## Vrije Universiteit Brussel

# On perturbations of non-diagonalizable stochastic matrices of order 3 

Pauwelyn, Pieter-Jan; Guerry, Marie-Anne

Published in:
Statistics and Probability Letters

DOI:
10.1016/j.spl.2019.108633

Publication date:
2020

License:
CC BY-NC-ND

Document Version:
Accepted author manuscript

Link to publication

Citation for published version (APA):
Pauwelyn, P-J., \& Guerry, M-A. (2020). On perturbations of non-diagonalizable stochastic matrices of order 3.
Statistics and Probability Letters, 157, [108633]. https://doi.org/10.1016/j.spl.2019.108633

## Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

## Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

# On perturbations of non-diagonalizable stochastic matrices of order 3 

P.J. PAUWELYN ${ }^{\text {a }}$ and M.A. GUERRY ${ }^{\text {b }}$


#### 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 [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) $\operatorname{spec}(\tilde{A})=\{1, \mu, \nu\} \subset \mathbb{R}, \mu \neq \nu$ and $\forall \delta>0, \exists E \in \mathbb{R}^{3 \times 3}:|\mu-\lambda|<\delta \wedge|\nu-\lambda|<\delta$.
(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 $A$, 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 matrixelements $\epsilon_{i j}$ 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 $A \overrightarrow{1}=\tilde{A} \overrightarrow{1}$, (where $\overrightarrow{1} \in \mathbb{R}^{3}$ is the columnvector with all components equal to 1) must be fullfilled. This equality implies

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

[^0]Therefore the rowsums of the perturbation matrix $E$ should be zero. Otherwise said:

$$
\begin{equation*}
\forall i \in\{1,2,3\}: \sum_{j=1}^{3} \epsilon_{i j}=0 \tag{1}
\end{equation*}
$$

Implementing condition (d) goes as follows. The eigenspaces corresponding to the eigenvalue 1 , are the left eigenspace $V_{L}(1)=\left\{v \in \mathbb{R}^{1 \times 3} \mid v\left(A-I_{3}\right)=0\right\}$ and the right eigenspace $V_{R}(1)=\left\{w \in \mathbb{R}^{3 \times 1} \mid\left(A-I_{3}\right) w=0\right\}$, with $I_{3}$ the identity matrix of order 3 . From here on we use the notation $V_{L}(\lambda)$ for the left eigenspace corresponding to the eigenvalue $\lambda$ of the matrix $A$ and $V_{R}(\lambda)$ for the right eigenspace 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 $v\left(\tilde{A}-I_{3}\right)=0$ and $\left(\tilde{A}-I_{3}\right) w=0$ 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 $V_{R}(1)=\tilde{V}_{R}(1)=$ $\left\{(k, k, k)^{T} \mid k \in \mathbb{R}\right\}$ and coincide. To obtain coinciding left eigenspaces $V_{L}(1)$ and $\tilde{V}_{L}(1)$, the perturbation matrix $E$ should suffice the equality $v\left(A+E-I_{3}\right)=0$, 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 $u E=0$.

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

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

Vice versa, if $u E=0$, then:

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

With this lemma, condition (d), the preservation of the eigenspaces corresponding to the eigenvalue 1 becomes $v_{L} E=0$ for $v_{L} \in V_{L}(1)$. If $v_{L} E=0$ for some $v_{L} \in V_{L}(1)$, then $v E=0$ for all $v \in V_{L}(1)$, since $V_{L}(1)$ is one-dimensional. This constraint $v E=0$ combined with the constraint $E \overrightarrow{1}=0$ implies the following form of $E$ :

$$
E=\left(\begin{array}{ccc}
v_{2}\left(\epsilon_{22}+\epsilon_{23}\right)+v_{3}\left(\epsilon_{32}+\epsilon_{33}\right) & -v_{2} \epsilon_{22}-v_{3} \epsilon_{32} & -v_{2} \epsilon_{23}-v_{3} \epsilon_{33}  \tag{2}\\
-v_{1}\left(\epsilon_{22}+\epsilon_{23}\right) & v_{1} \epsilon_{22} & v_{1} \epsilon_{23} \\
-v_{1}\left(\epsilon_{32}+\epsilon_{33}\right) & v_{1} \epsilon_{32} & v_{1} \epsilon_{33}
\end{array}\right)
$$

with $v=\left(v_{1}, v_{2}, v_{3}\right)$ 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:

$$
\begin{aligned}
& -\lambda^{3}+\operatorname{tr}(\tilde{A}) \lambda^{2}+[1-\operatorname{tr}(\tilde{A})-\operatorname{det}(\tilde{A})] \lambda+\operatorname{det}(\tilde{A})=0 \\
& \Leftrightarrow(\lambda-1)\left[-\lambda^{2}+(\operatorname{tr}(\tilde{A})-1) \lambda-\operatorname{det}(\tilde{A})\right]=0
\end{aligned}
$$

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

$$
\begin{equation*}
\Delta(\tilde{A})=[\operatorname{tr}(\tilde{A})-1]^{2}-4 \operatorname{det}(\tilde{A})>0 \tag{3}
\end{equation*}
$$

We consider the discriminant of a stochastic matrix as a function of its matrixelements. Therefore we define the set $U=\left\{\left(b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}\right) \mid \forall i, j: b_{i j} \geq 0 \wedge b_{i 2}+b_{i 3} \leq 1\right\}$, which is isomorph 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 $\mathbb{R}^{6}$.

$$
\Delta: U \rightarrow \mathbb{R}:\left(b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}\right) \mapsto \Delta(B)=\Delta\left(\begin{array}{ccc}
1-b_{12}-b_{13} & b_{12} & b_{13} \\
1-b_{22}-b_{23} & b_{22} & b_{23} \\
1-b_{32}-b_{33} & b_{32} & b_{33}
\end{array}\right)
$$

Next, the discriminant function can be written in fully as:

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

According to (3), the discriminant of $\tilde{A}$ must be positive. To determine what happens with the discriminant, when $A$ is perturbed into $\tilde{A}=A+E$, 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 $\vec{\epsilon}=\left(\epsilon_{12}, \epsilon_{13}, \epsilon_{22}, \epsilon_{23}, \epsilon_{32}, \epsilon_{33}\right) \in \mathbb{R}^{6}$. The directional derivative[1] of a non-diagonalizable stochastic $3 \times 3$ matrix $B$ in a direction $\vec{\epsilon}, D_{\vec{\epsilon}} \Delta(B)$ is

$$
\begin{aligned}
D_{\vec{\epsilon}} \Delta(B)= & D_{\vec{\epsilon}}\left(b_{12}, b_{13}, b_{22}, b_{23}, b_{32}, b_{33}\right)=\frac{\vec{\epsilon}}{\|\vec{\epsilon}\|} \bullet\left(\frac{\partial \Delta}{\partial b_{12}}, \frac{\partial \Delta}{\partial b_{13}}, \frac{\partial \Delta}{\partial b_{22}}, \frac{\partial \Delta}{\partial b_{23}}, \frac{\partial \Delta}{\partial b_{32}}, \frac{\partial \Delta}{\partial b_{33}}\right) \\
= & \frac{2}{\|\vec{\epsilon}\|}\left[\left(b_{12}-b_{13}-b_{22}+b_{33}\right)\left(\epsilon_{12}-\epsilon_{13}-\epsilon_{22}+\epsilon_{33}\right)+2\left(b_{13}-b_{23}\right)\left(\epsilon_{12}-\epsilon_{32}\right)\right. \\
& \left.+2\left(b_{12}-b_{32}\right)\left(\epsilon_{13}-\epsilon_{23}\right)\right]
\end{aligned}
$$

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:

$$
\begin{aligned}
D_{\vec{\epsilon}} \Delta(A)= & \left(a_{12}-a_{13}-a_{22}+a_{33}\right)\left(-v_{2} \epsilon_{22}-v_{3} \epsilon_{32}-v_{2} \epsilon_{23}-v_{3} \epsilon_{33}-v_{1} \epsilon_{22}+v_{1} \epsilon_{33}\right) \\
& \quad+2\left(a_{13}-a_{23}\right)\left(-v_{2} \epsilon_{22}-v_{3} \epsilon_{32}-v_{1} \epsilon_{32}\right)+2\left(a_{12}-a_{32}\right)\left(-v_{2} \epsilon_{23}-v_{3} \epsilon_{33}-v_{1} \epsilon_{23}\right) \\
= & K \epsilon_{22}+L \epsilon_{23}+M \epsilon_{32}+N \epsilon_{33}
\end{aligned}
$$

with

$$
\begin{aligned}
K & =\left(a_{12}-a_{13}-a_{22}+a_{33}\right)\left(v_{1}+v_{2}\right)+2\left(a_{13}-a_{23}\right) v_{2} \in \mathbb{R} \\
L & =\left(a_{12}-a_{13}-a_{22}+a_{33}\right) v_{2}-2\left(a_{12}-a_{32}\right)\left(v_{1}+v_{2}\right) \in \mathbb{R} \\
M & =\left(a_{12}-a_{13}-a_{22}+a_{33}\right) v_{3}+2\left(a_{13}-a_{23}\right)\left(v_{1}+v_{3}\right) \in \mathbb{R} \\
N & =\left(a_{12}-a_{13}-a_{22}+a_{33}\right)\left(v_{1}+v_{3}\right)-2\left(a_{12}-a_{32}\right) v_{3} \in \mathbb{R}
\end{aligned}
$$

From here, we consider the direction $\vec{\epsilon}$ to be the vector $\left(\epsilon_{22}, \epsilon_{23}, \epsilon_{32}, \epsilon_{33}\right)$.
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 $D_{\vec{\epsilon}}^{2} \Delta(A)$ 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 $D_{\vec{\epsilon}}^{2} \Delta(A)$ is positive, then $\Delta(A)$ will increase in the direction of $\vec{\epsilon}$. The inequality for $D_{\vec{\epsilon}}^{2} \Delta(A)$ is:

$$
D_{\vec{\epsilon}}^{2} \Delta(A)=\left(\epsilon_{22}\left(v_{1}+v_{2}\right)+\epsilon_{23} v_{2}+\epsilon_{32} v_{3}+\epsilon_{33}\left(v_{1}+v_{3}\right)\right)^{2}+4\left(\epsilon_{32} \epsilon_{23}-\epsilon_{22} \epsilon_{33}\right) v_{1}\left(v_{1}+v_{2}+v_{3}\right)>0
$$

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 $3 \times 3$ 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 $\tilde{A}$. 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 $\tilde{A}$ can be arbitrary close. Therefore, we can conclude that for every non-diagonalizable stochastic $3 \times 3$ 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 $3 \times 3$ 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 \(K:=\left(a_{12}-a_{13}-a_{22}+a_{33}\right)\left(v_{1}+v_{2}\right)+2\left(a_{13}-a_{23}\right) v_{2}\)
    Define \(L:=\left(a_{12}-a_{13}-a_{22}+a_{33}\right) v_{2}-2\left(a_{12}-a_{32}\right)\left(v_{1}+v_{2}\right)\)
    Define \(M:=\left(a_{12}-a_{13}-a_{22}+a_{33}\right) v_{3}+2\left(a_{13}-a_{23}\right)\left(v_{1}+v_{3}\right)\)
    Define \(N:=\left(a_{12}-a_{13}-a_{22}+a_{33}\right)\left(v_{1}+v_{3}\right)-2\left(a_{12}-a_{32}\right) v_{3}\)
    if \(K \neq 0\) or \(L \neq 0\) or \(M \neq 0\) or \(N \neq 0\) then
        Find a solution \(\overrightarrow{\epsilon^{*}}=\left(\epsilon_{22}^{*}, \epsilon_{23}^{*}, \epsilon_{32}^{*}, \epsilon_{33}^{*}\right)\) of \(K \epsilon_{22}+L \epsilon_{23}+M \epsilon_{32}+N \epsilon_{33}>0\)
        Define \(E:=\left(\begin{array}{ccc}v_{2}\left(\epsilon_{22}^{*}+\epsilon_{23}^{*}\right)+v_{3}\left(\epsilon_{32}^{*}+\epsilon_{33}^{*}\right) & -v_{2} \epsilon_{22}^{*}-v_{3} \epsilon_{32}^{*} & -v_{2} \epsilon_{23}^{*}-v_{3} \epsilon_{33}^{*} \\ -\epsilon_{22}^{*}-\epsilon_{23}^{*} & \epsilon_{22}^{*} & \epsilon_{23}^{*} \\ -\epsilon_{32}^{*}-\epsilon_{33}^{*} & \epsilon_{32}^{*} & \epsilon_{33}^{*}\end{array}\right)\)
        while \(A+E \nsupseteq 0\) do
            Find another solution \(\overrightarrow{\epsilon^{\prime}}=\left(\epsilon_{22}^{\prime}, \epsilon_{23}^{\prime}, \epsilon_{32}^{\prime}, \epsilon_{33}^{\prime}\right)\) of \(K \epsilon_{22}+L \epsilon_{23}+M \epsilon_{32}+N \epsilon_{33}>0\), linear independent of
            \(\epsilon^{*}\)
            Define \(E:=\left(\begin{array}{ccc}v_{2}\left(\epsilon_{22}^{\prime}+\epsilon_{23}^{\prime}\right)+v_{3}\left(\epsilon_{32}^{\prime}+\epsilon_{33}^{\prime}\right) & -v_{2} \epsilon_{22}^{\prime}-v_{3} \epsilon_{32}^{\prime} & -v_{2} \epsilon_{23}^{\prime}-v_{3} \epsilon_{33}^{\prime} \\ -\epsilon_{22}^{\prime}-\epsilon_{23}^{\prime} & \epsilon_{22}^{\prime} & \epsilon_{23}^{\prime} \\ -\epsilon_{32}^{\prime}-\epsilon_{33}^{\prime} & \epsilon_{32}^{\prime} & \epsilon_{33}^{\prime}\end{array}\right)\)
```


## end while

```
Define \(\tilde{A}:=A+E\)
else
Find a solution \(\overrightarrow{\epsilon^{*}}=\left(\epsilon_{22}^{*}, \epsilon_{23}^{*}, \epsilon_{32}^{*}, \epsilon_{33}^{*}\right)\) for \(\left(\epsilon_{22}\left(v_{1}+v_{2}\right)+\epsilon_{23} v_{2}+\epsilon_{32} v_{3}+\epsilon_{33}\left(v_{1}+v_{3}\right)\right)^{2}+\) \(4\left(\epsilon_{32} \epsilon_{23}-\epsilon_{22} \epsilon_{33}\right) v_{1}\left(v_{1}+v_{2}+v_{3}\right)>0\)
Define \(E:=\left(\begin{array}{ccc}v_{2}\left(\epsilon_{22}^{*}+\epsilon_{23}^{*}\right)+v_{3}\left(\epsilon_{32}^{*}+\epsilon_{33}^{*}\right) & -v_{2} \epsilon_{22}^{*}-v_{3} \epsilon_{32}^{*} & -v_{2} \epsilon_{23}^{*}-v_{3} \epsilon_{33}^{*} \\ -\epsilon_{22}^{*}-\epsilon_{23}^{*} & \epsilon_{22}^{*} & \epsilon_{23}^{*} \\ -\epsilon_{32}^{*}-\epsilon_{33}^{*} & \epsilon_{32}^{*} & \epsilon_{33}^{*}\end{array}\right)\)
while \(A+E \nsupseteq 0\) do
Find another solution \(\overrightarrow{\epsilon^{\prime}}=\left(\epsilon_{22}^{\prime}, \epsilon_{23}^{\prime}, \epsilon_{32}^{\prime}, \epsilon_{33}^{\prime}\right)\) of \(K \epsilon_{22}+L \epsilon_{23}+M \epsilon_{32}+N \epsilon_{33}>0\), linear independent of \(\overrightarrow{\epsilon^{*}}\)
Define \(E:=\left(\begin{array}{ccc}v_{2}\left(\epsilon_{22}^{\prime}+\epsilon_{23}^{\prime}\right)+v_{3}\left(\epsilon_{32}^{\prime}+\epsilon_{33}^{\prime}\right) & -v_{2} \epsilon_{22}^{\prime}-v_{3} \epsilon_{32}^{\prime} & -v_{2} \epsilon_{23}^{\prime}-v_{3} \epsilon_{33}^{\prime} \\ -\epsilon_{22}^{\prime}-\epsilon_{23}^{\prime} & \epsilon_{22}^{\prime} & \epsilon_{23}^{\prime} \\ -\epsilon_{32}^{\prime}-\epsilon_{33}^{\prime} & \epsilon_{32}^{\prime} & \epsilon_{33}^{\prime}\end{array}\right)\)
```


## end while

```
Define \(\tilde{A}:=A+E\)
end if
return \(\tilde{A}\)
```


## 4 Further spectral properties

If we consider the set $\mathcal{D} \mathcal{P}_{\delta}(A)$ of all matrices $\tilde{A}$ (for a given $\delta$ in condition (c)) which suffice the conditions (a)-(d), then $\mathcal{D P}_{\delta}(A) \cup\{A\}$ 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 $t E$, 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 $t E$. The multiplication with $t \in(0,1)$ preserves also the positivity of the directional derivatives.

$$
\left.\left.\begin{array}{l}
D_{t \vec{\epsilon}} \Delta(A)=K t \epsilon_{22}+L t \epsilon_{23}+M t \epsilon_{32}+N t \epsilon_{33}=t\left(K \epsilon_{22}+L \epsilon_{23}+M \epsilon_{32}+N \epsilon_{33}\right)>0 \\
D_{t \epsilon}^{2} \Delta(A)
\end{array}\right)\left[t \epsilon_{22}\left(v_{1}+v_{2}\right)+t \epsilon_{23} v_{2}+t \epsilon_{32} v_{3}+t \epsilon_{33}\left(v_{1}+v_{3}\right)\right]^{2}+4\left(t \epsilon_{32} t \epsilon_{23}-t \epsilon_{22} t \epsilon_{33}\right) v_{1}\left(v_{1}+v_{2}+v_{3}\right)\right] .
$$

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

The starconvexity of $\mathcal{D} \mathcal{P}_{\delta}(A) \cup\{A\}$ 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 $\tilde{A}(t)=A+t E$ 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 $\tilde{A}(t)=A+t E$ converge ${ }^{1}$ to the eigenspace of the non-principal eigenvalue of $A$, as 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 $\tilde{A}(t)=A+t E$ coincide with the space spanned by the eigenvector corresponding to the non-principal eigenvalue of $A$ and it's generalized eigenvector. This statement is true for both left and right eigenvectors.
Proof. (a) The analytic matrix function $\tilde{A}(t)$ has corresponding eigenvalues $1, \mu(t)$ 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 $\lambda$ 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 $\mu(0)=\nu(0)=\lambda, A(0)=A$ and

$$
\left\{\begin{array}{l}
v_{\mu}(0) A(0)=\mu(0) v_{\mu}(0) \\
v_{\nu}(0) A(0)=\nu(0) v_{\nu}(0)
\end{array}\right.
$$

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:

$$
\operatorname{Span}\left[v_{\mu}(0)\right]=\operatorname{Span}\left[v_{\nu}(0)\right]=\operatorname{Span}[v]
$$

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 $\operatorname{Ker}(A-\lambda I) \subset \operatorname{Ker}(A-\lambda I)^{2}$, 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]Then we have two linear independent vectors, $v_{\mu}(t)$ and $v_{\nu}(t)$, in the two dimensional kernel $\operatorname{Ker}\left[(A-\lambda I)^{2}\right]$, thus those form a basis.
Therefore we consider first the matrix $(A-\lambda I)^{2}$ :

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

where $T J T^{-1}$ is the corresponding Jordan decomposition of $A, J$ being the corresponding Jordanmatrix, $w \in V_{R}(1)$ 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 $w=c w(t)$, 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:

$$
\begin{aligned}
v_{\mu}(t) w(t) & =v_{\mu}(t)[\tilde{A}(t) w(t)]=\left[v_{\mu}(t) \tilde{A}(t)\right] w(t)=\mu v_{\mu}(t) w(t) \\
& \Rightarrow(1-\mu) v_{\mu}(t) w(t)=0 \\
& \Rightarrow v_{\mu}(t) w(t)=0
\end{aligned}
$$

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

[1] Adams, R.A. (2006) Calculus: A complete course (6th ed.), Pearson Education Canada
[2] Baumgärtel, H. (1985) Analytic perturbation theory for matrix and operators, Birkhäuser, Basel
[3] Casas-Alvero, E. (2000) Singularities of plane curves, London Math. Soc. Lecture Notes 276, Cambridge University Press
[4] Chaitin-Chatelin, F., Harrabi, A. \& Ilahi, A. (2000) About Hölder condition numbers and the stratification diagram for defective eigenvalues, Mathematics and Computers in Simulation 54, 397-402, Elsevier
[5] Conlisk, J. (1985) Comparative statics for markov chains, Journal of Economic Dynamics and Control 9, 139-151
[6] Hryniv, R., Lancaster, P. (1999) On the perturbation of analytic matrix functions, Integral Equations and Operator Theory 34, 325-338, Birkhäuser, Basel
[7] Kato, T. (1985) Perturbation theory for linear Operators, Springer-Verlag, Berlin-Heidelberg-New York
[8] Meyer, C.D. and Stewart, G.W. (1988) Derivatives and perturbations of eigenvectors, SIAM Journal on Numerical Analysis 25(3)
[9] Minc, H. (1988) Nonnegative matrices, Wiley Interscience Series in Discrete Mathematics and Optimization, John Wiley \& Sons
[10] Pauwelyn, P., Guerry, M.A.(2019) Working paper: Preservations of stochasticity after perturbations, ES working paper (or Researchgate)
[11] Rellich, F. (1968) Perturbation theory of eigenvalue problems, Gordon and Breach Science Publishers, New York-London-Paris
[12] Stewart, G.W. \& Sun, J.-G. (1990) Matrix perturbation theory, Academic Press, New York
[13] Stewart, G. W. (2001) Matrix algorithms volume 2: Eigensystems, Society for Industrial and Applied Mathematics


[^0]:    *Pieter-Jan.Pauwelyn@vub.be, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
    ${ }^{\dagger}$ Marie-Anne.Guerry@vub.be, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium

[^1]:    ${ }^{1}$ 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 $t \rightarrow a$ converges to $v$, the eigenspace $V_{L}(v(t))$ converges to the space spanned by to the limit vector $v$.

