Finding a continuous-time Markov chain via sparse stochastic matrices in manpower systems

A.A. Osagiede*, V.U. Ekhosuehi

Department of Mathematics, University of Benin, Benin City, Nigeria

Received 16 July 2013; received in revised form 11 May 2014; accepted 17 July 2014
Available online 4 December 2014

Abstract

We consider a manpower system with finite discrete non-overlapping states where recruitment is done to replace wastage and to achieve the desired growth. The states of the system are defined in terms of the ranks. Data for the evolution of manpower structure in the system may be obtained at any choice of time instants. The empirical stochastic matrix resulting from the evolution of the system at each time instant is sparse. We propose a transition model for the system where the multi-step empirical stochastic matrix is expressed as the exponential of a Markov generator. We give illustrations using academic staff data in a university setting.

© 2014 Nigerian Mathematical Society. Production and Hosting by Elsevier B.V. All rights reserved.

MSC: 60J20; 91D35

Keywords: Manpower systems; Markov generator; Sparse matrix; Stochastic matrix

1. Introduction

Consider a manpower system stratified into finite discrete non-overlapping states defined in terms of the ranks. Transitions within the system follow a natural order (i.e., from one state to the next higher state). Let \( S = \{1, 2, \ldots, k\} \) be the set of these states. The state of the system at any choice of time instant \( t_v \), \( v = 0, 1, 2, \ldots, \gamma \), is represented by the row vector

\[
\bar{q}(t_v) = [\bar{q}_1(t_v), \bar{q}_2(t_v), \ldots, \bar{q}_k(t_v)],
\]

where \( \bar{q}_i(t_v) \) is the expected relative proportion of members of the system in state \( i \) at time \( t_v \). The vector \( \bar{q}(t_v) \) is called the expected relative structure (or simply, the structure) of the system at time \( t_v \). Moreover, when the expected number of members of the system at time \( t_v \), denoted by \( \bar{n}(t_v) \), is known, the expected relative structure is obtained as

\[
\bar{q}(t_v) = \bar{n}(t_v)[\bar{n}(t_v)e']^{-1},
\]

1. Introduction

Consider a manpower system stratified into finite discrete non-overlapping states defined in terms of the ranks. Transitions within the system follow a natural order (i.e., from one state to the next higher state). Let \( S = \{1, 2, \ldots, k\} \) be the set of these states. The state of the system at any choice of time instant \( t_v \), \( v = 0, 1, 2, \ldots, \gamma \), is represented by the row vector

\[
\bar{q}(t_v) = [\bar{q}_1(t_v), \bar{q}_2(t_v), \ldots, \bar{q}_k(t_v)],
\]

where \( \bar{q}_i(t_v) \) is the expected relative proportion of members of the system in state \( i \) at time \( t_v \). The vector \( \bar{q}(t_v) \) is called the expected relative structure (or simply, the structure) of the system at time \( t_v \). Moreover, when the expected number of members of the system at time \( t_v \), denoted by \( \bar{n}(t_v) \), is known, the expected relative structure is obtained as

\[
\bar{q}(t_v) = \bar{n}(t_v)[\bar{n}(t_v)e']^{-1},
\]

1. Introduction

Consider a manpower system stratified into finite discrete non-overlapping states defined in terms of the ranks. Transitions within the system follow a natural order (i.e., from one state to the next higher state). Let \( S = \{1, 2, \ldots, k\} \) be the set of these states. The state of the system at any choice of time instant \( t_v \), \( v = 0, 1, 2, \ldots, \gamma \), is represented by the row vector

\[
\bar{q}(t_v) = [\bar{q}_1(t_v), \bar{q}_2(t_v), \ldots, \bar{q}_k(t_v)],
\]

where \( \bar{q}_i(t_v) \) is the expected relative proportion of members of the system in state \( i \) at time \( t_v \). The vector \( \bar{q}(t_v) \) is called the expected relative structure (or simply, the structure) of the system at time \( t_v \). Moreover, when the expected number of members of the system at time \( t_v \), denoted by \( \bar{n}(t_v) \), is known, the expected relative structure is obtained as

\[
\bar{q}(t_v) = \bar{n}(t_v)[\bar{n}(t_v)e']^{-1},
\]
where \( \mathbf{e}' \) is a \( k \times 1 \) vector of ones. The time instant \( t_v \) may not be an integer for \( v = 0, 1, 2, \ldots, \gamma \), where \( \gamma \) is the maximum time index for which data are available. Let \( \{\mathbf{P}(t_v)\}_{v=0}^{\gamma} \) be the sequence of \( k \times k \) empirical transition matrices between the states. Assume a state 0 which denotes the environment outside the manpower system and that recruitment is done to replace wastage and to achieve the desired growth. By these assumptions, we have a Markov population replacement process [1–4]. Let \( \{\mathbf{w}(t_v)\}_{v=0}^{\gamma} \) be the sequence of \( 1 \times k \) vectors of observed loss probabilities and let \( \{\mathbf{p}_0(t_v)\}_{v=0}^{\gamma} \) be the sequence of \( 1 \times k \) vectors of observed recruitment probabilities. The entries in each of the sequences \( \{\mathbf{P}(t_v)\}_{v=0}^{\gamma} \), \( \{\mathbf{w}(t_v)\}_{v=0}^{\gamma} \), and \( \{\mathbf{p}_0(t_v)\}_{v=0}^{\gamma} \) are non-negative. The transition probability entries in each of the sequences \( \{\mathbf{P}(t_v)\}_{v=0}^{\gamma} \) and \( \{\mathbf{w}(t_v)\}_{v=0}^{\gamma} \) are estimated using the maximum likelihood method [5]. Each observed loss probability vector can be expressed in terms of the empirical transition matrix between the states as

\[
\mathbf{w}'(t_v) = (\mathbf{I} - \mathbf{P}(t_v)) \mathbf{e}',
\]

where \( \mathbf{I} \) is a \( k \times k \) identity matrix.

To extrapolate the structure of the system, we need a transition matrix which is stationary. Note that if the transition process is homogeneous, then the \( \eta \)-period transition matrix can be obtained by raising the one-period transition matrix to \( \eta \) index. Nonetheless, situations may arise where the structure of the system is required at any time instant, e.g., a non-integer period. In this case, raising the one-period transition matrix to a non-integer index may not give a substantive transition matrix, especially when the transition matrix is sparse [6]. This is a problem. In this paper we attempt a solution to the problem by expressing the sequence of sparse empirical stochastic matrices \( \{\mathbf{S}(t_v)\}_{v=0}^{\gamma} \) as the exponential of a matrix with row sums equal to zero and non-negative off-diagonal elements within some error distance in the Euclidean sense. This approach is closely related to the well-known embedding problem [6,7]. Stochastic matrices and sparse matrices have gained attention in the literature [8–10]. The embedding problem has been acknowledged to have practical relevance in the modelling of social phenomena [7], educational systems [11] and credit risk behaviour [6,12]. The key theoretical underpinnings of this paper are credited to [11]. The approach is centred on embedding the multi-step empirical stochastic matrix into a stationary continuous-time Markov chain. Several aspects of the continuous-time Markov chain have been studied in the literature. These include: the Markov decision process with discounted cost criterion [13], the stability and rate of convergence [14,15], the quasi-birth–death queues [16], the elapsed time between the chain observations [17] and the similarity class [18].

2. Methodology

Suppose the growth rate of the system is unknown to the researcher, but can be estimated from available data [11]. Let \( \hat{g} \) denote the estimated growth rate. Then, by the assumption that recruitment is done to replace wastage and to achieve the desired growth, \( \hat{g} \), the total recruits at \( t_{v+1} \), denoted as \( R(t_{v+1}) \), is expressed as

\[
R(t_{v+1}) = \mathbf{n}(t_v) \left( (1 + \hat{g}) \mathbf{I} - \mathbf{P}(t_v) \right) \mathbf{e}'.
\]

Using Eqs. (1) and (3), the relative structure of the system in terms of the observed structure \( \mathbf{q}(t_v) \) is expressed in a form analogous to [19] as

\[
\tilde{\mathbf{q}}(t_{v+1}) = \mathbf{q}(t_v) \mathbf{S}(t_v), \quad v = 0, 1, 2, \ldots, \gamma,
\]

where \( \mathbf{S}(t_v) \) is given as

\[
\mathbf{S}(t_v) = (\mathbf{P}(t_v) + ((1 - \mathbf{P}(t_v)) \mathbf{e}' + \hat{g} \mathbf{e}') \mathbf{p}_0(t_v)) (1 + \hat{g})^{-1}.
\]

The matrix \( \mathbf{S}(t_v) \) is stochastic as \( \mathbf{S}(t_v) \mathbf{e}' = \mathbf{e}' \). We determine the multi-step empirical stochastic matrix by rewriting Eq. (4) for each \( v = 0, 1, 2, \ldots, \gamma \), as [20]:

\[
\begin{align*}
\tilde{\mathbf{q}}(t_1) &= \mathbf{q}(t_0) \mathbf{S}(t_0) \\
\tilde{\mathbf{q}}(t_2) &= \mathbf{q}(t_0) \mathbf{S}(t_0) \mathbf{S}(t_1) \\
&\vdots \\
\tilde{\mathbf{q}}(t_{\gamma+1}) &= \mathbf{q}(t_0) \prod_{v=0}^{\gamma} \mathbf{S}(t_v).
\end{align*}
\]
The multi-step \((\gamma + 1)\)-stochastic matrix, \(\prod_{v=0}^{\gamma} S(t_v)\), is necessary to reduce the sparsity (i.e., the number of zero entries) in the one-period empirical stochastic matrix. A zero entry is usually as a result of a state which is not accessible from another state.

We now demonstrate how to obtain a stationary Markov transition model for the system based on the multi-step stochastic matrix. The stationary Markov transition model is vital to extrapolate the long-range shift in the structure of the system. We are interested in the state of the system at time indices: \(t_{\gamma+1}, t_{\gamma+2}, t_{\gamma+3}, \ldots\).

Let

\[
\prod_{v=0}^{\gamma} S(t_v) = \tilde{Q}.
\]

Equivalently,

\[
\prod_{v=0}^{\gamma} S(t_v) = \exp(\ln(\tilde{Q})).
\]

Now if there is a Markov generator, \(G\), satisfying \(\exp(G) = \tilde{Q}\), then Eq. (8) becomes

\[
\ln(\prod_{v=0}^{\gamma} S(t_v)) = G.
\]

The task is to solve the system in Eq. (9) such that the following properties of a Markov generator are satisfied for \(G\):

- the off-diagonal entries are non-negative;
- the diagonal entries are negative; and
- the row sum equals zero.

In practice, however, the matrix \(\ln(\prod_{v=0}^{\gamma} S(t_v))\) may violate the properties of a Markov generator. For instance, some of the off-diagonal elements may be negative. Whenever this is the case, we employ the diagonal adjustment regularization method \([12]\). We augment \(G = (g_{ij})\) say, to obtain a generator \(G^\alpha = (g_{ij}^\alpha)\) by setting:

\[
g_{ij}^\alpha = \max(g_{ij}, 0), \quad j \neq i,
\]

and

\[
g_{ij}^\alpha = g_{ii} + \sum_{j=1}^{k} \min(g_{ij}, 0).
\]

The Markov generator \(G^\alpha = (g_{ij}^\alpha)\) is obtained from either the sum or the partial sum of the infinite series arising from the expansion of \(\ln(\prod_{v=0}^{\gamma} S(t_v))\) about \(I\). It is worth noting that the Markov generator, \(G^\alpha\), may not be unique as more than one generator may be linked to the same transition matrix \([6]\). To circumvent this problem, we propose the following criteria: Suppose there are \(\varphi\) number of Markov generators. Then select \(G^* = G^\alpha\) with \(\zeta\) such that

\[
\zeta = \min_{\alpha} \| \exp(G^\alpha) - \prod_{v=0}^{\gamma} S(t_v) \|, \quad \alpha = 1, 2, 3, \ldots, \varphi.
\]

The expression \(\| \exp(G^\alpha) - \prod_{v=0}^{\gamma} S(t_v) \|\) is the Euclidean norm. Since \(\exp(G^*)\) is close to \(\prod_{v=0}^{\gamma} S(t_v)\) within some error distance, the state of the system at time indices \(t_{\gamma+1}, t_{\gamma+2}, t_{\gamma+3}, \ldots\), can therefore be represented as

\[
\tilde{q}(t_{\gamma+\tau}) = q(t_0) \exp(\delta t G^*), \quad \delta t = [t_{\gamma+\tau} - t_\gamma], \quad t_{\gamma+\tau} > t_\gamma, \quad \tau = 1, 2, 3, \ldots
\]

3. Numerical illustration

We examine data on academic staff flows in a faculty in the University of Benin, Nigeria. The categories of academic staff in the university encompass—graduate assistant, assistant lecturer (or assistant research fellow), lecturer II
We try the following methods and compute the Euclidean norm 

\[ \hat{g} \]

The growth rate is estimated as \( \hat{g} = 2.44\% \). We compute the sequence of stochastic matrices, \( \{S(t_i)\}_{i=0}^{5} \), and obtain

\[
\ln \left( \prod_{i=0}^{5} S(t_i) \right) = \begin{bmatrix}
-2.2820 & 2.7258 & -1.2930 & 0.8998 & 0.0120 & -0.5390 & 0.4764 \\
0.1826 & -4.2109 & 6.2929 & -2.3979 & -0.1952 & 0.1101 & 0.2184 \\
0.3090 & -0.1071 & -2.6963 & 3.1787 & -0.7021 & 0.8832 & -0.8653 \\
0.1978 & 0.1402 & -0.2544 & -1.7872 & 1.7363 & -0.6173 & 0.5845 \\
0.2006 & -0.1555 & 0.0535 & 0.1992 & -1.2248 & 1.6519 & -0.7250 \\
-0.0270 & 0.3355 & 0.0493 & -0.3456 & 0.2622 & -3.5121 & 3.2377 \\
0.6062 & 0.1919 & -0.2825 & 0.3029 & -0.0484 & 0.1153 & -0.8854
\end{bmatrix}.
\]

Clearly, the matrix, \( \ln \left( \prod_{i=0}^{5} S(t_i) \right) \), is not a valid Markov generator. So, we have to search for a valid generator.

We try the following methods and compute the Euclidean norm

\[
\left\| \exp(G^a) - \prod_{i=0}^{5} S(t_i) \right\|,
\]

for each of them.

- We apply the diagonal adjustment method to \( \ln \left( \prod_{i=0}^{5} S(t_i) \right) \), and then obtain a generator \( G^1 \) as

\[
G^1 = \begin{bmatrix}
-4.1140 & 2.7258 & 0 & 0.8998 & 0.0120 & 0 & 0.4764 \\
0.1826 & -6.8040 & 6.2929 & 0 & 0 & 0.1101 & 0.2184 \\
0.3090 & 0 & -4.3708 & 3.1787 & 0 & 0.8832 & 0 \\
0.1978 & 0.1402 & 0 & -2.6588 & 1.7363 & 0 & 0.5845 \\
0.2006 & 0 & 0.0535 & 0.1992 & -2.1052 & 1.6519 & 0 \\
0 & 0.3355 & 0.0493 & 0 & 0.2622 & -3.8847 & 3.2377 \\
0.6062 & 0.1919 & 0 & 0.3029 & 0 & 0.1153 & -1.2163
\end{bmatrix},
\]

with Euclidean norm of 0.6018.

- We set

\[
G^2 = \left( \prod_{i=0}^{5} S(t_i) - I \right),
\]

and obtain

\[
G^2 = \begin{bmatrix}
-0.8263 & 0.1601 & 0.3305 & 0.2963 & 0.0179 & 0.0051 & 0.0165 \\
0.0845 & -0.9300 & 0.2398 & 0.4574 & 0.1149 & 0.0190 & 0.0144 \\
0.0786 & 0.0516 & -0.9095 & 0.3467 & 0.3141 & 0.0695 & 0.0489 \\
0.0801 & 0.0518 & 0.0314 & -0.8088 & 0.3985 & 0.1285 & 0.1185 \\
0.0755 & 0.0489 & 0.0330 & 0.0427 & -0.6492 & 0.2080 & 0.2410 \\
0.1436 & 0.1100 & 0.0898 & 0.0784 & 0.0093 & -0.9538 & 0.5226 \\
0.1661 & 0.1278 & 0.1055 & 0.0996 & 0.0166 & 0.0066 & -0.5222
\end{bmatrix},
\]

with Euclidean norm of 0.5020.
where $\delta$ as

Next we apply the diagonal adjustment method to the partial sum

with Euclidean norm of 0.6553.

Next we apply the diagonal adjustment method to the partial sum

to get $G^4$ as

with Euclidean norm of 0.3218.

Among them, the Markov generator $G^4$ has the minimum Euclidean norm. Thus, we set $G^* = G^4$ with $\xi = 0.3218$.

Since data for the 2011/2012 academic session was as at February, 2012 instead of October, 2011, we set $t_5 = 1\frac{5}{12}$.

From Eq. (13), the stationary Markov system associated with the state of the system at time indices $t_6, t_7, t_8, \ldots$, is therefore represented as

$\bar{q}(t_{\tau + 5}) = \begin{bmatrix} 0.0792 & 0.1089 & 0.1584 & 0.2277 & 0.1882 & 0.0594 & 0.1782 \\ -1.7135 & 0.3920 & 0.8161 & 0.5045 & 0 & 0 & 0.0009 \\ 0.1531 & -1.8449 & 0.6182 & 1.0736 & 0 & 0 & 0 \\ 0.1308 & 0.0804 & -1.6931 & 0.8407 & 0.5826 & 0.0586 & 0 \\ 0.1231 & 0.0701 & 0.0032 & -1.4056 & 0.9005 & 0.2360 & 0.3347 \\ 0.0882 & 0.0432 & 0 & 0.0083 & -0.9802 & 0.5058 & 0.3347 \\ 0.2583 & 0.1935 & 0.0964 & 0.0145 & 0 & -1.7596 & 1.1969 \\ 0.3310 & 0.2501 & 0.1335 & 0.0526 & 0 & 0 & -0.7672 \end{bmatrix}$

where $\delta t = [t_{\tau + 5} - 1\frac{5}{12}], t_{\tau + 5} > 1\frac{5}{12}, \tau = 1, 2, 3, \ldots$.

As at November, 2012, i.e., $t_6 = 2\frac{1}{12}$, when the 2012/2013 academic session commenced, the manpower stock for the faculty is observed as

$\mathbf{n}(t_6) = \begin{bmatrix} 29 & 15 & 12 & 24 & 27 & 10 & 24 \end{bmatrix}$.

Using the stationary continuous-time Markov chain model in Eq. (14), the expected stock of the faculty is projected as

$\mathbf{\bar{n}}(t_6) = \begin{bmatrix} 9 & 9 & 12 & 21 & 24 & 9 & 21 \end{bmatrix}$. 
Table 1
A faculty manpower data at different time instants.

| $i ightarrow j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Losses | Stocks |
|------------------|---|---|---|---|---|---|---|--------|--------|
| **Recruits: $n_{0j}(t_0)$** |   |   |   |   |   |   |   |        |        |
| 1                | 2 | 3 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 3 | 8 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 6 | 10| 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 17| 6 | 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 18| 0 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 5 | 1 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 17|        |        |
| **$n_{0j}(t_1)$** |   |   |   |   |   |   |   |        |        |
| 1                | 2 | 3 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 10| 0 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 16| 1 | 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 23| 3 | 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 23| 1 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 2 | 2 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 15|        |        |
| **$n_{0j}(t_2)$** |   |   |   |   |   |   |   |        |        |
| 1                | 3 | 1 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 12| 0 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 14| 2 | 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 23| 0 | 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 26| 0 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 3 | 0 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 13|        |        |
| **$n_{0j}(t_3)$** |   |   |   |   |   |   |   |        |        |
| 1                | 7 | 0 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 7 | 8 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 11| 3 | 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 17| 8 | 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 20| 7 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 1 | 2 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 13|        |        |
| **$n_{0j}(t_4)$** |   |   |   |   |   |   |   |        |        |
| 1                | 8 | 3 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 1 | 6 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 12| 0 | 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 9 | 11| 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 19| 8 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 8 | 0 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 14|        |        |
| **$n_{0j}(t_5)$** |   |   |   |   |   |   |   |        |        |
| 1                | 12| 3 | 0 | 0 | 0 | 0 | 0 |        |        |
| 2                | 0 | 7 | 0 | 0 | 0 | 0 | 0 |        |        |
| 3                | 0 | 0 | 11| 1 | 0 | 0 | 0 |        |        |
| 4                | 0 | 0 | 0 | 22| 0 | 0 | 0 |        |        |
| 5                | 0 | 0 | 0 | 0 | 26| 3 | 0 |        |        |
| 6                | 0 | 0 | 0 | 0 | 0 | 7 | 0 |        |        |
| 7                | 0 | 0 | 0 | 0 | 0 | 0 | 14|        |        |

where

$$\bar{q}(t_6) = [0.0882 \ 0.0818 \ 0.1145 \ 0.2022 \ 0.2297 \ 0.0843 \ 0.1994].$$

Comparing the expected stock with the actual stock, we find that the model gives a very poor fit for two grades – graduate assistant and assistant lecturer, while the fit is very good for the remaining five grades – lecturer II, lecturer
I, senior lecturer, associate professor and professor. Nonetheless, we do not expect our model to give the exact fit as we have earlier regularized the Markov generator using the diagonal adjustment method.

4. Concluding remarks

This study has focused on identifying a stationary continuous-time Markov chain from a multi-step empirical stochastic matrix for Markovian manpower systems. Each one-step empirical stochastic matrix is sparse and it is obtained at any choice of time instants (including non-integral time instants). Consequent on the possibility of non-integral time instants, we conclude that the appropriate technical apparatus to extrapolate the long-range shifts in the structure of the system is the stationary continuous-time Markov chain. This is our main contribution. This contribution opens a new window in Markov population replacement processes [1–4] and it is in line with [11,6,12]. More so, the multi-step empirical stochastic matrix may not have a valid Markov generator because the one-step empirical stochastic matrix is sparse. As a result, we make adjustments to find a Markov generator within some error distance to the multi-step empirical stochastic matrix. We have demonstrated the utility of our proposed method for a faculty academic staff manpower system. It should be noted that, although the closest rivalry to this study is [11], there are points of departure, in that in the present study, the system is not constrained by the carrying capacity (i.e. there is enough room for expansion of the workforce), the probability estimates are not constrained by the logistic equation, and the application area differs.

Acknowledgement

The author would like to thank the anonymous referee whose comments improved the original version of this manuscript.

Appendix. MATLAB source codes for the continuous-time Markov chain model

clc
% Data declaration
%
% Manpower data for Chemistry.
ch11=[2 1 1 1 1 1]; ch12=[0 1 0 0 0 0]; ch10=[0 0 0 0 0 0];
ch22=[1 1 2 0 0 1]; ch23=[0 0 0 2 0 0]; ch20=[0 0 0 0 0 0];
ch33=[3 3 3 3 2 2]; ch34=[0 0 0 0 2 0]; ch30=[0 0 0 0 1 0];
ch44=[3 3 3 2 1 3]; ch45=[0 0 0 1 1 0]; ch40=[0 0 0 0 0 0];
ch55=[8 8 8 4 3 4]; ch56=[0 0 0 4 1 0]; ch50=[0 0 0 0 1 0];
ch66=[1 1 1 0 4 0]; ch67=[0 0 0 1 0 5]; ch60=[0 0 0 0 0 0];
ch77=[9 8 5 5 6 6]; ch70=[0 1 3 1 0 0];
chs1=[2 2 1 1 1 1]; chs2=[1 1 2 2 0 1]; chs3=[3 3 3 3 5 2];
chs4=[3 3 3 3 2 3]; chs5=[8 8 8 5 4]; chs6=[1 1 1 4 5];
chs7=[9 9 8 6 6 6];
rch1=[0 0 0 0 0 0]; rch2=rch1; rch3=[0 0 0 0 0 0 0]; rch4=rch1;
rch5=[0 1 0 0 0 0]; rch6=[4 2 0 1 0 1];
%
% Manpower data for Computer Science.

c11=[0 0 0 0 0 0]; c12=[3 0 0 0 0 0]; c10=[1 0 0 0 0 0];
c22=[2 6 5 1 0 2]; c23=[3 0 0 4 1 0]; c20=[0 0 2 0 0 0];
c33=[0 5 5 4 2 3]; c34=[4 1 0 1 6 0]; c30=[0 0 0 0 0 0];
c44=[5 9 10 8 2 9]; c45=[0 0 0 2 7 0]; c40=[0 0 0 0 0 0];
c55=[1 1 1 1 2 9]; c56=[0 0 0 0 1 0]; c50=[0 0 0 0 0 0];
c66=[0 0 0 0 0 0]; c67=[0 0 0 0 0 0]; c60=[0 0 0 0 0 0];
c77=[1 1 1 1 1 1]; c70=[0 0 0 0 0 0];
cs1=[4 0 0 0 0 0]; cs2=[5 6 7 5 1 2]; cs3=[4 6 5 5 8 3];
cs4=[5 9 10 10 9 9]; cs5=[1 1 1 1 3 9]; cs6=[0 0 0 0 0 1];
cs7=[1 1 1 1 1 1];
rc1=[0 1 3 0 0 0 0]; rc2=[0 1 0 0 0 0 0]; rc3=rc2;
rc4=[0 0 0 0 0 0 0]; rc5=[0 2 0 1 0 0 0]; rc6=[5 2 0 0 0 0 0];

disp('Manpower data for Geology.')
g11=[0 0 0 0 2 6]; g12=[0 0 0 0 0 0]; g10=[0 0 0 0 0 0];
g22=[0 0 0 0 1 1]; g23=[0 0 0 0 0 0]; g20=[0 0 0 0 0 0];
g33=[3 3 3 2 2 1]; g34=[0 0 0 1 0 1]; g30=[0 0 0 0 0 0];
g44=[1 1 1 0 1 1]; g45=[0 0 0 1 0 0]; g40=[0 0 0 0 0 0];
g55=[4 4 4 3 4 3]; g56=[0 0 0 1 0 1]; g50=[1 0 0 0 0 0];
g66=[2 1 1 0 1 0]; g67=[0 1 0 1 0 1]; g60=[0 0 0 0 0 0];
g77=[2 2 3 3 3 3]; g70=[1 0 0 0 1 0];
gs1=[0 0 0 0 2 6]; gs2=[0 0 0 0 1 1]; gs3=[3 3 3 3 2 2];
gs4=[1 1 1 1 1 1]; gs5=[5 4 4 4 4 4]; gs6=[2 2 1 1 1 1];
gs7=[3 2 3 3 4 3];
rg1=[0 0 0 0 0 0 0]; rg2=rg1; rg3=rg1; rg4=[2 1 0 0 0 0 0];
rg5=[4 0 0 0 0 0 0];

disp('Manpower data for Mathematics.')
m11=[0 1 2 3 3 3]; m12=[2 0 1 0 2 1]; m10=[0 0 0 0 0 0];
m22=[0 2 4 3 0 2]; m23=[3 0 0 2 2 0]; m20=[0 0 0 0 1 0];
m33=[0 3 1 0 0 2]; m34=[6 0 2 1 2 0]; m30=[0 0 0 0 0 0];
m44=[3 7 7 6 5 7]; m45=[5 2 0 3 2 0]; m40=[0 0 0 0 0 0];
m55=[3 7 9 7 5 6]; m56=[0 1 0 2 5 1]; m50=[0 0 0 0 0 0];
m66=[0 0 1 1 3 6]; m67=[1 0 0 0 0 2]; m60=[0 0 0 0 0 0];
m77=[3 4 3 3 3 3]; m70=[0 0 1 0 0 0];
ms1=[2 1 3 3 5 4]; ms2=[3 2 4 5 3 2]; ms3=[6 3 3 1 2 2];
ms4=[8 9 7 9 7 7]; ms5=[3 8 9 9 10 7]; ms6=[1 0 1 1 3 8];
ms7=[3 4 4 3 3 3];
rm1=[1 0 0 0 0 0]; rm2=[2 2 0 0 0 0 0]; rm3=rm1;
rm4=[2 0 0 0 0 0 0];

disp('Manpower data for Physics.')
ph11=[0 0 0 3 2 2]; ph12=[0 0 0 0 1 2]; ph10=[0 0 0 0 0 0];
ph22=[0 1 1 3 0 1]; ph23=[2 0 0 0 3 0]; ph20=[0 0 0 0 0 0];
ph33=[0 2 2 2 0 3]; ph34=[0 0 0 0 2 0]; ph30=[0 0 0 0 0 0];
ph44=[5 3 2 1 0 2]; ph45=[1 1 0 1 1 0]; ph40=[0 1 0 1 0 0];
ph55=[2 3 4 5 5 4]; ph56=[0 0 0 0 1 1]; ph50=[0 0 0 0 0 1];
ph66=[2 0 0 0 0 1]; ph67=[0 1 0 0 0 0]; ph60=[0 1 0 0 0 0];
ph77=[2 0 1 1 1 1]; ph70=[0 2 0 0 0 0];
phs1=[0 0 0 3 3 4]; phs2=[2 1 1 3 3 1]; phs3=[0 2 2 2 2 3];
phs4=[6 5 3 2 1 2]; phs5=[2 3 4 5 6 6]; phs6=[2 2 0 0 0 1];
phs7=[2 2 1 1 1 1];
rph1=[0 1 0 0 0 0 0]; rph2=[0 0 0 0 0 0 0]; rph3=[3 0 0 0 0 0 0];
rph4=rph2; rph5=[2 0 0 0 0 0 0]; rph6=[5 0 0 0 0 0 0];

disp('Faculty manpower data')
f11=ch11+c11+g11+m11+ph11, f12=ch12+c12+g12+m12+ph12,
f10=ch10+c10+g10+m10+ph10, f22=ch22+c22+g22+m22+ph22,
f23=ch23+c23+g23+m23+ph23, f20=ch20+c20+g20+m20+ph20,
f33=ch33+c33+g33+m33+ph33, f34=ch34+c34+g34+m34+ph34,
f30=ch30+c30+g30+m30+ph30, f44=ch44+c44+g44+m44+ph44,
f45 = ch45 + g45 + m45 + ph45, f40 = ch40 + g40 + m40 + ph40,
f55 = ch55 + g55 + m55 + ph55, f56 = ch56 + g56 + m56 + ph56,
f50 = ch50 + g50 + m50 + ph50, f66 = ch66 + g66 + m66 + ph66,
f67 = ch67 + g67 + m67 + ph67, f60 = ch60 + g60 + m60 + ph60,
f77 = ch77 + g77 + m77 + ph77, f70 = ch70 + g70 + m70 + ph70,
fs1 = chs1 + cs1 + gs1 + ms1 + phs1, fs2 = chs2 + cs2 + gs2 + ms2 + phs2,
fs3 = chs3 + cs3 + gs3 + ms3 + phs3, fs4 = chs4 + cs4 + gs4 + ms4 + phs4,
fs5 = chs5 + cs5 + gs5 + ms5 + phs5, fs6 = chs6 + cs6 + gs6 + ms6 + phs6,
fs7 = chs7 + cs7 + gs7 + ms7 + phs7,
rf = rch1 + rcl + rgl + rm1 + rph1, rfb = rch2 + rcl + rgl + rm2 + rph2,
rfc = rch3 + rcl + rgl + rm3 + rph3, rfd = rch4 + rcl + rgl + rm4 + rph4,
rf = rch5 + rcl + rgl + rm5 + rph5, rff = rch6 + rcl + rgl + rm6 + rph6,

disp('The growth factor.')</s> = 7; e = ones(s, 1);
E1 = [fs1(1,1) fs2(1,1) fs3(1,1) fs4(1,1) fs5(1,1) fs6(1,1) fs7(1,1)];
E2 = [fs1(1,2) fs2(1,2) fs3(1,2) fs4(1,2) fs5(1,2) fs6(1,2) fs7(1,2)];
E3 = [fs1(1,3) fs2(1,3) fs3(1,3) fs4(1,3) fs5(1,3) fs6(1,3) fs7(1,3)];
E4 = [fs1(1,4) fs2(1,4) fs3(1,4) fs4(1,4) fs5(1,4) fs6(1,4) fs7(1,4)];
E5 = [fs1(1,5) fs2(1,5) fs3(1,5) fs4(1,5) fs5(1,5) fs6(1,5) fs7(1,5)];
E6 = [fs1(1,6) fs2(1,6) fs3(1,6) fs4(1,6) fs5(1,6) fs6(1,6) fs7(1,6)];
a = ones(6, 1);
N = [log(E1*e) log(E2*e) log(E3*e) log(E4*e) log(E5*e) log(E6*e)]',
T = length(N); c = [1:T]';
g = exp([0 1]*(inv([a c]'*[a c]')*[a c]'*N))-1,

disp('Empirical probabilities')
pf11 = f11./fs1; pf12 = f12./fs1; pf10 = f10./fs1; pf22 = f22./fs2;
pf23 = f23./fs2; pf20 = f20./fs2; pf33 = f33./fs3; pf34 = f34./fs3;
pf30 = f30./fs3; pf44 = f44./fs4; pf45 = f45./fs4; pf40 = f40./fs4;
pf55 = f55./fs5; pf65 = f65./fs5; pf50 = f50./fs5; pf66 = f66./fs6;
pf67 = f67./fs6; pf60 = f60./fs6; pf77 = f77./fs7; pf70 = f70./fs7;

disp('Empirical matrices & vectors')
Pfa = [pf11(1,1) pf12(1,1) 0 0 0 0 0; 0 pf22(1,1) pf23(1,1) 0 0 0 0;
 0 0 pf33(1,1) pf34(1,1) 0 0 0; 0 0 0 pf44(1,1) pf45(1,1) 0 0;
 0 0 0 0 pf55(1,1) pf56(1,1) 0 0; 0 0 0 0 0 pf66(1,1) pf67(1,1);
 0 0 0 0 0 0 pf77(1,1)],
Wfa = [pf10(1,1) pf20(1,1) pf30(1,1) pf40(1,1) pf50(1,1) pf60(1,1)... pf70(1,1)],
Rfa = rfa/sum(rfa), I = eye(s);
Sa = (Pfa - (I-Pfa)*e+g*e)*Rfa)*inv(1+g), SP1 = Sa*e,
Pfb = [pf11(1,2) pf12(1,2) 0 0 0 0 0; 0 pf22(1,2) pf23(1,2) 0 0 0 0;
 0 0 pf33(1,2) pf34(1,2) 0 0 0; 0 0 0 pf44(1,2) pf45(1,2) 0 0;
 0 0 0 0 pf55(1,2) pf56(1,2) 0 0; 0 0 0 0 0 pf66(1,2) pf67(1,2);
 0 0 0 0 0 0 pf77(1,2)],
Wfb = [pf10(1,2) pf20(1,2) pf30(1,2) pf40(1,2) pf50(1,2) pf60(1,2)... pf70(1,2)],
Rfb = rfb/sum(rfb),
Sb = (Pfb - (I-Pfb)*e+g*e)*Rfb)*inv(1+g), SP2 = Sb*e,
Pfc = [pf11(1,3) pf12(1,3) 0 0 0 0 0; 0 pf22(1,3) pf23(1,3) 0 0 0 0;
0 0 pf33(1,3) pf34(1,3) 0 0 0; 0 0 0 pf44(1,3) pf45(1,3) 0 0;
0 0 0 pf55(1,3) pf56(1,3) 0; 0 0 0 0 pf66(1,3) pf67(1,3);
0 0 0 0 0 pf77(1,3)],
Wfc=[pf10(1,3) pf20(1,3) pf30(1,3) pf40(1,3) pf50(1,3) pf60(1,3)...
pf70(1,3)],
Rfc=rfc/sum(rfc),
Sc=(Pfc+((I-Pfc)*e+g*e)*Rfc)*inv(1+g), SP3=Sc*e,

Pfd=[pf11(1,4) pf12(1,4) 0 0 0 0; 0 pf22(1,4) pf23(1,4) 0 0 0 0;
0 0 pf33(1,4) pf34(1,4) 0 0 0; 0 0 0 pf44(1,4) pf45(1,4) 0 0;
0 0 0 pf55(1,4) pf56(1,4) 0; 0 0 0 0 pf66(1,4) pf67(1,4);
0 0 0 0 0 pf77(1,4)],
Wfd=[pf10(1,4) pf20(1,4) pf30(1,4) pf40(1,4) pf50(1,4) pf60(1,4)...
pf70(1,4)],
Rfd=rfd/sum(rfd),
Sd=(Pfd+((I-Pfd)*e+g*e)*Rfd)*inv(1+g), SP4=Sd*e,

Pfe=[pf11(1,5) pf12(1,5) 0 0 0 0; 0 pf22(1,5) pf23(1,5) 0 0 0 0;
0 0 pf33(1,5) pf34(1,5) 0 0 0; 0 0 0 pf44(1,5) pf45(1,5) 0 0;
0 0 0 pf55(1,5) pf56(1,5) 0; 0 0 0 0 pf66(1,5) pf67(1,5);
0 0 0 0 0 pf77(1,5)],
Wfe=[pf10(1,5) pf20(1,5) pf30(1,5) pf40(1,5) pf50(1,5) pf60(1,5)...
pf70(1,5)],
Rfe=rfe/sum(rfe),
Se=(Pfe+((I-Pfe)*e+g*e)*Rfe)*inv(1+g), SP5=Se*e,

Pff=[pf11(1,6) pf12(1,6) 0 0 0 0; 0 pf22(1,6) pf23(1,6) 0 0 0 0;
0 0 pf33(1,6) pf34(1,6) 0 0 0; 0 0 0 pf44(1,6) pf45(1,6) 0 0;
0 0 0 pf55(1,6) pf56(1,6) 0; 0 0 0 0 pf66(1,6) pf67(1,6);
0 0 0 0 0 pf77(1,6)],
Wff=[pf10(1,6) pf20(1,6) pf30(1,6) pf40(1,6) pf50(1,6) pf60(1,6)...
pf70(1,6)],
Rff=rff/sum(rff),
Sf=(Pff+((I-Pff)*e+g*e)*Rff)*inv(1+g), SP6=Sf*e,
Saf=Sa*Sb*Sc*Sd*Se*Sf, SP7=Saf*e,

G0=logm(Saf), s0=G0*e,

disp('The logarithmic matrix.')
G0 = [-2.2820 2.7258 -1.2930 0.8998 0.0120 -0.5390 0.4764;
0.1826 -4.2109 6.2929 -2.3979 -0.1952 0.1101 0.2184;
0.3090 -0.1071 -2.6963 3.1787 -0.7021 0.8832 -0.8653;
0.1978 0.1402 -0.2544 -1.7872 1.7363 -0.6173 0.5845;
0.2006 -0.1555 0.0535 0.1992 -1.2248 1.6519 -0.7250;
-0.0270 0.3355 0.0493 -0.3456 0.2622 -3.5121 3.2377;
0.6062 0.1919 -0.2825 0.3029 -0.0484 0.1153 -0.8854];

G1 = [ (-2.2820-1.2930-0.5390) 2.7258 0 0.8998 0.0120 0 0.4764;
0.1826 (-4.2109-2.3979-0.1952) 6.2929 0 0 0.1101 0.2184;
0.3090 0 (-0.1071-2.6963-0.7021-0.8653) 3.1787 0 0.8832 0;
0.1978 0.1402 0 (-0.2544-1.7872-0.6173+0.1*1.0e-003) 1.7363 0 0.5845;
0.2006 0 0.0535 0.1992 (-0.1555-1.2248-0.7250+0.1*1.0e-003) 1.6519 0; 0 0.3355 0.0493 0 0.2622 (-0.0270-0.3456-3.5121) 3.2377; 0.6062 0.1919 0 0.3029 0 0.1153 (-0.2825-0.0484-0.8854) , s2=G1*e, Eu1=norm(expm(G1)-Saf), G2=(Saf-I), s2=G2*e, Eu2=norm(expm(G2)-Saf), G3=G1+G2, s3=G3*e, Eu3=norm(expm(G3)-Saf), G4a=(Saf-I)-((Saf-I)∧2)/2+((Saf-I)∧3)/3; G4 =([-1.4267-0.2339-0.0530+0.1*1.0e-003] 0.3920 0.8161 0.5045 0 0 0.0009; 0.1531 (-1.6939-0.0367-0.0604-0.0540+1.0e-003-0.9*1.0e-003) 0.9005 0.2360 0.0727; 0.0882 0.0432 0 0.0083 (-0.9745-0.0058+0.1*1.0e-003) 0.5058 0.3347; 0.2583 0.1935 0.0964 0.0145 0 (-1.7150-0.0446) 1.1969; 0.3310 0.2501 0.1335 0.0526 0 0 (-0.0441-0.0039-0.7191-0.1*1.0e-003), Eu4=norm(expm(G4)-Saf), disp('Goodness of fit') q0 = [0.0792 0.1089 0.1584 0.2277 0.1882 0.0594 0.1782]; q6=q0*expm((25/12-17/12)*G4), n6=ceil(sum(E1)*q6), E=[29 15 12 24 27 10 24], x2=sum(((E-n6).∧2)./n6), Ea=[n6(1,1) n6(1,1) 12 24 27 10 24], x2a=sum(((Ea-n6).∧2)./n6),

References