Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 418 (2006) 682–692
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Correlation matrices of yields and total positivity

Ernesto Salinelli [a], Carlo Sgarra [b,*]

[a] *Dipartimento di Scienze Economiche e Metodi Quantitativi, Università del Piemonte Orientale "A. Avogadro", Via Perrone 18, 28100 Novara, Italy*
[b] *Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

Received 17 November 2005; accepted 5 March 2006
Available online 5 May 2006
Submitted by S. Fallat

Abstract

It has been empirically observed that correlation matrices of forward interest rates have the first three eigenvalues which are simple and their corresponding eigenvectors, termed as shift, slope and curvature respectively, with elements presenting changes of sign in a regular way. These spectral properties are very similar to those exhibited by Strictly Totally Positive and Oscillatory matrices. In the present paper we investigate how these spectral properties are related with those characterizing the correlation matrices considered, i.e. the positivity and the monotonicity of their elements. On the basis of these relations we prove the simplicity of the first two eigenvalues and provide an estimate of the second one.
© 2006 Elsevier Inc. All rights reserved.

AMS classification: 15A48; 62H25; 91B28

Keywords: Forward rates; Correlation matrices; Principal component analysis; Total positivity

1. Introduction

Since the middle of the eighties an intensive effort has been devoted to develop statistical methods in order to describe the movements of the yield curve and to empirically justify some important models based on one or more factors (see e.g. [8,9,20,21]). Principal component analysis (PCA from now on) has turned out to be one of the most useful tools to find such factors.

* Corresponding author. Tel.: +39 02 2399 4546; fax: +39 02 2399 4621.
E-mail addresses: ernesto.salinelli@eco.unipmn.it (E. Salinelli), carlo.sgarra@mate.polimi.it (C. Sgarra).

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.03.021

Consider a yield curve completely described by an n-dimensional vector of forward rates corresponding to maturities $t_1 < t_2 < \cdots < t_n$. These forward rates can be assumed (see [13]) as the components of a random vector (r.v.) $\mathbf{X}^{\mathrm{T}} = [X_1, X_2, \ldots, X_n]$ with zero mean $\mathbb{E}[\mathbf{X}]$ and unit variances $\mathbb{E}[X_i^2]$. This vector can be expressed as a linear transformation $V\mathbf{Y}$ of a n-dimensional r.v. $\mathbf{Y}$ with uncorrelated elements (factors) $Y_j = \mathbf{V}_j^{\mathrm{T}}\mathbf{X}$ termed as (see [12]) the *principal components* (PCs) of $\mathbf{X}$. The vectors $\mathbf{V}_j$, for $j = 1, 2, \ldots, n$, are the normalized eigenvectors of the correlation matrix $R = \mathbb{E}[\mathbf{X}'\mathbf{X}]$ of $\mathbf{X}$ and the corresponding eigenvalues λ_j are the variances of the PCs, taken in decreasing order: $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$. The elements V_{ij} of each vector $\mathbf{V}_j$ represent the contribution of the ith original variable to the jth PC.

We want to mention here that PCA has been extensively used for applications both to fixed income derivatives pricing and to risk management (as in [15,22]); several indices have been introduced in literature to measure the risk related to yield curve dynamics, and several different approaches have been presented to manage exposure of portfolios to such risk, but many of them have been proved to be related to PCA in some way [7].

An interesting feature of the correlation matrices for forward rates is that they systematically exhibit some structural properties which have a very simple financial explanation:

(a) interest rates at different maturities are always positively correlated;
(b) the correlation coefficients decrease when the distance between the indices increases: this is a far obvious consequence of the decreasing degree of correlation when the variables are more distant in time;
(c) the previous reduction in the correlation between variables corresponding to the same difference in the indices tends to decrease as the maturities of both the variables are greater.

It is important to stress at this point that in the PCA for yield curve models, given the high degree of correlation between the variables involved, a small number of PCs can explain a large part of the variability, i.e. if n is the dimension of $\mathbf{X}$ then

$$\sum_{s=1}^{k} \lambda_s \simeq \operatorname{tr}(R) = n \quad \text{for } k \ll n.$$

More precisely, the first three PCs are usually considered to be sufficient to account for the total variability. Moreover, it has been empirically observed (see [20]) that the first eigenvector $\mathbf{V}_1$ has approximately equal components, hence it is interpreted as the average level of the yield curve and called the **shift** (or **level**); the second eigenvector $\mathbf{V}_2$ has elements approximately equal in magnitude and opposite signs at the extremes of the maturity range and for this reason it is named the **slope** of the yield curve; the third eigenvector $\mathbf{V}_3$ has components approximately equal at the extremes of the maturity range and twice large and of the opposite sign in the middle, hence it is generally called the **curvature**. It has been observed also that the eigenvalues of the correlation matrices appearing in this context are always simple.

In this paper we study how the empirically observed results afore mentioned are related with the structure of the correlation matrix of the data, by using some tools originally developed in the framework of totally positive matrices (TP). TP matrices have been thoroughly investigated in the past: we just mention the books by Gantmacher [3], Gantmacher and Krein [4], Karlin [10] and the more recent review by Ando [1]. In particular, in [4] the applications to continuum mechanics and the relations with oscillating systems have been systematically studied. A more detailed study of the financial implications of the present approach will be the subject of a forthcoming paper.

The structure of the exposition is the following: in Section 2 we present the main hypotheses, based on empirical evidence, on which our results will be proved. In Section 3 we recall the basic properties of TP matrices and illustrate the main relations between them and the correlation matrices of our interest. In Section 4 we study in detail the properties of the second eigenvalue of R, in particular its simplicity, and give an estimate for this eigenvalue on the basis of some other, more recent results, obtained in the framework of total positivity. We present in Section 5 some future perspectives and possible developments of the present work.

2. Notation and basic assumptions

An n-dimensional correlation matrix $R = [\rho_{ij}]$ is a symmetric semidefinite positive matrix with $\rho_{ij} \in [-1, 1]$ and $\rho_{ii} = 1$ for all i. Based on the empirical evidence, we will consider correlation matrices with the following properties:

P0 **positivity**, i.e. $\rho_{ij} > 0, \forall i, j = 1, 2, \ldots, n$;
P1 **strictly decreasing subdiagonal column elements**:

$$1 \leqslant j < s < i \leqslant n \quad \text{imply } \rho_{ij} < \rho_{sj};$$

P2 **strictly increasing superdiagonal column elements**:

$$1 \leqslant s < i < j \leqslant n \quad \text{imply } \rho_{sj} < \rho_{ij}.$$

Properties P0–P2 express the positivity and the monotonic behavior of correlations between forward rates with respect to difference in maturities we just mentioned in the introduction as points (a) and (b). Note that, by the symmetry of R, properties P1 and P2 can be obviously stated in terms of rows. A suitable formalization of point (c) will be introduced later in Section 3.

Example 1. In [20] it is shown that the correlation matrix R of yields can be approximately described by the following correlation function of the maturities $t_s > 0$:

$$\rho_{ij} = \exp\{-\beta|t_j - t_i|\} \quad \beta > 0. \tag{1}$$

Since $\{t_s\}$ is an increasing sequence, the correlation matrix $R = [\rho_{ij}]$ satisfies P0–P2.

Remark 2. If $R = [\rho_{ij}]$ is an n-dimensional correlation matrix R satisfying properties P0–P2 then $\rho_{ij} < 1$ for all $i \neq j$, and $\min_{ij}\{\rho_{ij}\} = \rho_{n1}$.

Having formalized the properties of the empirical correlation matrices of forward rates, now we characterize the spectral properties of our interest.

Definition 3. Given a correlation matrix R having the first three eigenvalues simple, we define:

- its *first eigenvector*: **pure shift** if it is $\mathbf{1} = [1, 1, \ldots, 1]^{\mathrm{T}}$; **shift** if all its components are positive, increasing and then decreasing; **weak shift** if all its components are positive;
- its *second eigenvector*: **slope** if its elements are increasing (decreasing) presenting a unique change of sign; **weak slope** if its elements present a unique change of sign;
- its *third eigenvector*: **curvature** if its elements are decreasing (increasing) and then increasing (decreasing) presenting only two changes of sign; **weak curvature** if its elements present only two changes of sign.

It is possible to find examples of correlation matrices with all the first three eigenvectors satisfying at least one of the requirements given in the afore mentioned definition.

Remark 4. It is easy to verify that the dominant eigenvector $\mathbf{V}_1$ of an n-dimensional positive correlation matrix R is pure shift if and only if each sum r_i of the ith row elements of R is equal to the dominant eigenvalue λ_1.

The distinction among shift, slope and curvature (eventually in a weak sense) depends on the number of sign variations in their components. So, it is necessary to clarify the role of zero elements. More precisely, all zero components can be considered both positive or negative, but according to the sign attributed to each one of them, for every eigenvector a maximum and minimum number of sign variations can be defined.

Definition 5 (*see* [3]). Given a vector $\mathbf{V} \in \mathbb{R}^n$ and denoting by $S_{\mathbf{V}}^+$ and $S_{\mathbf{V}}^-$ respectively the maximum and minimum number of sign variations for its components, if $S_{\mathbf{V}}^+ = S_{\mathbf{V}}^-$ we define this common value the **number of sign variations** of $\mathbf{V}$.

It is important to remark that the minimum and maximum number of sign variations coincide only if the first and the last components of a vector are not zero and if for every zero component the preceding and the following components are of different sign.

Example 6 (*Perfect correlation*). Given a correlation matrix R of dimension n, $n \geqslant 2$, we have $\rho_{ij} = 1$ for all i, j, if and only if

$$\lambda_1 = n, \quad \lambda_k = 0 \quad \text{for } k = 2, 3, \ldots, n.$$

In this case the dominant eigenvector is pure shift.

Example 7 (*Equicorrelation*). If $n \geqslant 2$ and $\rho_{ij} = \rho \in (0, 1)$ for $i \neq j$, then (see e.g. [11,12]) R admits the simple dominant eigenvalue $\lambda_1 = 1 + (n-1)\rho$ with associated eigenvector which is pure shift. The other $n-1$ eigenvalues $\lambda_k = 1 - \rho$, for $k = 2, \ldots, n$, have corresponding eigenvectors

$$\mathbf{V}_k = \Bigg[\underbrace{\frac{1}{\sqrt{k(k-1)}}, \ldots, \frac{1}{\sqrt{k(k-1)}}}_{k-1 \text{ times}}, -\frac{k-1}{\sqrt{k(k-1)}}, 0, \ldots, 0\Bigg]^{\mathrm{T}}.$$

In particular, we can conclude that any 2-dimensional correlation matrix (with $\rho \in (0, 1)$) has dominant eigenvector pure shift and the second eigenvector slope.

Example 8. If $n = 3$, it easy to show that a positive correlation matrix R has pure shift dominant eigenvector if and only if $\rho_{ij} = \rho$ for all $i \neq j$ (see Example 7). In this case the second and third eigenvectors are:

$$\mathbf{V}_2 = [-1, 0, 1]^{\mathrm{T}}, \qquad \mathbf{V}_3 = [-1, 1, 0]^{\mathrm{T}}.$$

According to Definition 5, $\mathbf{V}_2$ is slope, while $\mathbf{V}_3$ is not of proper curvature kind since $S_{\mathbf{V}_3}^+ = 2$ but $S_{\mathbf{V}_3}^- = 1$. We conclude that it does not exist *a three-dimensional positive correlation matrix with eigenvectors that are pure shift, slope and curvature*.

Every positive correlation matrix has weak shift dominant eigenvector: this is a part of the content of the celebrated Frobenius–Perron Theorem (see e.g. [3]).

Theorem 9. *If A is an $n \times n$ positive matrix, then there exists a positive real number λ_1 and a positive vector $\mathbf{V}_1 \in \mathbb{R}^n$ such that $A\mathbf{V}_1 = \lambda_1 \mathbf{V}_1$. If $\lambda \neq \lambda_1$ is any other eigenvalue of A, then $|\lambda| < \lambda_1$, and λ_1 is an eigenvalue of geometric and algebraic multiplicity* 1.

The following Lemma allows us to easily show that correlation matrices of our interest cannot have a pure shift dominant eigenvector.

Lemma 10. *If R is an $n \times n$ positive correlation matrix, property* P1 *implies $r_n < r_{n-1}$, property* P2 *implies $r_1 < r_2$.*

By Remark 4 and Lemma 10 the following result holds:

Proposition 11. *If R is a positive correlation matrix satisfying property* P1 *or* P2, *then* $\mathbf{V}_1 \neq \mathbf{1}$.

We will not consider in this paper the existence problem of shift eigenvectors. Similarly, in the sequel our interest will be devoted to eigenvectors we have called weak slope and weak curvature, then only the sign changes in their components will be of interest.

Given an $n \times n$ correlation matrix R we shall denote by $R[i_1, i_2|j_1, j_2]$ all its minors of order 2, with $1 \leqslant i_1 < i_2 \leqslant n$ and $1 \leqslant j_1 < j_2 \leqslant n$.

3. Totally positive matrices

In this section we will point out the relationship between the correlation matrices under study and some classes of matrices with relevant spectral properties. We need some well-known definitions [1,3,10].

Definition 12. Let A be an $n \times n$ matrix and let $p \in \mathbb{N}$ with $p \leqslant n$. The pth compound matrix A_p of A is defined as the $\binom{n}{p}$-square matrix of the p minors of A.

Definition 13. An $n \times n$ matrix A is called:

- strictly totally positive of order k, denoted by STP_k, if A_p is positive for all $p = 1, \ldots, k$;
- totally positive of order k, denoted by TP_k, if A_p is nonnegative for all $p = 1, \ldots, k$;
- (strictly) totally positive, denoted by TP (STP), if it is TP_n (STP_n);
- oscillatory (O) if it is TP and there exists a $q \in \mathbb{N}\backslash\{0\}$ such that A^q is STP.

Spectral properties of TP and O matrices have been thoroughly investigated in [4] because of their important role in the study of vibrations of mechanical systems. We recall here the following theorem (see [3,17]):

Theorem 14. *An n-dimensional oscillatory matrix A has always n simple positive eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. The eigenvector $\mathbf{V}_k$, for $k = 1, 2, \ldots, n$, belonging to the eigenvalue λ_k has exactly $k - 1$ variations of sign.*

Theorem 14 gives a complete characterization of the two properties we are investigating: the simplicity of the eigenvalues and the sign changes of the corresponding eigenvectors. We have to point out here that this result is much stronger than we need since we are interested only in the first three eigenvectors of a correlation matrix. Anyway, we are going to study under what conditions the conclusions of Theorem 14 can be applied.

We have to stress that our fundamental requirement for correlation matrices of interest rates, namely properties P0–P2, are not enough to assure neither strict total positivity (or total positivity) nor oscillatory properties, as the following example shows.

Example 15. The correlation matrix

$$R = \begin{bmatrix} 1 & 0.85 & 0.8 \\ 0.85 & 1 & 0.9 \\ 0.8 & 0.9 & 1 \end{bmatrix}$$

satisfies properties P0–P2 but is not TP, since $R[1, 2|2, 3] = -0.035$.

Conversely, it is possible to prove that total positivity implies properties P0–P2.

Theorem 16. *An $n \times n$ STP$_2$ correlation matrix $R = [\rho_{ij}]$ satisfies properties* P0–P2.

Proof. Since R is STP$_2$, $\rho_{ij} > 0$ for all i, j; further, the positivity of principal minors of order 2 implies $\rho_{ij} < 1$ for all $i \neq j$.

Consider now the minors $R[i_1, i_2|j_1, j_2]$ with $1 \leqslant i_1 < i_2 = j_1 < j_2 \leqslant n$. Since R is an STP$_2$ matrix, it follows that

$$\rho_{i_1,j_1} \cdot \rho_{i_2,j_2} - \rho_{i_1,j_2} > 0$$

and $\rho_{ij} \in (0, 1)$ for $i \neq j$, implies the following:

$$\begin{cases} 1 \leqslant i_1 < j_1 < j_2 \leqslant n & \rho_{i_1,j_2} < \rho_{i_1,j_1} \Leftrightarrow P1, \\ 1 \leqslant i_1 < i_2 < j_2 \leqslant n & \rho_{i_1,j_2} < \rho_{i_2,j_2} \Leftrightarrow P2. \end{cases} \qquad \square$$

Remark 17. The same conclusion of Theorem 16 can be obtained if R is TP$_2$ and $\rho_{ij} < 1$ for $i \neq j$.

If R is only TP$_2$, then it is not possible to exclude that some of its elements are 1 but, with a proof similar to the one of Theorem 16, the following result can be derived:

$$\begin{cases} 1 \leqslant i_1 < i_2 < j_2 \leqslant n & \rho_{i_1,j_2} \leqslant \rho_{i_2,j_2} \leqslant 1, \\ 1 \leqslant i_1 < j_1 < j_2 \leqslant n & \rho_{i_1,j_2} \leqslant \rho_{i_1,j_1} \leqslant 1. \end{cases}$$

Although, in general, correlation matrices under study are neither STP nor TP, we want to show that in some relevant cases they are.

Example 18. A relevant class of correlation matrices for forward rates is given by

$$\rho_{ij} = \exp\{-\beta|i - j|^q\} \tag{2}$$

with $\beta > 0$, $q \geqslant 1$. For all q these Toepliz matrices are oscillatory. It is not difficult to prove that they are TP by using the criterion of total positivity for Polya Frequency Sequences, explained for example in [1, p. 215], i.e. by direct inspection of the associated Laurent series $\sum_{n=-\infty}^{+\infty} a_n z^n$. In this case $a_n = \rho^{q|n|}$ and since $0 < \rho < 1$, the series converges in the ring of the complex plane $\rho^q < |z| < 1/\rho^q$, hence:

$$\sum_{n=-\infty}^{+\infty} a_n z^n = \sum_{n=0}^{+\infty} \rho^{qn} z^n + \sum_{n=0}^{+\infty} \frac{z^n}{\rho^{qn}} - 1 = \frac{1-\rho^{2q}}{(1-\rho^q z)(1-\rho^q/z)}.$$

Oscillatory character follows then from the positivity of the elements ρ_{ij} with $|i-j| = 1$.

If we identify indices with maturities in the model of the Example 1, we obtain the special case $q = 1$ in (2). Empirically this assumption will be fulfilled only on a partial range of the typical maturity set. Actually, the time intervals $t_{i+1} - t_i$ corresponding to a couple $(i, i+1)$ of successive indices exhibit this feature: for small values of i the time intervals are less than one year, typically of a few months, for big values they become multiples of one year. It is tempting to describe this behavior with a quadratic or, more in general, with a power dependence in maturity difference. A different proof of (strict) total positivity in the case $q = 2$ is given in [10].

Remark 19. We want to recall here the empirically observed property mentioned in the introduction, according to which the reduction of correlations between variables with the same difference in maturities tends to decrease for increasing time indices. This property can be suitably formalized as follows:

$$\rho_{i,j+1} - \rho_{i,j} > \rho_{i,l+1} - \rho_{i,l} \quad i = 1, \ldots, n-3,\ i < j < l \leqslant n-1.$$

It is immediate to verify that in the previous example this property is verified, except for the case $q = 1$, in which equality holds.

4. Total positivity of order 2 and weak slope

The results provided in the previous Section concern all the eigenvectors of a correlation matrix. Now we want to give some results concerning the second eigenvalue based on some properties of STP_k matrices.

Definition 20. If $A = [a_{ij}]$ is an $n \times n$ matrix with positive real entries, define the Birkhoff contraction ratio (or Hopf oscillation ratio) of A by the quantity

$$N(A) = \frac{\kappa(A) - 1}{\kappa(A) + 1},$$

where

$$\kappa^2(A) = \max_{i,j,p,q} \frac{a_{ij} \cdot a_{pq}}{a_{iq} \cdot a_{pj}}.$$

Note that (recall Remark 2) for an $n \times n$ correlation matrix R satisfying properties P0–P2 $\kappa^2(R) = (\rho_{1n})^{-2}$, hence

$$N(R) = \frac{1 - \rho_{1n}}{1 + \rho_{1n}}.$$

We now restate in terms of STP_k two results given in [2].

Theorem 21. *Let A be an $n \times n$ matrix STP_k, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be its eigenvalues, counted according to their multiplicities and ordered so that $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$. Then for $1 \leqslant p \leqslant k$, λ_p is a simple eigenvalue and $\lambda_p > 0$. For $1 \leqslant p \leqslant k-1$*

$$\lambda_{p+1} = \lambda_p \inf_{j \in \mathbb{N}} N\left(A_p^j\right)^{1/j}$$

and

$$|\lambda_{k+1}| = \lambda_k \inf_{j\in\mathbb{N}} N\left(A_k^j\right)^{1/j}.$$

Corollary 22. *Under the hypotheses of Theorem 21, for $1 \leqslant p \leqslant k$, let $m_p(A)$ be the minimum entry in A_p and $M_p(A)$ be the maximum entry in A_p. Then, for $1 \leqslant p \leqslant k-1$,*

$$\lambda_{p+1} \leqslant \frac{M_p(A) - m_p(A)}{M_p(A) + m_p(A)}\lambda_p$$

and

$$|\lambda_{k+1}| \leqslant \frac{M_k(A) - m_k(A)}{M_k(A) + m_k(A)}\lambda_k.$$

In the last years several papers have provided criteria for the determination of TP or STP of a matrix by a reduced number of its minors (see e.g. [1,5,6,19]). In the same spirit, we now investigate more in detail the relations between STP_2 and P0–P2 properties of a correlation matrix by expanding the study sketched in Theorem 16.

Proposition 23. *An $n \times n$ correlation matrix R with $n \geqslant 3$, satisfying properties* P0–P2 *has positive all its minors $R[i_1, i_2|j_1, j_2]$ of the following types*:

(a) $i_1 = j_1$ *and* $i_2 = j_2$;
(b) $i_1 = j_1$ *or* $j_2 = i_2$;
(c) $i_1 < j_1$ *and* $i_2 > j_2$, *or* $j_1 < i_1$ *and* $i_2 < j_2$.

Proof

(a) By Remark 2 all the $\binom{n}{2}$ principal minors of order 2

$$R[i_1, i_2|i_1, i_2] = 1 - (\rho_{i_1,i_2})^2 \quad 1 \leqslant i_1 < i_2 \leqslant n$$

are positive.

(b) By the symmetry of R we can confine ourselves to study only the half of these cases, considering always the elements ρ_{ij} with the row index less than the column one.
If $i_1 = j_1 < i_2 < j_2$ then

$$R[i_1, i_2|j_1, j_2] = \rho_{i_2,j_2} - \rho_{i_1,j_2}\rho_{j_1,i_2}. \tag{3}$$

By property P2 $\rho_{i_2,j_2} > \rho_{i_1,j_2}$; moreover $\rho_{j_1,i_2} \in (0, 1)$, hence positivity of (3) follows. In the case $i_1 < j_1 < j_2 = i_2$ the same conclusion is obtained by property P1.

(c) If $i_1 < j_1$ and $i_2 > j_2$ then

$$R[i_1, i_2|j_1, j_2] = \rho_{i_1,j_1}\rho_{j_2,i_2} - \rho_{i_1,j_2}\rho_{j_1,i_2}$$

and the conclusion descends by the inequalities

$$\rho_{i_1,j_1} > \rho_{i_1,j_2} > 0 \quad \text{and} \quad \rho_{j_2,i_2} > \rho_{j_1,i_2} > 0.$$

By a symmetry argument the case $j_1 < i_1$ and $j_2 < i_2$ immediately follows. □

Remark 24. Note that if R is a correlation matrix with elements $\rho_{ij} \in (0, 1)$ (for $i \neq j$) then the positivity of its minors

$$R[i_1, i_2|i_2, j_2] \quad \text{or} \quad R[i_1, i_2|j_1, i_1]$$

implies the positivity of all the minors

$$R[i_1, i_2|i_1, j_2] \quad \text{or} \quad R[i_1, i_2|j_1, i_2].$$

In fact, if $i_2 = j_1$ by

$$R[i_1, i_2|i_2, j_2] = \rho_{i_1,i_2} \cdot \rho_{i_2,j_2} - \rho_{i_1,j_2} > 0$$

and $\rho_{ij} \in (0, 1)$, it follows:

$$\rho_{i_1,i_2} > \rho_{i_1,j_2} \cdot \frac{1}{\rho_{i_2,j_2}} > \rho_{i_1,j_2} \cdot \rho_{i_2,j_2}$$

and

$$\rho_{i_2,j_2} > \rho_{i_1,j_2} \cdot \frac{1}{\rho_{i_1,i_2}} > \rho_{i_1,j_2} \cdot \rho_{i_1,i_2}.$$

Theorem 25. *An $n \times n$ correlation matrix R with $n \geqslant 3$, satisfying property* P0 *and such that for $i = 2, \ldots, n-1$ and $s = 1, \ldots, i-1$*

$$R[s, s+1|i, i+1] > 0 \tag{4}$$

is STP_2.

Proof. For any fixed $i = 2, \ldots, n-1$, combining assumption $R[i-1, i|i, i+1] > 0$ equivalent to

$$\rho_{i-1,i} > \frac{\rho_{i-1,i+1}}{\rho_{i,i+1}}$$

with assumptions $R[i-1, i|i+k, i+k+1] > 0$ for $k = 1, \ldots, n-i-1$, equivalent to

$$\frac{\rho_{i-1,i+k}}{\rho_{i,i+k}} > \frac{\rho_{i-1,i+k+1}}{\rho_{i,i+k+1}}$$

one obtains

$$\rho_{i-1,i} > \frac{\rho_{i-1,i+k+1}}{\rho_{i,i+k+1}}, \quad k = 1, \ldots, n-i-1$$

i.e. the positivity of all the minors $R[i_1, i_2|i_2, j_2]$ (and by symmetry $R[i_1, i_2|j_1, i_1]$). By the proof of Theorem 16 properties P1 and P2 follow. Then from Proposition 23 the positivity of all the principal minors of order 2 descends.

By symmetry, all 2×2 minors of R with consecutive rows and columns are positive. Then the result follows after an iterated application of the same Theorem 3.2 of Chapter 2 of [10] to the following submatrices of R: the submatrix formed by the first and second columns, the submatrix formed by the second and third columns, and so on, until the submatrix formed by the two last columns. $\square$

Remark 26. The positivity of the principal minors of order 2 in the previous proof is a consequence of properties P0–P2. This argument does not hold for minors of order $\geqslant 3$, as the following example shows:

$$\det \begin{bmatrix} 1 & 0.8 & 0.6 \\ 0.8 & 1 & 0.96 \\ 0.6 & 0.96 & 1 \end{bmatrix} = 0.$$

We can finally state our result on the second eigenvalue of a correlation matrix for interest rates.

Corollary 27. *The second eigenvalue λ_2 of an $n \times n$ correlation matrix R with $n \geqslant 3$, satisfying the hypotheses of Theorem 25 is simple. Furthermore, the following estimate holds*:

$$\lambda_2 \leqslant \frac{1-\rho_{1n}}{1+\rho_{1n}}\lambda_1.$$

Proof. The first conclusion is a consequence of Theorems 21 and 25. The second statement is a direct consequence of Remark 2 and Corollary 22. $\square$

Our estimate of the second eigenvalue of the correlation matrix R involves the first eigenvalue λ_1. It is worth mentioning that many sharp estimates have been given in the literature for the greatest eigenvalue of a nonnegative matrix. We just mention the results in [14,16,18].

Remark 28. Corollary 22 cannot be immediately applied to estimate the third eigenvalue $\lambda_3 \geqslant 0$. In fact, it is immediate to observe that the greatest minor of order 2 of a correlation matrix that satisfies properties P0–P2 is $M_2 = 1 - \rho_{1,n}^2$ but it is not possible to identify in a definite way the minimum m_2, as the following two examples clarify:

$$R_1 = \begin{bmatrix} 1 & 0.87 & 0.7 & 0.55 \\ 0.87 & 1 & 0.92 & 0.75 \\ 0.7 & 0.92 & 1 & 0.95 \\ 0.55 & 0.75 & 0.95 & 1 \end{bmatrix},$$

$$\min R_1[i_1, i_2 | j_1, j_2] = R[1, 2|3, 4] = 0.019,$$

$$R_2 = \begin{bmatrix} 1 & 0.87 & 0.7 & 0.55 \\ 0.87 & 1 & 0.92 & 0.83 \\ 0.7 & 0.92 & 1 & 0.95 \\ 0.55 & 0.83 & 0.95 & 1 \end{bmatrix},$$

$$\min R_2[i_1, i_2 | j_1, j_2] = R[2, 3|3, 4] = 0.044.$$

5. Concluding remarks

In this paper we have studied how some spectral properties of correlation matrices describing the interest rate dynamics can be formally justified. The starting point has been that these properties, namely the simplicity of the first three eigenvalues and regularity in the changes of sign of the corresponding eigenvectors are a restricted version of the properties exhibited by STP and O matrices. We have then investigated how the properties of positivity and monotonicity of the elements of these correlation matrices are related to strict total positivity and oscillatory features. More precisely, we have shown that STP implies positivity and monotonicity, whereas some financially relevant models give correlation matrices that are STP or O. Furthermore, we have proved that for our class of correlation matrices only the positivity of a restricted set of minors of order two implies STP_2, obtaining as corollaries the simplicity and an estimate of the second eigenvalue. We want to outline now the subjects for future developments of the present work. A first step is the extension of the results obtained for the second eigenvalue to the third one. Furthermore, in general, since the properties exhibited by the correlation matrices in interest rate modelling are not sufficient to assure their strict total positivity of order 3, the properties of the first three eigenvectors must be justified in a more general framework. Finally, the investigation, in our terminology, of shift, slope and curvature cases is relevant. Results about these themes are forthcoming.

Acknowledgments

We are grateful to three anonymous referees for their useful suggestions in particular for the simplification of the proof of Proposition 25. In any case, every mistake is ascribed entirely only to the authors.

References

[1] T. Ando, Totally positive matrices, Linear Algebra Appl. 90 (1987) 165–219.
[2] S.P. Eveson, The eigenvalue distribution of oscillatory and strictly sign-regular matrices, Linear Algebra Appl. 246 (1996) 17–21.
[3] F.R. Gantmacher, The Theory of Matrices, Chelsea Publishing Company, New York, 1964.
[4] F.R. Gantmacher, M.G. Krein, Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems, Revised ed., AMS Chelsea Publishing, Providence, RI, 2002.
[5] M. Gasca, J.M. Peña, Total positivity and Neville elimination, Linear Algebra Appl. 165 (1992) 25–44.
[6] G.M.L. Gladwell, Total positivity and the QR algorithm, Linear Algebra Appl. 271 (1998) 257–272.
[7] B.W. Golub, L.M. Tilman, Measuring yield curve risk using principal components analysis, value at risk, and key rate durations, J. Portfolio Management (1997) 72–84.
[8] T. Ho, S. Lee, Term structure movements and pricing interest rate contingent claims, J. Finance 41 (1986) 1011–1029.
[9] J. Hull, A. White, Pricing interest-rate derivative securities, Rev. Fin. Studies 3 (1990) 573–592.
[10] S. Karlin, Total Positivity, vol. 1, Stanford University Press, Stanford CA, 1968.
[11] S. Kotz, W.L. Pearn, D.W. Wichern, Eigenvalue–eigenvectors analysis for a class of patterned correlation matrices with an application, Statist. Probab. Lett. 2 (1984) 119–125.
[12] I.T. Jolliffe, Principal Component Analysis, Springer-Verlag, Berlin, 1986.
[13] S. Lardic, P. Priaulet, S. Priaulet, PCA of the yield curve dynamics: questions of methodologies, J. Bond Trading Management 1 (2003) 327–349.
[14] W. Ledermann, Bounds for the greatest latent roots of a positive matrix, J. London Math. Soc. 35 (1960) 265–268.
[15] R. Litterman, J. Scheinkman, Common factors affecting bond returns, J. Fixed Income 1 (1991) 54–61.
[16] H. Minc, On the maximal eigenvector of a positive matrix, SIAM J. Numer. Anal. 7 (1970) 424–427.
[17] H. Minc, Nonnegative Matrices, Wiley, New York, 1988.
[18] A.M. Ostrowski, Bounds for the greatest latent root of a positive matrix, J. London Math. Soc. 27 (1952) 253–256.
[19] J.M. Peña, Determinantal criteria for total positivity, Linear Algebra Appl. 332–334 (2001) 131–137.
[20] R. Rebonato, Interest-Rate Option Models, II ed., Wiley, New York, 2002.
[21] O. Vasicek, An equilibrium characterization of the term structure, J. Financial Econom. 5 (1977) 177–188.
[22] R. Willner, A new tool for portfolio managers: level, slope, and curvature durations, J. Fixed Income (1996) 48–59.