi_k)^T$, $Z = (Z_1, ..., Z_k)$ be the matrix of the eigenvectors, and $\bar{Z} = Z^{-1} = (\bar{Z}_1^T, ..., \bar{Z}_k^T)^T$ be the corresponding inverse matrix (here $\bar{Z}_\eta$ is the η-th row of $\bar{Z}$). Then (see Bellman 1969)

$$P(t) = exp(tA) = Zdiag(exp(\xi t))Z^{-1} = \sum_{i=1}^{k} Z_i exp(t\xi_i)\bar{Z}_i. \quad (2)$$

Let us fix initial state i and final state j of Markov chain $X(t)$ and consider the sojourn time $T_\nu(t)$ in the state $\nu \in N$ on the interval $(0, t)$. Then for expectation $\tau_\nu(t, i, j) = E(T_\nu(t)J_j(t)|X(0) = i, \ X(t) = j)$, we have:

$$\tau_\nu(t, i, j) = \int_0^t P_{i,\nu}(u)P_{\nu,j}(t-u)du, \quad \nu = 1, ..., k. \quad (3)$$

The conditional mixed second moments $\tau_{v,v*}(t,i,j) = E(T_v(t)T_{v*}(t)|X(0) = i, X(t) = j)$ and conditional covariance $C_{v,v*}(t,i,j) = Cov(T_v(t), T_{v*}(t)|X(0) = i,\ X(t) = j)$ of the sojourn time in the states $v, v* \in N$ on the interval $(0,t)$ are calculated as

$$\tau_{v,v*}(t,i,j) = \frac{1}{P_{i,j}(t)}(\int_0^t P_{i,v}(u)\tau_{v*}(t-u,v,j)du + \int_0^t P_{i,v*}(u)\tau_v(t-u,v*,j)du). \tag{4}$$

$$C_{v,v^*}(t,i,j) = \tau_{v,v*}(t,i,j) - P_{i,j}(t)^{-2}\tau_v(t,i,j)\tau_{v^*}(t,i,j). \tag{5}$$

Let us consider the vectors $T(t) = (T_1(t),\dots,T_k(t))^T$, $\tau(t,i,j) = (\tau_1(t,i,j),\dots,\tau_k(t,i,j))^T$, $\beta = (\beta_1,\dots,\beta_k)$, and conditional covariance matrix $C(t,i,j) = (C_{v,v^*}(t,i,j))$ of $T(t)$. The expectation $EY(t)$ and the variance $DY(t)$ of the reward are calculated as

$$EY(t) = \beta^T ET(t),\ DY(t) = \beta^T Cov(T(t))\beta. \tag{6}$$

We wish to estimate parameter $\beta = (\beta_1,\dots,\beta_k)^T$ using the following statistical data on r observations: (1) t_η, the duration of the η-th observation; (2) i_η, the initial state of the η-th observation $X(0)$; (3) j_η, the final state of the η-th observation $X(t_\eta)$; and (4) y_η, the acquired reward $Y(t_\eta)$.

Here we should mention the following reasons, taking into account that $\beta_i \neq \beta_j$ for $i \neq j$. For two observations with numbers η and $\eta*$ let $i_\eta = i_\eta* = j_\eta = j_\eta* = i0$, $Y(t_\eta)/t_\eta = Y(t_\eta*)/t_\eta*$. This means that the initial state $i0$ of the Markov chain retains itself during both observations with probability 1. Therefore, $\beta_{i0} = Y(t_\eta)/t_\eta = Y(t_\eta*)/t_\eta*$. This simplifies the problem of the estimation essentially. Further, we suppose that durations of the observations are sufficiently big that this situation does not take place.

It is supposed that matrix λ is known and has relative small dimension $k \leq 10$; therefore, vector $ET(t)$ and covariance matrix $Cov(T(t))$ are known too for each observation η (needed expressions were represented by Eqs. (3) and (5)).

We begin with the method of weighted least squares.

3. Method of weighted least squares

The method of weighted least squares (Turkington 2002) presupposes minimizing the sum

$$LSS(\beta) = (\sum_{\eta=1}^r \omega_\eta)^{-1}\sum_{\eta=1}^r \omega_\eta(y_\eta - EY(t_\eta))^2 = (\sum_{\eta=1}^r \omega_\eta)^{-1}\sum_{\eta=1}^r \omega_\eta(y_\eta - \beta^T ET(t_\eta))^2$$

$$= (\sum_{\eta=1}^r \omega_\eta)^{-1}(y - H\beta)^T diag(\omega_1,\dots,\omega_r)(y - H\beta), \tag{7}$$

$$\text{where } \omega_\eta = DY(t_\eta)^{-1} = (\beta^T Cov(T(t_\eta))\beta)^{-1},$$

$$y = (y_1,\dots,y_r)^T,\quad H = (ET(t_1),\dots,ET(t_r))^T.$$

If weights $\{\omega_\eta\}$ do not depend on unknown parameters β, then the classical estimator is calculated as

$$b = (H^T diag(\omega_1,\dots,\omega_r)H)^{-1}H^T diag(\omega_1,\dots,\omega_r)y. \tag{8}$$

As weights $\{\omega_\eta\}$ depend on unknown parameters β, the usual two-step procedure is repeated. First, weights are calculated for the given estimator of β; then a new estimator is recalculated. Usually a small number of iterations is sufficient for convergence.

4. Method of the saddle point

This method uses the explicit form of Laplace transformation $\Pi_{i,j}(s,\beta,t) = E(exp(-sY(t)) \times J_j(t)|X(0) = i)$, $i,j = 1,\ldots,k$, of the reward $Y(t)$ and the indicator function $J_j(t)$ of the event $\{X(t) = j\}$. Let $\Pi(s,\beta,t) = (\Pi_{i,j}(s,\beta,t))_{k\times k}$ be corresponding $k \times k$ matrix of the Laplace transformations.

Further, let $B = diag(\beta_1,\ldots,\beta_k)$. It is known (Bladt, Meini, Neuts, and Sericola 2002; Sericola 2000) that

$$\Pi(s,\beta,t) = exp(t(A - sB)) = \sum_{\psi=0}^{\infty} \frac{1}{\psi!}(t(A - sB))^{\psi}. \tag{9}$$

Let us consider a mini–max criterion

$$\hat{R}(s,b) = \sum_{\eta=1}^{r} \left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,b,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta)\right)^2, \tag{10}$$

where notations $t(\eta)$, $i(\eta)$, and $j(\eta)$ are used instead of t_η, i_η, and j_η to avoid double indexes.

We must find the saddle point $(s*, b^*)$ of this function: It is maximum with respect to s and minimum with respect to b (Minoux, 1989). The maximum with respect to s is searched in the positive neighborhood of zero. Let us find the derivative of Eq. (9) with respect to s:

$$\begin{aligned}\frac{\partial}{\partial s}\Pi(s,\beta,t) &= \frac{\partial}{\partial s} exp(t(A - sB)) = \frac{\partial}{\partial s}\sum_{\psi=0}^{\infty}\frac{1}{\psi!}(t(A - sB))^{\psi} \\ &= \sum_{\psi=1}^{\infty}\frac{1}{\psi!}t^{\psi}\sum_{m=0}^{\psi-1}(A - sB)^m(-B)(A - sB)^{\psi-1-m} \\ &= -tB - \sum_{\psi=2}^{\infty}\frac{1}{\psi!}t^{\psi}\sum_{m=0}^{\psi-1}(A - sB)^m B(A - sB)^{\psi-1-m}.\end{aligned} \tag{11}$$

It is necessary to find the gradient of Eq. (9) with respect to β. Let I_γ be the $k \times k$ matrix having unique nonzero element 1 on the γ-th place of the main diagonal. Let $M^{(\gamma)}$ and $M_{(\gamma)}$ be the γ-th column and the γ-th row of matrix M correspondingly. Then the partial derivative with respect to β_γ is the following:

$$\begin{aligned}\frac{\partial}{\partial\beta_\gamma}\Pi(s,\beta,t) &= \frac{\partial}{\partial\beta_\gamma}exp(t(A-sB)) = \frac{\partial}{\partial\beta_\gamma}\sum_{\psi=0}^{\infty}\frac{1}{\psi!}(t(A-sB))^{\psi}\\ &= \sum_{\psi=1}^{\infty}\frac{1}{\psi!}t^{\psi}\sum_{m=0}^{\psi-1}(A-sB)^m(-sI_\gamma)(A-sB)^{\psi-1-m}\\ &= -s\sum_{\psi=1}^{\infty}\frac{1}{\psi!}t^{\psi}\sum_{m=0}^{\psi-1}(A-sB)^m I_\gamma I_\gamma(A-sB)^{\psi-1-m}\\ &= -s\sum_{\psi=1}^{\infty}\frac{1}{\psi!}t^{\psi}\sum_{m=0}^{\psi-1}((A-sB)^m)^{(\gamma)}((A-sB)^{\psi-1-m})_{(\gamma)}.\end{aligned} \tag{12}$$

Therefore,

$$\begin{aligned}\tfrac{\partial}{\partial s}\hat{R}(\beta,s) = 2\sum_{\eta=1}^{r}&(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta))\times\\ &(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\tfrac{\partial}{\partial s}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} + y_\eta exp(-sy_\eta)).\end{aligned} \tag{13}$$

$$\begin{aligned}\tfrac{\partial}{\partial\beta_\gamma}\hat{R}(\beta,s) &= \tfrac{\partial}{\partial\beta_\gamma}\sum_{\eta=1}^{r}(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta))^2\\ &= 2\sum_{\eta=1}^{r}(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta))\times\\ &\qquad(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\tfrac{\partial}{\partial\beta_\gamma}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)}).\end{aligned} \tag{14}$$

Also, the gradient of $\hat{R}(\beta,s)$ with respect to vector $\beta=(\beta_1,\ldots,\beta_k)$ can be calculated.

Now we are able to use gradient method to find a saddle point $(s*,\beta^*)$. It appears that the maximum with respect to s depends on β negligibly. Now minimization with respect to $\beta=(\beta_1,\ldots,\beta_k)$ for fixed $s*$ does not pose any problem. Note that calculation of the infinity sum is restricted by the finite number of the addends.

5. Weighted method of the saddle point

Usually weights are used with the purpose of improving the acquired estimators. The optimal weight of the given observation is equal to an inverse value of the variance of the observation. Thus, it is necessary to calculate the variance $exp(-sY_\eta)$ at the point s for the η-th observation. As

$$E_\beta\left(exp(-sY_\eta)\right) = \frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)},$$

then

$$\begin{aligned}Var_\beta(exp(-sY_\eta)) = E_\beta(exp(-sY_\eta))^2 - (E_\beta(exp(-sY_\eta)))^2 =\\ \tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(2s,\beta,t(\eta))_{i(\eta),j(\eta)} - \left(\tfrac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)}\right)^2.\end{aligned}$$

Therefore, the absolute weight $w_\eta(\beta)$ of the η-th observation is

$$
\begin{aligned}
w_\eta(\beta) &= Var_\beta(exp(-sY_\eta))^{-1} \\
&= \left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(2s,\beta,t(\eta))_{i(\eta),j(\eta)} - \left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)}\right)^2\right)^{-1}.
\end{aligned} \tag{15}
$$

The relative weight of the η-th observation equals $w_\eta(\beta)/w_\Sigma(\beta)$ where $w_\Sigma(\beta)$ is the sum of all absolute weights:

$$
w_\Sigma(\beta) = \sum_{\eta=1}^{r} w_\eta(\beta). \tag{16}
$$

Now we can represent the weighted mini–max criterion as

$$
R_w(\beta,s) = \sum_{\eta=1}^{r} \frac{w_\eta(\beta)}{w_\Sigma(\beta)}\left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta)\right)^2. \tag{17}
$$

The derivative with respect to s is

$$
\begin{aligned}
\frac{\partial}{\partial s}R_w(\beta,s) &= 2\sum_{\eta=1}^{n} \frac{w_\eta(\beta)}{w_\Sigma(\beta)}\left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta)\right)\times \\
&\times \left(\frac{1}{P_{i(\eta),j(\eta)}t(\eta)}\frac{\partial}{\partial s}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} + y_\eta exp(-sy_\eta)\right).
\end{aligned} \tag{18}
$$

The gradient with respect to β is calculated according to Eq. (12) using the two-step repeated operations as follows. First, weights $\{w_\eta(\beta)\}$ are calculated for given β, and then the gradient with respect to β is calculated for fixed $\{w_\eta(\beta)\}$:

$$
\begin{aligned}
\frac{\partial}{\partial \beta_\gamma}R_w(\beta,s) &= \frac{\partial}{\partial_{\beta_\gamma}}\sum_{\eta=1}^{r} \frac{w_\eta(\beta)}{w_\Sigma(\beta)}\left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta)\right)^2 = \\
&- 2s\sum_{\eta=1}^{r} \frac{w_\eta(\beta)}{w_\Sigma(\beta)}\left(\frac{1}{P_{i(\eta),j(\eta)}(t(\eta))}\Pi(s,\beta,t(\eta))_{i(\eta),j(\eta)} - exp(-sy_\eta)\right)\frac{1}{P_{i(\eta),j(\eta)}t(\eta)}\times \\
&\times \sum_{\psi=1}^{\infty}\frac{1}{\psi!}t^\psi \sum_{m=0}^{\psi-1}\left(((A-sB)^m)^{(\gamma)}((A-sB)^{\psi-1-m})\right)_{(\gamma)})_{i(\eta),j(\eta)}.
\end{aligned} \tag{19}
$$

These operations are repeated until a convergence; usually two to three iterations are sufficient.

The considered criterion (17) is such that each observation $\eta = 1, \ldots, r$ gives its square of the deviation. This makes the criterion a very unsmooth function. To increase the smoothing, it is reasonable to unite observations with the same indices $i(\eta)$, $j(\eta)$, $t(\eta)$. This approach will be considered further.

6. Modified weighted mini–max criterion

Let $\Psi(t,i,j)$ be the set of the observations with the same values t,i,j; $n(t,i,j) = |\Psi(t,i,j)|$ be a number of such observations. We unite rewards $\{y_\eta\}$ as follows:

$$y(t,i,j) = \sum_{\eta \in \Psi(t,i,j)} y_\eta.$$

Now $exp(-s\,y(t,i,j))$ is the Laplace transform of reward $y(t,i,j)$. The absolute weight $w_{(t,i,j)}(\beta)$ of the observation $exp(-s\,y(t,i,j))$ is the following:

$$w_{t,i,j}(\beta) = (n(t,i,j)(\frac{1}{P_{i,j}(t)}\Pi(2s,\beta,t)_{i,j} - (\frac{1}{P_{i,j}(t)}\Pi(s,\beta,t)_{i,j})^2))^{-1}. \tag{20}$$

The corresponding theoretical Laplace transformation is $(\Pi(s,\beta,t)_{i,j})^{n(t,i,j)}$. Therefore, criterion (17) has the following form:

$$\begin{aligned} R_M(s,\beta) = \sum_t \sum_i \sum_j \tfrac{w_{t,i,j}(\beta)}{w_\Sigma(\beta)} \times \\ \times((\tfrac{1}{P_{i,j}(t)}\Pi(s,\beta,t)_{i,j})^{n(t,i,j)} - exp(-sy(t,i,j)))^2. \end{aligned} \tag{21}$$

where w_Σ is the sum of all weights.

This criterion has a good theoretical foundation. Random variable $y(t,i,j)$ is a sum of many independent identically distributed random addends. Therefore, its distribution is close to a normal distribution. We need to estimate parameters $\beta = (\beta_1, \ldots, \beta_k)$ on the basis of samples $\{y(t,i,j)\}$ for various t,i,j. The maximum likelihood method leads to criterion (21) .

Now formulas (18) and (19) are of the form:

$$\begin{aligned} \tfrac{\partial}{\partial s} R_M(s,\beta) = 2\sum_t \sum_i \sum_j \tfrac{w_{t,i,j}(\beta)}{w_\Sigma(\beta)}((\tfrac{1}{P_{i,j}(t)}\Pi(s,\beta,t)_{i,j})^{n(t,i,j)} - exp(-sy(t,i,j)))\times \\ \times((\tfrac{1}{P_{i,j}(t)})^{n(t,i,j)} n(t,i,j)\Pi(s,\beta,t)_{i,j}^{n(t,i,j)-1} \tfrac{\partial}{\partial s}\Pi(s,\beta,t)_{i,j} + y(t,i,j)exp(-sy(t,i,j))). \end{aligned} \tag{22}$$

$$\begin{aligned} \tfrac{\partial}{\partial \beta_y} R_M(\beta,s) = 2\sum_t \sum_i \sum_j \tfrac{w_{t,i,j}(\beta)}{w_\Sigma(\beta)}((\tfrac{1}{P_{i,j}(t)}\Pi(s,\beta,t)_{i,j})^{n(t,i,j)} - exp(-sy(t,i,j)))\times \\ \times(\tfrac{1}{P_{i,j}(t)})^{n(t,i,j)} n(t,i,j)\Pi(s,\beta,t)_{i,j}^{n(t,i,j)-1} \tfrac{\partial}{\partial \beta_\gamma}\Pi(s,\beta,t)_{i,j}. \end{aligned} \tag{23}$$

7. Simulation study

We check the efficiency of suggested methods of the estimation, using the following initial data. The Markov chain has three states and the following transition matrix:

$$\lambda = \begin{pmatrix} 0 & 0.2 & 0.3 \\ 0.4 & 0 & 0.2 \\ 0.2 & 0.2 & 0 \end{pmatrix}.$$

The eigenvalues and eigenvectors of the generator for A are the following:

$$\chi = \begin{pmatrix} 0 \\ -0.8 \\ -0.7 \end{pmatrix}, \quad Z = \begin{pmatrix} -0.577 & -0.302 & -0.100 \\ -0.577 & -0.905 & -0.796 \\ -0.577 & -0.302 & 0.597 \end{pmatrix}.$$

Further three durations of the observations will be considered: $t = 8, 10,$ and 12. The matrix of transition probabilities between states for various durations are the following:

$$P(8) = \begin{pmatrix} 0.359 & 0.250 & 0.391 \\ 0.359 & 0.251 & 0.390 \\ 0.355 & 0.250 & 0.395 \end{pmatrix}, \; P(10) = \begin{pmatrix} 0.357 & 0.250 & 0.393 \\ 0.358 & 0.250 & 0.392 \\ 0.357 & 0.250 & 0.393 \end{pmatrix},$$
$$P(12) = \begin{pmatrix} 0.357 & 0.250 & 0.393 \\ 0.357 & 0.250 & 0.393 \\ 0.357 & 0.250 & 0.393 \end{pmatrix}.$$

The vector of the expectation (3) $\tau_v(t,i,j) = E(T_v(t)|X(0) = i, X(t) = j)$ of sojourn time for various initial i and final j states of the Markov chain and observation time $t = 8$ are the following:

$$\tau(8,1,1) = (4.514 \;\; 1.508 \;\; 1.978)^T, \quad \tau(8,1,2) = (3.235 \;\; 2.618 \;\; 2.147)^T,$$
$$\tau(8,1,3) = (3.212 \;\; 1.260 \;\; 3.528)^T, \quad \tau(8,2,1) = (3.428 \;\; 2.758 \;\; 1.814)^T,$$
$$\tau(8,2,2) = (2.152 \;\; 3.883 \;\; 1.3965)^T, \quad \tau(8,2,3) = (2.156 \;\; 2.490 \;\; 3.354)^T,$$
$$\tau(8,3,1) = (3.087 \;\; 1.524 \;\; 3.389)^T, \quad \tau(8,3,2) = (1.829 \;\; 2.618 \;\; 3.553)^T,$$
$$\tau(8,3,3) = (1.793 \;\; 1.248 \;\; 4.959)^T.$$

The covariance matrices (5) of sojourn time $C(t,i,j) = (Cov(T_v(t), T_{v*}(t)|X(0) = i, X(t) = j))_{3\times3}$ are presented for time $t = 8$, $X(0) = i = 1$, and $X(8) = j = 1, 2, 3$:

$$C(8,1,1) = \begin{pmatrix} 3.776 & -1.446 & -2.330 \\ -1.446 & 2.347 & -0.901 \\ -2.330 & -0.901 & 3.231 \end{pmatrix}, \; C(8,1,2) = \begin{pmatrix} 3.360 & -1.470 & -1.890 \\ -1.470 & 2.955 & -1.485 \\ -1.890 & -1.485 & 3.375 \end{pmatrix},$$
$$C(8,1,3) = \begin{pmatrix} 3.346 & -0.807 & -2.539 \\ -1.470 & 2.050 & -1.243 \\ -2.539 & -1.243 & 3.782 \end{pmatrix}.$$

We apply the simulation to get the observed data, using the following reward rates: $\beta = (1 \; 2 \; 3)^T$. The acquired data are presented in the Tables 1, 2, and 3. Each table contains 30 reward values y (odd rows) and 30 final states j (even rows) for initial states 1, 2, and 3 correspondingly and for the observation time $t = 8, 10, 12$.

Using these data, the initial value of estimate $\beta^* = (0.759 \; 1.987 \; 3.334)^T$ was acquired by the ordinary method of least squares. Using this value the method of weighted least squares gives estimate $\beta* = (0.822 \; 1.882 \; 3.346)^T$. To investigate the rate of the convergence, the number of the observations was increased by 90.

The results are presented in Table 4. The last row of this table contains values of the precision criterion $\delta = (\beta - \beta^*)^T diag(\beta)^{-1}(\beta - \beta^*)$, where $diag(\beta)$ is a diagonal matrix with vector β on the main diagonal.

Now we consider the method of the saddle point. First, we note that the value of s, maximizing criterion (17), depends on β negligibly, when β varies in the range presented in Table 4. This value for fixed β is the root of function (18) and is calculated easily. Figure 1 illustrates the small dependence of criterion (17) on β, where curves $R1(s)$ and

Table 1. The observed data (y, j) for initial state $X(0) = 1$.

$t = 8$	16.50	12.63	10.76	15.24	19.49.	16.67	17.09	15.84	11.05	13.26
$t = 8$	1	1	1	1	3	3	3	1	1	2
$t = 10$	11.60	17.71	12.24	16.45	14.19	13.38	13.15	11.69	20.34	14.63
$t = 10$	0	0	0	1	0	1	2	0	2	1
$t = 12$	14.82	19.13	15.87	18.59	20.04	16.25	21.98	17.10	19.08	22.25
$t = 12$	1	1	3	2	1	1	1	2	3	3

Table 2. The observed data (y, j) for initial state $X(0) = 2$.

$t = 8$	23.92	18.48	21.29	12.46	14.63.	13.50	18.38	19.54	19.14	13.55
$t = 8$	3	3	3	3	1	2	1	3	1	1
$t = 10$	21.18	24.53	23.04	20.22	22.90	15.50	18.54	13.35	17.19	11.00
$t = 10$	3	1	2	1	2	2	2	1	1	1
$t = 12$	19.29	21.65	26.37	26.19	20.50	18.64	17.97	24.42	20.91	24.99
$t = 12$	2	2	3	3	3	1	1	1	3	1

Table 3. The observed data (y, j) for initial state $X(0) = 3$.

$t = 8$	32.28	24.70	27.70	17.80	34.37	20.19	16.49	26.29	29.59	12.00
$t = 8$	2	3	2	2	3	1	1	1	3	1
$t = 10$	29.83	27.53	25.33	22.21	31.11	17.70	22.58	30.13	24.93	0.91
$t = 10$	3	1	1	1	2	3	1	3	3	3
$t = 12$	29.75	22.80	30.85	27.17	33.77	28.00	23.18	23.66	27.12	24.71
$t = 12$	2	1	3	1	2	2	1	1	1	2

Table 4. Convergence for the method of weighted least squares.

Sample size	90	180	270	360	450	540
β_1^*	0.822	0.917	1.068	0.953	0.961	0.968
β_2^*	1.882	1.821	1.743	1.878	1.856	1.840
β_3^*	3.346	3.376	3.231	3.214	3.153	3.111
δ	0.079	0.070	0.055	0.025	0.020	0.018

$R2(s)$ correspond to $\beta = (0.822\ 1.882\ 3.346)$ and $\beta = (0.968\ 1.840\ 3.111)$. The extremal value of s equals 0.052 in our calculations.

The convergence of the method of the saddle point is illustrated in Table 5, corresponding to the same data as for Table 4 and initial vector $\beta = (0.822\ 1.882\ 3.346)$.

Two last tables show that the method of the saddle point does not have considerable advantages compared with the method of weighted least squares. Nevertheless, it can be preferable when there is no valuable information such as the expectation and the variance of the sojourn time in the various states.

Additionally, it should be remarked that an estimate of the weighted least squares can be improved by use of the saddle-point procedure. For example, using the last to the final estimate of weighted least squares estimate $\beta = (0.968\ 1.840\ 3.111)^T$, we get estimate $\beta = (0.997\ 1.862\ 3.141)^T$ with $\delta = 0.016$. This and other examples show that better results are reached if the acquired estimates of the least squares method are made more precise by the method of the saddle point.

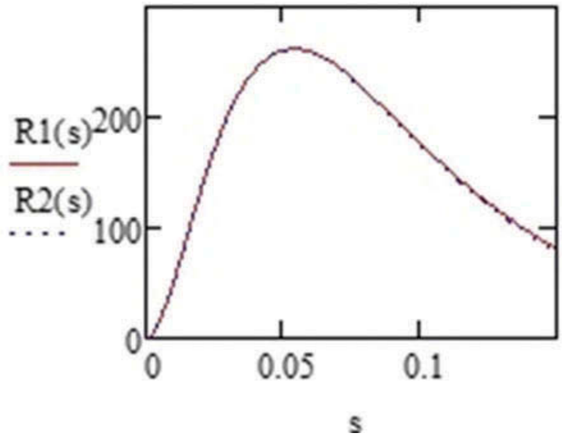

Figure 1. Values of criterion (17) as a function of *s*.

Table 5. Convergence for the method of the saddle point.

Sample size	90	180	270	360	450	540
β_1^*	0.845	0.884	0.980	0.966	0.951	0.952
β_2^*	1.896	1.902	1.799	1.845	1.841	1.869
β_3^*	3.375	3.416	3.319	3.268	3.209	3.216
δ	0.076	0.076	0.055	0.037	0.030	0.026

8. Conclusion

Two estimators of the reward rates for continuous-time homogeneous irreducible Markov chains have been considered: the weighted least squares method and the saddle-point method. It is ascertained that the best estimates are observed if estimates of the least squares are improved by the method of the saddle point. Our future research will be connected with an investigation of asymptotic properties of the suggested approach.

References

Abate, J., and W. Whitt. 1992. The Fourier-series method for inverting transform of probability distributions. *Queueing Systems* 19:5–88.

Andronov, A. M. 1992. Parameter statistical estimates of Markov-modulated linear regression. In *Statistical methods of parameter estimation and hypothesis testing*, vol. 24, 163–80. Perm, Russia: Perm State University (in Russian).

Andronov, A. M. 2014. Markov-modulated samples and their applications. In *Topics in statistical simulation*, ed. V. B. Melas, S. Mignani, P. Monari, and L. Salmoso, vol. 114, 29–35. New York, NY: Springer Proceedings in Mathematics & Statistics, Springer.

Bellman, R. 1969. *Introduction to matrix analysis*. New York, NY: McGraw-Hill.

Bladt, M., B. Meini, M. F. Neuts, and B. Sericola. 2002. Distributions of reward functions on continuous-time Markov chain. In *4th International Conference on Matrix Analytic Methods. Theory and applications*, 1–24. Adelaide, Australia.

Crawford, F. W., V. N. Minin, and M. A. Suchard. 2014. Estimation for general birth–death processes. *Journal of the American Statistical Association* 109 (506): 730–47.

Kijima, M. 1997. *Markov processes for stochastic modeling*. London, UK: Chapman & Hall.

Minoux, M. 1989. *Programmation matematique. Theorie et Algorithmes*. Paris, France: Bordas.
Pacheco, A., L. C. Tang, and N. U. Prabhu. 2009. *Markov-modulated processes* & *semiregenerative phenomena*. Hoboken, NJ: World Scientific.
Sericola, B. 2000. Occupation times in Markov processes. *Stochastic Models* 16:479–510.
Turkington, D. A. 2002. *Matrix calculus and zero-one matrices. Statistical and econometric applications*. Cambridge, UK: Cambridge University Press.