On perturbations of non-diagonalizable stochastic matrices of order 3

P.J. PAUWELYN\textsuperscript{a} and M.A. GUERRY\textsuperscript{b}

Abstract

In this paper, the interest lies with the non-diagonalizable stochastic matrices. We show that it is possible for every non-diagonalizable stochastic $3 \times 3$ matrix to be perturbed into a diagonalizable stochastic matrix with the eigenvalues, arbitrarily close to the eigenvalues of the original matrix. Moreover, this perturbed matrix is a stochastic matrix with the same principal (left and right) eigenspaces as the original matrix. An algorithm is presented to determine a perturbation matrix, which preserves these spectral properties. Additionally, a relation is proved between the eigenvectors and generalized eigenvectors of the original matrix and the perturbed matrix.

Keywords: Stochastic matrices; Non-diagonalizable matrices; Perturbation theory; Markov chains

MSC: 15B51, 15A18, 47A55, 60J10

1 Introduction and main result

In general, perturbation theory deals with the following question \cite{2,5,7,8,11,12,13}: What happens to certain matrix quantities or properties if the matrix is perturbed in a certain way? The present paper examines another question: Which perturbation do we need in order to attain certain matrix properties? In particular, we start from a non-diagonalizable stochastic $3 \times 3$ matrix $A$ with eigenvalues $1$ and $\lambda$. We determine a perturbation matrix $E$ such that the perturbed matrix $\tilde{A}$ suffices the following conditions:

(a) $\tilde{A} = A + E$ is an additive perturbation of $A$.

(b) $\tilde{A}$ is a stochastic matrix.

(c) $\text{spec}(\tilde{A}) = \{1, \mu, \nu\} \subset \mathbb{R}$, $\mu \neq \nu$ and $\forall \delta > 0$, $\exists E \in \mathbb{R}^{3 \times 3}$: $|\mu - \lambda| < \delta \land |\nu - \lambda| < \delta$.

(d) The eigenspaces corresponding to the eigenvalue $1$ of the matrices $A$ and $\tilde{A}$ coincide.

These conditions are all chosen with good reason. Condition (a) allows to directly see and interpret the effects of a perturbation on $\tilde{A}$, especially in an applied situation. By condition (b), after perturbation, the matrix remains stochastic. Condition (c) demands that $\tilde{A}$ is a diagonalizable matrix with real eigenvalues. Each eigenvalue of $\tilde{A}$ can be arbitrary close to an eigenvalue of $A$, depending on the perturbation matrix $E$. Condition (d) implies that the principal (left and right) eigenspaces of $A$ and $\tilde{A}$ coincide. A side-effect of this condition is that the corresponding Markov chains of $A$ and $\tilde{A}$ have the same stationary distribution.

We obtained the result that for all non-diagonalizable stochastic $3 \times 3$ matrices $A$ it is possible to find a perturbation matrix $E$ such that $A + E$ suffices the conditions (a)-(d).

2 Construction of a perturbation matrix

We have to translate the conditions (a)-(d) into conditions on the perturbation matrix $E$ or conditions on the matrix elements $e_{ij}$ of $E$. We consider additive perturbations of $A$ such that the stochastic property of $A$ can be easily preserved. In order to preserve the rowsums of $A$, the condition $A\bar{1} = \tilde{A}\bar{1}$, (where $\bar{1} \in \mathbb{R}^3$ is the columnvector with all components equal to $1$) must be fulfilled. This equality implies

$$E\bar{1} = (\tilde{A} - A)\bar{1} = \tilde{A}\bar{1} - A\bar{1} = 0$$

\textsuperscript{a}Pieter-Jan.Pauwelyn@vub.be, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
\textsuperscript{b}Marie-Anne.Guerry@vub.be, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
Therefore the rowsums of the perturbation matrix $E$ should be zero. Otherwise said:

$$\forall i \in \{1, 2, 3\}: \sum_{j=1}^{3} e_{ij} = 0 \tag{1}$$

Implementing condition (d) goes as follows. The eigenspaces corresponding to the eigenvalue 1, are the left eigenspace $V_L(1) = \{ v \in \mathbb{R}^{1 \times 3} | v(A - I_3) = 0 \}$ and the right eigenspace $V_R(1) = \{ w \in \mathbb{R}^{3 \times 1} | (A - I_3)w = 0 \}$, with $I_3$ the identity matrix of order 3. From here on we use the notation $V_L(\lambda)$ for the left eigenvalue $\lambda$ of the matrix $A$ and $V_R(\lambda)$ for the right eigenvalue corresponding to the eigenvalue $\lambda$ of the matrix $A$. For the left and right eigenspaces corresponding to the perturbed matrix $\tilde{A}$, we use the notation $\tilde{V}_L(\lambda)$ and $\tilde{V}_R(\lambda)$. Since $\tilde{A}$ must have the same eigenspaces as the matrix $A$ corresponding to the eigenvalue 1, the equalities $v(\tilde{A} - I_3)w = 0$ should hold for all $v \in V_L(1)$ and for all $w \in V_R(1)$. Since $A$ and $\tilde{A}$ are both stochastic matrices, their right eigenspaces are both $V_R(1) = \{ (k,k,k)^T | k \in \mathbb{R} \}$ and coincide. To obtain coinciding left eigenspaces $V_L(1)$ and $\tilde{V}_L(1)$, the perturbation matrix $E$ should suffice the equality $v(A + E - I_3) = 0$, for all $v \in V_L(1)$. A method to determine such $E$ is given by the following lemma:

**Lemma 2.1.** Given a matrix $A$ with an eigenvalue $\lambda$ and corresponding left eigenvector $u$ and a perturbation matrix $E$. Then $\lambda$ is also an eigenvalue of $A + E$ with corresponding left eigenvector $u$ if and only if $uE = 0$.

**Proof.** If $\lambda$ is an eigenvalue of $A + E$ with corresponding left eigenvector $u$, then $u(A + E) = \lambda u$. Consequently (since $uA = \lambda u$):

$$uE = u(A + E) - uA = \lambda u - \lambda u = 0$$

Vice versa, if $uE = 0$, then:

$$u(A + E) = uA + uE = \lambda u + 0 = \lambda u$$

With this lemma, condition (d), the preservation of the eigenspaces corresponding to the eigenvalue 1 becomes $v_LE = 0$ for $v_L \in V_L(1)$. If $v_LE = 0$ for some $v_L \in V_L(1)$, then $vE = 0$ for all $v \in V_L(1)$, since $V_L(1)$ is one-dimensional. This constraint $vE = 0$ combined with the constraint $E1 = 0$ implies the following form of $E$:

$$E = \begin{pmatrix} v_2(\epsilon_{22} + \epsilon_{23}) + v_3(\epsilon_{32} + \epsilon_{33}) & -v_2\epsilon_{22} - v_3\epsilon_{32} & -v_2\epsilon_{23} - v_3\epsilon_{33} \\ -v_1(\epsilon_{22} + \epsilon_{23}) & v_1\epsilon_{22} & v_1\epsilon_{23} \\ -v_1(\epsilon_{32} + \epsilon_{33}) & v_1\epsilon_{32} & v_1\epsilon_{33} \end{pmatrix} \tag{2}$$

with $v = (v_1, v_2, v_3)$ for some $v \in V_L(1)$.

The eigenvalues of the perturbed matrix $\tilde{A}$ should remain real after the perturbation of $A$. $A$ has two real eigenvalues 1 and $\lambda$, 1 is always a semisimple eigenvalue[9]. Therefore $\lambda$ must have multiplicity 2, if $A$ is non-diagonalizable[4]. In order for $\tilde{A}$ to have three distinct eigenvalues, its characteristic equation needs to have 3 distinct roots. Since $\tilde{A}$ is also a stochastic matrix, 1 is an eigenvalue of $\tilde{A}$.

For the characteristic equation of $\tilde{A}$, we have:

$$-\lambda^3 + tr(\tilde{A})\lambda^2 + [1 - tr(\tilde{A}) - det(\tilde{A})]\lambda + det(\tilde{A}) = 0$$

$$\Leftrightarrow (\lambda - 1) \left[ -\lambda^2 + (tr(\tilde{A}) - 1)\lambda - det(\tilde{A}) \right] = 0$$

For $\tilde{A}$ to have three distinct real eigenvalues, the equation $-\lambda^2 + (tr(\tilde{A}) - 1)\lambda - det(\tilde{A}) = 0$ must have two distinct real roots. This is only the case if the discriminant $\Delta(\tilde{A})$ is strictly positive.

$$\Delta(\tilde{A}) = \left[ tr(\tilde{A}) - 1 \right]^2 - 4 \det(\tilde{A}) > 0 \tag{3}$$

We consider the discriminant of a stochastic matrix as a function of its matrixelements. Therefore we define the set $U = \{ (b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}) | \forall i,j : b_{ij} \geq 0 \land b_{12} + b_{13} \leq 1 \}$, which is isomorphic with the set of the stochastic $3 \times 3$ matrices. Now, we can define the discriminant in function of six variables on the subset $U$ of $\mathbb{R}^6$.

$$\Delta : U \rightarrow \mathbb{R} : (b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}) \mapsto \Delta(B) = \Delta \begin{pmatrix} 1 - b_{12} - b_{13} & b_{12} & b_{13} \\ 1 - b_{22} - b_{23} & b_{22} & b_{23} \\ 1 - b_{32} - b_{33} & b_{32} & b_{33} \end{pmatrix}$$
Next, the discriminant function can be written in fully as:

\[
\Delta(B) = \Delta \begin{pmatrix} 1 - b_{12} - b_{13} & b_{12} & b_{13} \\ 1 - b_{22} - b_{23} & b_{22} & b_{23} \\ 1 - b_{32} - b_{33} & b_{32} & b_{33} \end{pmatrix} = \left[\text{tr}(B) - 1\right]^2 - 4\det(B) \\
= (b_{12} - b_{13} - b_{22} + b_{33})^2 + 4 (b_{12} - b_{32}) (b_{13} - b_{23})
\]

According to (3), the discriminant of \( \hat{A} \) must be positive. To determine what happens with the discriminant, when \( A \) is perturbed into \( \hat{A} = A + E \), the directional derivative of \( \Delta(A) \) will be examined. Since the rowsums of \( E \) must be 0 (see (1)), \( E \) can be mapped on a direction \( \vec{\epsilon} = (\epsilon_{12}, \epsilon_{13}, \epsilon_{22}, \epsilon_{23}, \epsilon_{32}, \epsilon_{33}) \in \mathbb{R}^6 \). The directional derivative[1] of a non-diagonalizable stochastic 3 × 3 matrix \( B \) in a direction \( \vec{\epsilon} \), \( D_\vec{\epsilon}\Delta(B) \) is

\[
D_\vec{\epsilon}\Delta(B) = D_\vec{\epsilon}(b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}) = \frac{\vec{\epsilon}}{||\vec{\epsilon}||} \cdot \left( \frac{\partial \Delta}{\partial b_{12}}, \frac{\partial \Delta}{\partial b_{13}}, \frac{\partial \Delta}{\partial b_{22}}, \frac{\partial \Delta}{\partial b_{23}}, \frac{\partial \Delta}{\partial b_{32}}, \frac{\partial \Delta}{\partial b_{33}} \right)
\]

\[
= \frac{2}{||\vec{\epsilon}||} [(b_{12} - b_{13} - b_{22} + b_{33}) (\epsilon_{12} - \epsilon_{13} - \epsilon_{22} + \epsilon_{33}) + 2 (b_{13} - b_{23}) (\epsilon_{12} - \epsilon_{32}) + 2 (b_{12} - b_{32}) (\epsilon_{13} - \epsilon_{23})]
\]

If the calculation of the directional derivative is applied on a direction \( \vec{\epsilon} \) corresponding to \( E \) as in (2), we have that the directional derivative of the discriminant is a function of four perturbation variables, \( \epsilon_{22}, \epsilon_{23}, \epsilon_{32} \) and \( \epsilon_{33} \) and then the directional derivative becomes:

\[
D_\vec{\epsilon}\Delta(A) = (a_{12} - a_{13} - a_{22} + a_{33})(-v_{22}v_{33} - v_{23}v_{32} - v_{32}v_{12} - v_{12}v_{13}) + 2(a_{13} - a_{23}) (-v_{22}v_{33} - v_{23}v_{32} - v_{32}v_{12} - v_{12}v_{13}) + 2(a_{12} - a_{32}) (-v_{22}v_{33} - v_{23}v_{32} - v_{32}v_{12} - v_{12}v_{13})
\]

\[
= K \epsilon_{22} + L \epsilon_{23} + M \epsilon_{32} + N \epsilon_{33}
\]

with

\[
K = (a_{12} - a_{13} - a_{22} + a_{33})(v_1 + v_2) + 2(a_{13} - a_{23})v_2 \in \mathbb{R}
\]

\[
L = (a_{12} - a_{13} - a_{22} + a_{33})v_2 - 2(a_{12} - a_{32})v_1 \in \mathbb{R}
\]

\[
M = (a_{12} - a_{13} - a_{22} + a_{33})v_3 + 2(a_{13} - a_{23})(v_1 + v_3) \in \mathbb{R}
\]

\[
N = (a_{12} - a_{13} - a_{22} + a_{33})(v_1 + v_3) - 2(a_{12} - a_{32})v_3 \in \mathbb{R}
\]

From here, we consider the direction \( \vec{\epsilon} \) to be the vector \((\epsilon_{22}, \epsilon_{23}, \epsilon_{32}, \epsilon_{33})\).

For the given non-diagonalizable stochastic matrix \( A \), we need to determine a direction \( \vec{\epsilon} \) such that \( D_\vec{\epsilon}\Delta(A) > 0 \). This means solving the following inequality:

\[
D_\vec{\epsilon}\Delta(A) > 0
\]

Since \( D_\vec{\epsilon}\Delta(A) \) is a homogeneous polynomial of degree 1 for its variables \( \epsilon_{22}, \epsilon_{23}, \epsilon_{32} \) and \( \epsilon_{33} \), the inequality \( D_\vec{\epsilon}\Delta(A) > 0 \) has a solution, unless \( K = L = M = N = 0 \). If \( K = L = M = N = 0 \), then the second order directional derivative \( D^2_\vec{\epsilon}\Delta(A) \) must be calculated[1]. If the first order directional derivative \( D_\vec{\epsilon}\Delta(A) \) is zero in a direction and the second order directional derivative \( D^2_\vec{\epsilon}\Delta(A) \) is positive, then \( \Delta(A) \) will increase in the direction of \( \vec{\epsilon} \). The inequality for \( D^2_\vec{\epsilon}\Delta(A) \) is:

\[
D^2_\vec{\epsilon}\Delta(A) = (\epsilon_{22}(v_1 + v_2) + \epsilon_{23}v_2 + \epsilon_{32}v_3 + \epsilon_{33}(v_1 + v_3))^2 + 4(\epsilon_{22}\epsilon_{23} - \epsilon_{22}\epsilon_{33})v_1(v_1 + v_2 + v_3) > 0
\]

The left hand side is a non-zero homogeneous polynomial of degree 2, thus the existence of a solution to this inequality is guaranteed.

Before we can conclude that for every non-diagonalizable stochastic 3 × 3 matrix there exists a perturbation matrix \( E \) such that the conditions (a)-(d) are sufficed, we have to verify that after perturbation no negative elements or elements greater than 1 arise in the matrix \( \hat{A} \). This is shown in a working paper[10].

And since the eigenvalues of a matrix depend continuously on the matrix-elements[12], the eigenvalues of \( A \) and \( \hat{A} \) can be arbitrary close. Therefore, we can conclude that for every non-diagonalizable stochastic 3 × 3 matrix \( A \), there exists a perturbation matrix \( E \) such that the matrix \( A + E \) suffices the conditions (a)-(d).

### 3 SCEP-algorithm

The algorithm to determine a 3 × 3 perturbation matrix, which suffices conditions (a)-(d) is called the Stochastic Conserving Eigenspace Perturbation algorithm (SCEP-algorithm) and goes as follows:
Algorithm 1 SCEP-algorithm

Calculate a principal eigenvector $v$ of $A$  
Define $v_i$, ($i = 1, 2, 3$) as the components of $v$.  
Define $K := (a_{12} - a_{13} - a_{22} + a_{33})(v_1 + v_2) + 2(a_{13} - a_{23})v_2$  
Define $L := (a_{12} - a_{13} - a_{22} + a_{33})v_2 - 2(a_{12} - a_{32})(v_1 + v_2)$  
Define $M := (a_{12} - a_{13} - a_{22} + a_{33})v_3 + 2(a_{13} - a_{23})(v_1 + v_3)$  
Define $N := (a_{12} - a_{13} - a_{22} + a_{33})(v_1 + v_3) - 2(a_{12} - a_{32})v_3$

if $K \neq 0$ or $L \neq 0$ or $M \neq 0$ or $N \neq 0$ then

Find a solution $\tilde{c} = (\tilde{c}_{22}, \tilde{c}_{23}, \tilde{c}_{32}, \tilde{c}_{33})$ of $K\tilde{c}_{22} + L\tilde{c}_{23} + M\tilde{c}_{32} + N\tilde{c}_{33} > 0$

Define $E := \begin{pmatrix}
    v_2(\tilde{c}_{22} + \tilde{c}_{23}) + v_3(\tilde{c}_{32} + \tilde{c}_{33}) & -v_2\tilde{c}_{22} - v_3\tilde{c}_{32} & -v_2\tilde{c}_{23} - v_3\tilde{c}_{33} \\
    -\tilde{c}_{22} - \tilde{c}_{23} & \tilde{c}_{22} & \tilde{c}_{23} \\
    -\tilde{c}_{32} - \tilde{c}_{33} & \tilde{c}_{32} & \tilde{c}_{33}
\end{pmatrix}$

while $A + E \geq 0$ do

Find another solution $\tilde{c}' = (\tilde{c}'_{22}, \tilde{c}'_{23}, \tilde{c}'_{32}, \tilde{c}'_{33})$ of $K\tilde{c}'_{22} + L\tilde{c}'_{23} + M\tilde{c}'_{32} + N\tilde{c}'_{33} > 0$, linear independent of $\tilde{c}$

Define $E' := \begin{pmatrix}
    v_2(\tilde{c}'_{22} + \tilde{c}'_{23}) + v_3(\tilde{c}'_{32} + \tilde{c}'_{33}) & -v_2\tilde{c}'_{22} - v_3\tilde{c}'_{32} & -v_2\tilde{c}'_{23} - v_3\tilde{c}'_{33} \\
    -\tilde{c}'_{22} - \tilde{c}'_{23} & \tilde{c}'_{22} & \tilde{c}'_{23} \\
    -\tilde{c}'_{32} - \tilde{c}'_{33} & \tilde{c}'_{32} & \tilde{c}'_{33}
\end{pmatrix}$

end while

Define $\tilde{A} := A + E$

end if

return $\tilde{A}$
4 Further spectral properties

If we consider the set $\mathcal{DP}_{\delta}(A)$ of all matrices $\hat{A}$ (for a given $\delta$ in condition (c)) which suffice the conditions (a)-(d), then $\mathcal{DP}_{\delta}(A) \cup \{A\}$ is starconvex, according to the following theorem.

**Lemma 4.1.** If a perturbation matrix $E$ suffices the conditions (a)-(d) for a given non-diagonalizable matrix $A$, then does also $tE$, for all $t \in (0,1)$.

**Proof.** If we consider the form (2) of $E$, then conditions (a), (b) and (d) are immediately sufficed for $tE$. The multiplication with $t \in (0,1)$ preserves also the positivity of the directional derivatives.

$$D_{t\epsilon} \Delta(A) = K t \epsilon_{22} + L t \epsilon_{23} + M t \epsilon_{32} + N t \epsilon_{33} = t (K \epsilon_{22} + L \epsilon_{23} + M \epsilon_{32} + N \epsilon_{33}) > 0$$

$$D_{t\epsilon}^2 \Delta(A) = \left[t \epsilon_{22}(v_1 + v_2) + t \epsilon_{23}v_2 + t \epsilon_{32}v_3 + t \epsilon_{33}(v_1 + v_3) \right] + 4 (\epsilon_{32}\epsilon_{23} - \epsilon_{22}\epsilon_{33}) v_1 (v_1 + v_2 + v_3)$$

$$> 0$$

Therefore, we have that if $D_{t\epsilon} \Delta(A) > 0$, then $D_{t\epsilon} \Delta(A) > 0$ and if $D_{t\epsilon}^2 \Delta(A) > 0$, then $D_{t\epsilon}^2 \Delta(A) > 0$, both implications are valid for all $t \in (0,1)$. \[\square\]

The starconvexity of $\mathcal{DP}_{\delta}(A) \cup \{A\}$ is a necessary fact to be able to state the following theorem. In section 2, we already showed that every non-diagonalizable stochastic $3 \times 3$ matrix $A$ can be perturbed such that the conditions (a)-(d) are sufficed for $A+E$. Moreover, according to lemma 4.1, the matrices $\hat{A}(t) = A + tE$ suffice the condition (a)-(d), $\forall t \in (0,1)$.

**Theorem 4.1.** Suppose that $A$ is a non-diagonalizable stochastic $3 \times 3$-matrix. For all perturbation matrices $E$, resulting from the SCEP-algorithm hold the following statements:

(a) The eigenspaces of the non-principal eigenvalues of $\hat{A}(t) = A + tE$ converge \(^1\) to the eigenspace of the non-principal eigenvalue of $A$, as $t$ tends to zero (considering $E$ fixed as a result of the SCEP-algorithm and $t > 0$).

(b) The space spanned by the eigenvectors of the non-principal eigenvalues of $\hat{A}(t) = A + tE$ coincide with the space spanned by the eigenvector corresponding to the non-principal eigenvalue of $A$ and its generalized eigenvector. This statement is true for both left and right eigenvectors.

**Proof.** (a) The analytic matrix function $\hat{A}(t)$ has corresponding eigenvalues $\mu(t)$ and $\nu(t)$. The eigenvalues $\mu(t)$ and $\nu(t)$ can be expanded as Puiseux series \(^3\), being the branches of one Puiseux series. Both the eigenvalues $\mu(t)$ and $\nu(t)$ tend to $\lambda_0$ as $t$ tends to $0$. Since the eigenvalues $\mu(t)$ and $\nu(t)$ are branches of one Puiseux-series, we can choose \(^6\) corresponding normalized left eigenvectors $v_\mu(t) \in V_L(\mu(t))$ and $v_\nu(t) \in V_L(\nu(t))$ such that $v_\mu(0) = v_\nu(0)$. Now we have that $\mu(0) = \nu(0) = \lambda_0$, $A(0) = A$ and

$$\begin{cases} v_\mu(0)A(0) = \mu(0)v_\mu(0) \\ v_\nu(0)A(0) = \nu(0)v_\nu(0) \end{cases}$$

Since $\lambda$ has geometric multiplicity 1, it follows that an eigenvector $v$ corresponding the eigenvalue $\lambda$ is a multiple of $v_\mu(0)$, since both $v$ and $v_\mu(0)$ are eigenvectors corresponding to the eigenvalue $\lambda$ of $A = A(0)$. The same goes for $v_\nu(0)$. Thus:

$$\text{Span}[v_\mu(0)] = \text{Span}[v_\nu(0)] = \text{Span}[v]$$

which proves part (a) of the theorem.

(b) It is known that the eigenvectors $v_\mu(t)$ and $v_\nu(t)$ which correspond to different eigenvalues, are linear independent. Since $\text{Ker}(A - \lambda I) \subset \text{Ker}(A - \lambda I)^2$, the remainder to show is:

$$v_\mu(t) (A - \lambda I)^2 = 0 \quad \text{and} \quad v_\nu(t) (A - \lambda I)^2 = 0$$

\(^1\)Convergence of one-dimensional eigenspaces is defined here as follows:

The eigenspace is spanned by one eigenvector $v(t)$, depending on a continuous parameter $t$. As this parameter $t$ approaches a certain value $a$, the corresponding normalized vector $v(t)$ might converge to a certain vector $v$. If this limit for $v(t)$ as $t \to a$ converges to $v$, the eigenspace $V_L(v(t))$ converges to the space spanned by to the limit vector $v$. 

\(^3\)The eigenvalue is a Puiseux series.

\(^6\)Normalized means that the eigenvectors are normalized such that the corresponding eigenvalues are 1.
Then we have two linear independent vectors, \( v_\mu(t) \) and \( v_\nu(t) \), in the two dimensional kernel \( \text{Ker}[(A - \lambda I)^2] \), thus those form a basis.

Therefore we consider first the matrix \( (A - \lambda I)^2 \):

\[
(A - \lambda I)^2 = (TJT^{-1} - \lambda I)^2 = [T(J - \lambda I)T^{-1}]^2 = T(J - \lambda I)^2 T^{-1} = (1 - \lambda)^2 w u
\]

where \( TJT^{-1} \) is the corresponding Jordan decomposition of \( A \), \( J \) being the corresponding Jordan-matrix, \( w \in V_R(1) \) and \( u \in V_L(1) \). Since \( \tilde{A}(t) \) is also a stochastic matrix, its eigenvector corresponding to 1, \( w(t) \) is also a multiple of \( (1, 1, 1)^T \) and thus \( w = cw(t) \), for some \( c \in \mathbb{R} \). We have:

\[
v_\mu(t) (A - \lambda I)^2 = (1 - \lambda)^2 v_\mu(t) w u = c(1 - \lambda)^2 v_\mu(t) w(t) u = 0
\]

since \( v_\mu(t) w(t) = 0 \), which follows from:

\[
v_\mu(t) w(t) = v_\mu(t) \left[ \tilde{A}(t) w(t) \right] = \left[ v_\mu(t) \tilde{A}(t) \right] w(t) = \mu v_\mu(t) w(t)
\]

\[
\Rightarrow (1 - \mu)v_\mu(t) w(t) = 0
\]

\[
\Rightarrow v_\mu(t) w(t) = 0
\]

The proof for the right eigenvectors is completely analogue.

The presented procedure has many advantages. If we would compare this well-chosen perturbation, found by using the SCEP-algorithm, to a random generated perturbation matrix, this well-chosen perturbation preserves spectral properties from the original matrix \( A \) onto the perturbed matrix \( \tilde{A} \). Therefore the behaviour of the perturbed matrix \( \tilde{A} \) resembles the behaviour of the matrix \( A \). Another advantage, via this procedure, we know that for every non-diagonalizable stochastic 3 \( \times \) 3 matrix \( A \), there exists a sequence of diagonalizable stochastic 3 \( \times \) 3 matrices which all have coinciding principal eigenspaces and this sequence has \( A \) as its limit. This is helpful for proving matrix identities for non-diagonalizable matrices.

References


6