An EZI method to reduce the rank of a correlation matrix

Massimo Morini*
Cass Business School and
Universita' di Milano Bicocca
Piazza dell'Ateneo Nuovo 1
20126 Milano
Italia
Email: massimo.morini@unimib.it

Nick Webber
City University
Cass Business School
London EC2Y 8HB
United Kingdom
Tel: +44 20 7040 5171
Email: Nick.Webber@city.ac.uk

April 14, 2004

First version: October 2003

Abstract

Reducing the number of factors in a model by reducing the rank of a correlation matrix is a problem that can arise in many areas of finance, for instance pricing interest rate derivatives with Libor market models.

We introduce and describe a simple iterative algorithm for correlation rank reduction, the eigenvalue zeroing by iteration, EZI, algorithm.

We investigate its convergence to an optimal solution and compare its performance with those of other methods. Several test data sets are used including an empirical forward Libor correlation matrix.

The EZI algorithm is very fast even in computationally complex situations, and achieves a level of precision comparable to that of a computationally intensive optimization method.

From our results, the EZI algorithm has superior perfomance in practice to each of the two main methods in current use.

*Corresponding author. We are grateful to Damiano Brigo for useful comments on an earlier manuscript, to Umberto Magnani for helpful discussion, and to participants at the Fifth Quantitative Finance Workshop, Siena.

1 Introduction

The value of financial derivatives often depends on the values and dynamics of more than one underlying market securities. Models for pricing and hedging derivatives can involve many underlying state variables, each modelled as a stochastic process. When explicit solutions for derivatives prices are not available, numerical methods are used, for instance Monte Carlo simulation. Simulation and other methods become computationally burdensome when a high number of independent stochastic driving factors are used. For instance in market models for the term structure of interest rates there may be as many state variables as forward Libor rates, perhaps more than thirty or forty. Hence the problem arises of reducing the number of independent factors to a number small enough for the numerical method to cope with.

When the interdependency among the instantaneous stochastic shocks driving the different variables is Gaussian, with some correlation matrix ρ, then reducing the number of independent stochastic factors is equivalent to reducing the rank of the correlation matrix.

The problem of reducing the rank of an exogenous correlation matrix in financial modelling is the focus of this work. Rebonato and Jäckel (1999) [14] investigate a related problem, discussing the use of an eigenvalue zeroing method (EZN) and a method based on an angles parameterisation (OAP). Brigo (2002) [2] extended their work, specifically examining the same problem as this paper. Zhang and Wu (2003) [18] use Lagrange multiplier techniques, and Grubisic and Pietersz (2003) [9] use geometric optimization.

We introduce an iterative algorithm for correlation rank reduction called eigenvalue zeroing by iteration (EZI). This algorithm is intuitive and simple to implement. Empirical results demonstrate the superiority of our algorithm over the two alternative methods commonly used in finance. It is much more accurate than EZN, achieving an accuracy comparable to the high precision OAP method in a fraction of the time.

In section 2 the role of correlation in multi-dimensional financial models is discussed. In section 3 we formally state the correlation rank reduction problem, present some properties of correlation matrixes, and describe two existing methods for solving the correlation rank reduction problem. In section 4 we introduce and describe the EZI algorithm. In section 4.1 we analyse convergence and in section 4.3 we discuss the optimality. In section 5 we give empirical results for EZI on various data sets including financial data from the interest rate market, comparing its performance with two existing methods. Section 6 concludes.

2 Correlation in multivariate Financial Models

Consider a multivariate model involving M state variables $F_k(t)$, $k = 1, 2, \ldots M$, where the vector $F(t) = (F_1(t), \ldots F_M(t))'$ has the dynamics

$$\mathrm{d}F(t) = \Psi(F, t)\,\mathrm{d}t + \Gamma^i \Upsilon(t)\,\mathrm{d}Y(t), \ i = 1, 2, \tag{1}$$

where, for $r \leqslant M$, $Y(t)$ is a standard r-dimensional uncorrelated Wiener process under a measure Q, $\Psi(F,t)$ is an $M \times 1$ vector, $\Upsilon(t)$ is an $M \times r$ matrix function of time only, and Γ^i is an $M \times M$ diagonal matrix with $\Gamma^1 = I_M$, the identity matrix, and $\Gamma^2 = diag(F(t))$.

When $i = 1$ the diffusion part is modelled as an arithmetic Brownian Motion with time-varying diffusion coefficient. When $i = 2$ the diffusion part is modelled as a geometric Brownian motion with time-varying covariance term $\Upsilon(t)$. If $F(t)$ is a vector of discretely-compound forward rates, then for a suitable $\Psi(F,t)$ the latter model might be a Libor market model.

In (1) $\Upsilon(t)$ represents the covariance structure, describing explicitly the functional relationship between individual shocks. If (1) is used as a starting point to simulate F, the number r of independent stochastic factors heavily affects the computational burden, so that r is often chosen to be much smaller than M, the number of state variables.

Parameters in the functional form for $\Upsilon(t)$ can be found by minimizing some loss function representing the distance between model and market prices. However, in practice the most liquid products may bear imprecise or little, if any, information about an implied correlation matrix ρ, and calibrated correlations are often unreliable and not significant. Therefore ρ is often computed exogenously, for instance estimated from historical market data.

When the correlation matrix ρ is given exogenously we separate out volatility and correlation components. A natural alternative representation to (1) is

$$dF(t) = \Psi(F,t)\,dt + \Gamma^i \Sigma(t)\,dZ(t), \quad i = 1,2, \tag{2}$$

where $Z(t)$ is a standard M-dimensional Wiener process under Q with instantaneous correlation matrix ρ, and $\Sigma(t) = diag\{\sigma_i(t)\}_{i=1,\ldots,M}$ is a matrix of volatilities.

Representations (1) and (2) are equivalent when

$$\Upsilon(t)\,\Upsilon(t)' = \Sigma(t)\,\rho\Sigma(t), \tag{3}$$

in which case $r = rank(\rho)$.

In general an exogenously given correlation matrix ρ is a full M-rank matrix. If we wish to use (1) with $r < M$, we must replace ρ in the representation (3) with a matrix $\hat{\rho}$ approximating ρ such that $rank(\hat{\rho}) = r$.

This is the correlation rank reduction problem in financial modelling. In the next section we present some facts about correlation matrices and describe two existing solution methods.

3 Correlation Rank Reduction

The four properties characterizing a correlation matrix $\rho = \left\{\rho_{i,j}\right\}_{i,j=1,\ldots,M}$ are:

1. Symmetry: $\rho_{i,j} = \rho_{j,i}$, $\forall\ i,j$
2. Positive semidefiniteness: $x'\rho x \geq 0$, $\forall x \in \mathbb{R}^M$

3. Unitary diagonal: $\rho_{i,i} = 1, \ \forall \ i$

4. Normalization: $\left|\rho_{i,j}\right| \leq 1, \quad \forall \ i, j$

Property 4 is redundant. Recall that an $M \times M$ (non-null) matrix A is symmetric positive semidefinite of rank r if and only if there exist an $M \times r$ matrix B of rank r such that

$$A = BB'.$$

As a consequence, through the Cauchy-Schwartz inequality,

$$\left|\rho_{ij}\right| \leq \sqrt{\rho_{ii}\rho_{jj}} \tag{4}$$

for any symmetric positive semidefinite matrix. Hence $\left|\rho_{ij}\right| \leq 1$, and the first three properties are necessary and sufficient conditions for ρ to be a viable correlation matrix.

Define the following sets

$$\begin{aligned}
S &= \left\{Y \in \mathbb{R}^{M\times M} \mid Y = Y'\right\}, \\
P &= \left\{Y \in \mathbb{R}^{M\times M} \mid x'Yx \geq 0, \ x \in \mathbb{R}^{M}\right\}, \\
D &= \left\{Y \in \mathbb{R}^{M\times M} \mid y_{ij} = 0, \ i \neq j\right\} \\
U &= \left\{Y \in R^{M\times M} \mid y_{ii} \leq 1, \ i = 1, \ldots, M\right\}, \\
U_x &= \left\{Y \in U \mid y_{ii} = x, \ i = 1, \ldots, M\right\}.
\end{aligned}$$

Let S be the set of $M \times M$ symmetric matrixes, P the set of $M \times M$ positive semidefinite matrixes, D the set of diagonal matrixes, and U_1 the set of $M \times M$ matrixes with unit diagonal. S and U_0 are vector subspaces and P is a cone. $U_x = U + xI$ is an affine space. S, P, and U_1 are closed and convex. The set C of $M \times M$ correlation matrices is

$$C = S \cap P \cap U_1,$$

which is also closed and convex.

We define K_r to be the set of all $M \times M$ matrices of rank less than or equal to r

$$K_r = \left\{Y \in \mathbb{R}^{M\times M} \mid rank\left(Y\right) \leq r\right\}.$$

K_r is closed but not convex.

We can now state the correlation rank reduction problem. Given a correlation matrix $\rho \in C$, the problem is to find a matrix $\hat{\rho} \in C \cap K_r$ such that

$$\hat{\rho} = \underset{Y \in C \cap K_r}{\arg\min} \left\{\|Y - \rho\|\right\}, \tag{5}$$

where $\|\cdot\|$ is some chosen metric. Following Rebonato and Jäckel (1999) [14], the we use the metric

$$s^2 = \|A - B\|^2 = \sum_{i,j=1}^{M} \left|A_{ij} - B_{ij}\right|^2, \tag{6}$$

induced by the Frobenius norm.

In a market application context there are additional issues to consider in solving (5).

Firstly, an algorithm based on an exact optimization may be computationally intensive. In practical applications accuracy must be traded off against computation time. It may be sensible to use algorithms that do not generate an optimal $\hat{\rho}$.

Secondly, in practical applications it is necessary to exploit a decomposition $\hat{\rho} = JJ'$. The characterstics of J affect the results when $\hat{\rho}$ is used for pricing and hedging (an example is given in section 5.)

We briefly present some results that will be used later in our discussion of correlation rank reduction methods.

The angles parametrisation of a correlation matrix (Rebonato and Jäckel (1999) [14].) A parametric form for $A \in C \cap K_r$ of rank r is

$$A = BB'$$

where $B = \{b_{i,k}\}_{i=1,\ldots,M,\ k=1,\ldots,r}$ is an $M \times r$ matrix whose i^{th} row is given by

$$\begin{aligned} b_{i,1} &= \cos\theta_{i,1}, \\ b_{i,k} &= \sin\theta_{i,1} \ldots \sin\theta_{i,k-1}\cos\theta_{i,k}, \quad 1 < k < r, \\ b_{i,r} &= \sin\theta_{i,1} \ldots \sin\theta_{i,r-1}. \end{aligned} \tag{7}$$

Set $\theta = \{\theta_{i,k}\}_{i=1,\ldots,M,\ k=1,\ldots,r}$. We write $A(\theta)$ for the angles parameterisation of A.

Spectral decomposition of a correlation matrix $\rho \in C$ is symmetric positive definite so it admits M linearly independent eigenvectors forming an orthonormal basis with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$. Let $\Delta = diag\{\lambda_i\}_{i=1,\ldots,M}$ and form the eigenvectors into a matrix X so that

$$\rho X = X\Delta,$$

and $\rho = X\Delta X^{-1} = X\Delta X'$. We can write $\Delta = \Lambda\Lambda'$, where $\Lambda = \Lambda'$ is diagonal and $\Lambda_{ii} = \sqrt{\lambda_i}$, leading to

$$\rho = X\Lambda\Lambda' X' = (X\Lambda)(X\Lambda)'. \tag{8}$$

We have the following theorem for symmetric matrices.

Theorem 1 *(Optimal rank reduction) Let $A \in S \cap K_k$ be of rank $k \geq r$ and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ with $|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_k|$, be its non-zero eigenvalues. Consider the spectral decomposition $A = X\Delta X'$ where $\Delta = diag\{\lambda_1, \ldots, \lambda_k, 0, \ldots, 0\}_{i=1,\ldots,k}$. Then for any matrix $B \in K_r$*

$$\|B - A\|^2 = \sum_{i,j=1}^{M} (B_{ij} - A_{ij})^2 \geqslant \lambda_{r+1}^2 + \cdots + \lambda_k^2$$

and equality is attained when

$$B = X\tilde{\Delta}X', \;\; with \;\; \tilde{\Delta} = diag\left(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0\right).$$

(See Harville (1999) [11]).

3.1 Rank reduction methods

Accounts of two techniques for correlation rank reduction are given by Brigo (2002) [2] and Rebonato and Jäckel (1999) [14].[1] We now describe these techniques.

The first method, eigenvalue zeroing with normalization (EZN), is based on the spectral decomposition of ρ, followed by a rescaling of the resulting matrix entries to recover a viable correlation matrix. The second method is called optimization of the angles parameterization (OAP) in the following. It is based on the angles parameterization of ρ allowing standard unconstrained optimisation techniques to be applied to the correlation rank reduction problem.

3.1.1 Eigenvalue zeroing with normalization (EZN)

EZN is based on the spectral decomposition (8) of a symmetric positive definite matrix, $\rho = (X\Lambda)(X\Lambda)'$. Let $\tilde{\Lambda}$ be the matrix obtained by setting to zero the $M - r$ last entries of Λ, corresponding to the smallest $M - r$ eigenvalues of ρ. Set $L_r = S \cap P \cap U \cap K_r$ and define

$$f_1 \;\; : \;\; C \to L_r, \tag{9}$$

$$f_1 \;\; : \;\; \rho \mapsto \tilde{\rho} = \left(X\tilde{\Lambda}\right)\left(X\tilde{\Lambda}\right)'. \tag{10}$$

$f_1(\rho)$ is symmetric and its eigenvalues are non-negative so it is positive semi-definite, but the diagonal values $\tilde{\rho}$ of are not necessarily equal to 1. Since

$$\rho_{ii} = \sum_{j=1}^{M} X_{ij}^2 \lambda_j = 1, \; i = 1, \ldots, M,$$

setting to zero the $M - r$ smallest eigenvalues implies

$$0 \leq \tilde{\rho}_{ii} = \sum_{j=1}^{r} X_{ij}^2 \lambda_j \leq 1, \; i = 1, \ldots, M,$$

so although from (4) property 4 is still satisfied, property 3 can be lost.

Brigo (2002) [2] and Rebonato and Jäckel (1999) [14] propose a rescaling. Set $B = X\tilde{\Lambda}$ and write B_i for the *i-th* row[2] of B. Set $\hat{\rho} = \left\{\hat{\rho}_{ij}\right\}_{i,j=1,\ldots,M}$ to be

$$\hat{\rho}_{ij} = \frac{\tilde{\rho}_{ij}}{\sqrt{\tilde{\rho}_{ii}\tilde{\rho}_{jj}}} = \frac{(BB')_{ij}}{\sqrt{(BB')_{ii}(BB')_{jj}}} = \left(\frac{B_i}{\sqrt{B_i(B_i)'}}\right)\left(\frac{(B_j)'}{\sqrt{B_j(B_j)'}}\right). \tag{11}$$

[1] Rebonato and Jäckel invetigate a problem related to ours, that of finding some correlation matrix close to a general matrix.

[2] Given a matrix A, we write A_i for its i-th row and A^j for its j-th column.

This defines a map $\langle\cdot\rangle : S \cap P \cap K_r \to C$. $\langle\tilde{\rho}\rangle = \hat{\rho}$ is the EZN r-rank correlation matrix. $\langle\cdot\rangle$ is well defined only if $\tilde{\rho}_{ii} > 0$ for all i.

EZN appears to be a method commonly used by practitioners. As we will see in section 5, it is very fast but also inaccurate.

3.1.2 Optimization of the angles parameterization (OAP)

The OAP method exploits the angles parameterisation of a correlation matrix. The problem (5) is reduced to the unconstrained optimization

$$\hat{\rho} = \arg\min_{\theta} \|\rho - A(\theta)\|$$

where $A(\theta)$ is the angles parameterisation of A. The result $\hat{\rho}$ of this optimisation is the OAP reduced rank correlation matrix.

The OAP method is much more accurate than EZN, as will be shown in empirical tests in Section 5. However it relies on a complex optimisation which can be very slow, in particular when M and r are large.

Brigo, Mercurio and Rapisarda (2002) [6] point out that there is a redundancy in the angles parameterization, leading, in the full-rank case $r = M$, to twice as many parameters as degrees of freedom in an $M \times M$ correlation matrix. They suggest a way to remove this redundancy that could ease the optimisation. However an efficient choice of initial parameter values θ_0 for the optimization has not been found for their method, while for standard OAP an efficient θ_0 can be found by using the EZN rank reduced correlation matrix. This makes standard OAP more efficient in practice, in spite of its redundancy, so we will use it, rather than the alternative method, in the sequel.

4 The eigenvalue zeroing by iteration algorithm, EZI

The OAP method requires the use of general optimization techniques to solve problem (5). The only specific structure it imposes is the use of the angles parameterization. On the other hand EZN has a strong theoretical justification, based on Theorem 1; setting to zero the $M - r$ smallest eigenvalues is always the optimal procedure to reduce the rank of a symmetric matrix.[3] Since correlation matrices are symmetric, it induces the map f_1 for reducing the rank of a correlation matrix ρ. $\tilde{\rho} = f_1(\rho)$ is the optimal matrix satisfying properties 1, 2 and 4 but $\tilde{\rho}$ is not guaranteed to be in C.

The second step (11) in the EZN algorithm modifies $\tilde{\rho}$ to obtain property 3, but the resulting matrix is no longer optimal.

It is easy to see that when $A \in S \cap P$ with $a_{ii} \leq 1$, $i = 1, \ldots, M$, then for any $B \in S \cap P$ with $b_{ii} = 1$, $i = 1, \ldots, M$, the quantity $\|B - A\|^2$ is minimized when $B = A + H$ where

$$H = diag(1 - a_{11}, 1 - a_{22}, \ldots, 1 - a_{MM}).$$

[3] With respect to the metric $\|\cdot\|$ defined in (6).

Hence we can define a map

$$\begin{aligned} f_2 &: S \cap P \cap U \to C, \\ f_2 &: \tilde{\rho} \mapsto \rho = \tilde{\rho} + diag\left(1 - \tilde{\rho}_{11}, \ldots, 1 - \tilde{\rho}_{MM}\right). \end{aligned}$$

We also denote by f_2 the induced map $f_2 : L_r = S \cap P \cap U \cap K_r \to C$. f_2 yields the correlation matrix ρ closest to $\tilde{\rho}$, but there is now no guarantee that $f_2(\tilde{\rho}) \in K_r$.

This suggests the possibility of iterating the two maps f_1 and f_2. Successive application of the map $g = f_1 \circ f_2 : C \to C$ may give an improvement in accuracy compared to EZN while, unlike OAP, exploiting the efficiency of the eigenvalues zeroing procedure.

We formally state the algorithm. We start from an $M \times M$ correlation matrix ρ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$, whose rank is to be reduced to r.

Algorithm 2 ***Eigenvalues zeroing by iteration (EZI)***

1. *Set the iteration number* $s = 1$, $\rho^s = \rho$ *and* $a_s = I$.
2. *Reduce the rank:* Set $\tilde{\rho}^s = f_1(\rho^s)$ and $a_s = \|\rho^s - \tilde{\rho}^s\|$
3. *If a stopping condition is true, stop.* Return $\hat{\rho} = \langle\tilde{\rho}^s\rangle$ where $\langle\tilde{\rho}^s\rangle$ is defined by (11).
4. *Recover a unit diagonal:* Set $\rho^{s+1} = f_2(\tilde{\rho}^s)$ and go to step (1).

There are two stopping conditions.

1. For a tolerance level ε_1, stop when $\|\tilde{\rho}^s - \langle\tilde{\rho}^s\rangle\| < \varepsilon_1$.
2. Stop when $\|a_s - a_{s-1}\| < \varepsilon_2$.

In the first case $\langle\tilde{\rho}^s\rangle$ is surely a viable correlation of the required rank, and it is not distinguishable from $\tilde{\rho}^s$ to the tolerance ε_1.[4]

In the second case the algorithm can proceed no further, and convergence to $\hat{\rho} \in C \cap K_r$ has not been achieved.

We now investigate convergence of the algorithm and identify structures in ρ^s that can prevent the algorithm from converging to a rank reduced correlation matrix.

4.1 Convergence

We analyse the convergence of the sequence $a_s = \|\rho^s - \tilde{\rho}^s\|$. Anticipating the empirical results presented in Section 5, we find that for financially relevant initial matrices the EZI algorithm converges to a matrix in $C \cap K_r$ and gives

[4] Other stopping rules can be considered based upon distances of $\tilde{\rho}^s$ or ρ^s from the initial ρ or on $\|\rho^{s+1} - \rho^s\|$. Using a rule of the former type we found the performance of the algorithm was affected very little.

considerable improvements in accuracy compared to EZN. This convergence is achieved in a small number of iterations, so that EZI is computationally efficient.

We use s-superscripts to denote matrixes at the sth iteration, so that $\rho^s = X^s \Delta^s X^{s\prime}$ for instance.

Let $\bar{\Delta}^s = diag\left\{0, ..., 0, \lambda_{r+1}^s, ...\lambda_M^s\right\} \in D$ and define

$$\begin{aligned} Z^s &= X^s \bar{\Delta}^s X^{s\prime} = \sum_{j=r+1}^{M} \lambda_j^s \left(X^s\right)^j \left(X^s\right)^{j\prime}, \\ D^s &= diag\left\{Z_{ii}^s\right\}_{i=1,...,M} \end{aligned}$$

so that $\rho^s - \tilde{\rho}^s = Z^s$, and

$$\rho^{s+1} - \tilde{\rho}^s = diag\left\{1 - \tilde{\rho}_{ii}^s\right\} = diag\left\{\sum_{j=r+1}^{M} \lambda_j^s \left(X_{ij}^s\right)^2\right\} = diag\left\{Z_{ii}^s\right\} = D^s.$$

By Theorem 1, $\left\|\rho^{s+1} - \tilde{\rho}^{s+1}\right\| \leq \left\|\rho^{s+1} - \tilde{\rho}^s\right\|$, so $\left\|Z^{s+1}\right\| \leq \|D^s\|$ and

$$0 \leq \left\|Z^{s+1}\right\| \leq \|D^s\| \leq \|Z^s\| \tag{12}$$

so the sequence $a_s = \|\rho^s - \tilde{\rho}^s\| = \|Z^s\|$ converges uniformly to $a \geq 0$.

If $\left\|Z^{s+1}\right\| = \|Z^s\|$ for some s then $\|D^s\| = \|Z^s\|$ so $Z^s = D^s$ is diagonal and $\|Z^t\| = \|Z^s\|$ for all $t \geq s$. Conversely if $Z^s = D^s$ is diagonal for some s then $\|Z^t\| = \|Z^s\|$ for all $t \geq s$.

If $\left\|Z^{s+1}\right\| < \|Z^s\|$ for all s then $\|Z^s\|$ converges to some $a \geq 0$. Since $a < \|D^s\| < \|Z^s\|$ we have $\|Z^s\| - \|D^s\| = \|Z^s - D^s\| \to 0$, that is, the non-diagonal elements of Z^s go to zero. Since $D_{ii}^s \geq 0$ for all i we conclude $D^s \to D^\infty \in D$ and $Z^s \to D^\infty \in D$ also. Note that $a_{s+1} < a_s$ if and only if $Z^s \notin D$.

Since $\rho^{s+1} - \rho^s = D^s - Z^s$ we have $\left\|\rho^{s+1} - \rho^s\right\| \to 0$, achieving the limit if $Z^s = D^s$ for some s. Similarly, $\tilde{\rho}^{s+1} - \tilde{\rho}^s = D^s - Z^{s+1}$ and $\left\|\tilde{\rho}^{s+1} - \tilde{\rho}^s\right\| \to 0$.

We also have $\left\|\rho^{s+1} - \tilde{\rho}^s\right\| = \|D^s\| \leq \|Z^s\| = \|\rho^s - \tilde{\rho}^s\|$ and since $\rho^s - \tilde{\rho}^{s+1} = Z^s - D^s + Z^{s+1}$ we see that

$$\begin{aligned} \left\|\rho^s - \tilde{\rho}^{s+1}\right\|^2 &\leq \|Z^s - D^s\|^2 + \left\|Z^{s+1}\right\|^2 = \|Z^s\|^2 - \|D^s\|^2 + \left\|Z^{s+1}\right\|^2 \\ &\leq \|Z^s\|^2 = \|\rho^s - \tilde{\rho}^s\|^2. \end{aligned}$$

and $\left\|\rho^s - \tilde{\rho}^{s+1}\right\| \leq \|\rho^s - \tilde{\rho}^s\|$.

In fact, note that

$$\rho - \tilde{\rho}^s = \sum_{i=1}^{s-1} \left(Z^i - D^i\right) + Z^s,$$

and $\rho - \tilde{\rho}^{s+1} = \rho - \tilde{\rho}^s - D^s + Z^{s+1}$. Since

$$diag\left\{\rho - \tilde{\rho}^s\right\} = diag\left\{Z_{ii}^s\right\} = D^s \tag{13}$$

we have $\|\rho - \tilde{\rho}^s - D^s\| = \|\rho - \tilde{\rho}^s\| - \|D^s\|$ so

$$\begin{aligned} \left\|\rho - \tilde{\rho}^{s+1}\right\| &\leq \|\rho - \tilde{\rho}^s - D^s\| + \left\|Z^{s+1}\right\| \\ &\leq \|\rho - \tilde{\rho}^s\| - \|D^s\| + \left\|Z^{s+1}\right\| \\ &\leq \|\rho - \tilde{\rho}^s\|. \end{aligned}$$

Hence $\left\|\rho - \tilde{\rho}^{s+1}\right\| \leq \|\rho - \tilde{\rho}^s\| \leq \left\|\rho - \tilde{\rho}^1\right\|$ for all s.

If $a_s \longrightarrow 0$ then there exists $\rho^\infty = \langle \tilde{\rho}^\infty \rangle \in C \cap K_r$ such that $\rho^s \longrightarrow \rho^\infty$. The EZI algorithm converges to a matrix ρ^∞ which is a correlation matrix of the required rank r.

If $\|Z^s\| = 0$ for some s we have convergence to zero so $\langle \tilde{\rho}^s \rangle = \tilde{\rho}^s = \rho^s = \rho^\infty$ and the algorithm has converged to a correlation matrix of the desired rank. If $0 \neq Z^s \in D$ for some s then $a = \|Z^s\|$ and the algorithm stops. If $0 \neq Z^s \notin D$ is never diagonal then we have seen that Z^s converges to a diagonal matrix D^∞ and stopping rule 2 will apply.

Even if $a_s \nrightarrow 0$ as $s \longrightarrow \infty$ then EZI can be better than EZN. EZI returns $\langle \tilde{\rho}^s \rangle$ for some $s \geq 1$. EZN returns $\langle \tilde{\rho}^1 \rangle$. Since $\left\|\rho - \tilde{\rho}^{s+1}\right\| \leq \left\|\rho - \tilde{\rho}^1\right\|$ EZI always returns a matrix rescaled from one at least as close to ρ as that returned by EZN.

4.2 Eigenvector structure

Suppose that for some s, $0 \neq Z^s \in D$ so that $a = \|Z^s\| \neq 0$. We investigate implications for the structure of eigenvectors of X^s and of ρ^s.

Since

$$Z^s = X^s \bar{\Delta}^s X^{s\prime} = X^s \bar{\Delta}^s \left(X^s\right)^{-1}, \tag{14}$$

the matrices Z^s and $\bar{\Delta}^s$ are similar, and have the same set of eigenvalues. In a diagonal matrix the diagonal values are equal to the eigenvalues. Therefore when Z^s and $\bar{\Delta}^s$ are both diagonal they have the same elements, possibly in a different order. (14) implies

$$Z^s \left(X^s\right)^i = \left(\bar{\Delta}^s\right)_{ii} \left(X^s\right)^i, \quad i = 1, \ldots, M, \tag{15}$$

where $\left(X^s\right)^i$ is the i-th column of X^s.

Let $\mathbb{M} = \{1, 2, \ldots, M\}$ and define

$$\alpha^Z = \left\{j \in \mathbb{M} \mid Z^s_{jj} = 0\right\}, \tag{16}$$

$$\alpha^\Delta = \left\{i \in \mathbb{M} \mid \bar{\Delta}^s_{ii} = 0\right\}, \tag{17}$$

to be the set of all indices for Z^s and $\bar{\Delta}^s$ corresponding to null diagonal elements. α^Z and α^Δ have the same number of elements. Write $\bar{\alpha}^Z = \mathbb{M} \backslash \alpha^Z$ and $\bar{\alpha}^\Delta = \mathbb{M} \backslash \alpha^\Delta$ for the set of indices not in α^Z or α^Δ respectively.

From (15),

$$Z^s \left(X^s\right)^i = \begin{cases} 0, & i \in \alpha^\Delta, \\ \bar{\Delta}^s_{ii} \left(X^s\right)^i, \ \bar{\Delta}^s_{ii} > 0, & i \notin \alpha^\Delta. \end{cases}$$

This implies

$$\left.\begin{array}{l} i \in \alpha^{\Delta},\ j \in \bar{\alpha}^{Z},\ \text{or} \\ i \in \bar{\alpha}^{\Delta},\ j \in \alpha^{Z} \end{array}\right\} \Longrightarrow X^{s}_{ij} = 0.$$

If $0 \neq Z^s \in D$, the columns $(X^s)^i$ are divided in two sets: those with index in α^{Δ} have zeros in every position not in α^{Z}, while those with indexes not in in α^{Δ} have zeros in every position in α^{Z}.

The Structure of ρ^s We analyse now the structure of ρ^s. X^s is also the eigenvector matrix for ρ^s, and

$$\rho^{s}_{i,j} = \left(X^{s}\sqrt{\Delta^{s}}\right)_{i}\left(X^{s}\sqrt{\Delta^{s}}\right)'_{j}.$$

Since X^s comes from the spectral decomposition of the symmetric matrix ρ^s, its rows (and its columns) form an orthonormal set. Each row X_i, $i \in \alpha^{Z}$, is orthogonal to each row X_j, $j \in \bar{\alpha}^{Z}$, and this property is preserved when X^s is post-multiplied by Δ^s. In fact

$$\left.\begin{array}{l} i \in \alpha^{Z},\ j \in \bar{\alpha}^{Z},\ \text{or} \\ i \in \bar{\alpha}^{Z},\ j \in \alpha^{Z} \end{array}\right\} \Longrightarrow \rho^{s}_{ij} = \left(X^{s}\sqrt{\Delta^{s}}\right)_{i}\left(X^{s}\sqrt{\Delta^{s}}\right)'_{j} = 0.$$

Furthermore, when $i, j \in \bar{\alpha}^{Z}$,

$$\rho^{s}_{ij} = \sum_{k=1}^{M} \lambda^{s}_{k} X^{s}_{ik} X^{s}_{jk} = \sum_{k \notin \alpha^{D}} \lambda^{s}_{k} X^{s}_{ik} X^{s}_{jk} = Z^{s}_{ij}.$$

Since we assumed $Z^s \in D$, if $i \neq j$ and one of i or j is in $\bar{\alpha}^{Z}$, then $\rho^{s}_{ij} = 0$. Set $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 is in the ith position. Then for $i \in \bar{\alpha}^{Z}$,

$$\rho^{s}_{i} = \left((\rho^{s})^{i}\right)' = e_i. \tag{18}$$

Consider a general $M \times M$ correlation matrix $\rho \in C$. If there exists $\phi \neq H \subseteq \mathbb{M}$ such that $i \in H \Longrightarrow \rho^{s}_{i} = e_i$, then say that ρ has block elementary structure for indices $i \in H$. When ρ has a block elementary structure then for each index $i \in H$, ρ has an uncorrelated factor.

We have shown that Z^s diagonal and non-null always implies a block elementary structure with uncorrelated factors for all $i \in \bar{\alpha}^{Z}$. We now determine conditions for an uncorrelated factor to have index in $\bar{\alpha}^{Z}$.

Clearly, given a matrix A with block elementary structure, there exists a permutation matrix Q such that QAQ' is block-diagonal. This implies that A is reducible; there exists $H \subseteq \mathbb{M}$ such that

$$i \in H,\ j \in \mathbb{M} \setminus H \Rightarrow a_{ij} = 0.$$

QAQ' is the Gantmacher Normal Form (GNF) of the reducible matrix A (De Giuli and Magnani (1999) [8]). The spectral properties of a matrix are invariant when moving to its GNF, since QAQ' is similar to A.

Therefore if ρ^s has block elementary structure, λ is an eigenvalue of ρ^s if and only if λ is an eigenvalue of one of the diagonal blocks in the GNF of ρ^s. The diagonal block represented by a diagonal value ρ^s_{ii} associated with an uncorrelated factor has a unit eigenvalue.

Consequently for $0 \neq Z^s \in D$, since $\bar{\alpha}^Z \subseteq \{r+1, \ldots, M\}$ if $i \in \bar{\alpha}^Z$ then the associated eigenvector $\lambda = 1$ is one of the $M-r$ smallest eigenvalues. When there exist uncorrelated factors in a block elementary structure belong to the set of $M-r$ smallest eigenvalues we call this relevant block elementary structure (RBE structure) and call these uncorrelated factors relevant uncorrelated factors.[5]

We have shown that

$$0 \neq Z^s \in D \Longrightarrow \rho^s \; has \; RBE \; structure. \tag{19}$$

A stronger statement is possible. Suppose ρ^s has χ relevant uncorrelated factors. When $0 \neq Z^s \in D$, $\chi = \left|\bar{\alpha}^Z\right|$. From the definition of $\bar{\alpha}^Z$, these are the only non-zero eigenvalues in the set of the $M-r$ smallest eigenvalues. Therefore

$$rank\,(\rho^s) = r + \chi$$

with r eigenvalues ≥ 1.

When a correlation matrix with χ relevant uncorrelated factors has rank $r+\chi$, say that it has fixed RBE (FRBE) structure. So

$$0 \neq Z^s \in D \Longrightarrow \rho^s \; has \; FRBE \; structure. \tag{20}$$

Matrixes with FRBE structure are fixed points of the EZI algorithm. When $\chi = 0$ a matrix with FRBE structure is in $C \cap K_r$.

RBE Structure and convergence We have seen if $\|\rho^s - \tilde{\rho}^s\| \longrightarrow a = \|Z^s\|$ for some s then ρ^s has RBE structure for some s. We now show the converse.

Suppose ρ^s has RBE structure with $\chi = |H|$ relevant uncorrelated factors for indices $i \in H \subseteq \mathbb{M}$. Let Q be the permutation matrix moving rows with indices in H to the first χ positions. Then $Q\rho^s Q' = QX^s D^s X^{s\prime} Q'$ is block diagonal,

$$Q\rho^s Q' = \begin{pmatrix} I & 0 \\ 0 & \rho^t_{SE} \end{pmatrix},$$

where ρ^t_{SE} is $(M-\chi) \times (M-\chi)$. Set to zero the $M-r$ smallest eigenvalues. These include the unit eigenvalues in the top left block. We find again a block diagonal matrix, where the top left block is null,

$$Q\tilde{\rho}^s Q' = QX^s \tilde{D}^s X^{s\prime} Q' = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\rho}^s_{SE} \end{pmatrix},$$

where $\tilde{\rho}^s_{SE}$ is $(M-\chi) \times (M-\chi)$. Then $Q^{'} QX^s \tilde{D}^s X^{s\prime} Q' Q = X^s \tilde{D}^s X^{s\prime} = \tilde{\rho}^s$ has

$$\tilde{\rho}^s_i = \left(\left(\tilde{\rho}^s\right)^i\right)' = 0, \; i \in H.$$

[5] When $\lambda_r = \lambda_{r+1} = 1$ there is no unique choice of eigenvalue to set to zero. However the algorithm always chooses the same eigenvalue to set to zero.

Correct the diagonal to recovery a correlation matrix ρ^{s+1}. Then

$$\rho_i^{s+1} = \left(\left(\rho^{s+1} \right)^i \right)' = e_i,\ i \in H.$$

Since in EZI the r biggest eigenvalues never decrease in size,[6] if a unit eigenvalue belongs to the set of the $M-r$ smallest eigenvalues at iteration s, then it belongs to this set for any $t \geq s$ so ρ^s and $\tilde{\rho}^s$ will always differ by 1's in the i-th diagonal values, $i \in H$. Hence

$$\|\rho^s - \tilde{\rho}^s\| \geq \chi \nrightarrow 0 \ as\ s \longrightarrow \infty.$$

and we conclude

Proposition 3 *The EZI algorithm hits a fixed point with $a = \|Z^t\|$ if and only if ρ^s has RBE structure, for some $s \leq t$.*

This implies that if the initial matrix ρ has RBE structure, the EZI algorithm does not ultimately converge to a correlation matrix of rank r. Instead, stopping rule 2 will apply at some point.

We have identified a set of correlation matrixes ρ for which the EZI algorithm will not converge to a rank-r correlation matriz. Even so, EZI will always improve on the initial correlation matrix ρ (unless ρ already has FRBE structure).

An extreme case is when $\rho = I$, the identity matrix. In this case there is no correlation at all among the underlying variables. Here the problem of reducing model dimension via reducing the rank of the correlation matrix is not determined.

Note that if $\tilde{\rho} \in L_r$ has RBE structure, then the map $\langle \cdot \rangle$ (11) is undefined (since some denominators would be zero). Thus neither EZN nor EZI with the stopping condition 1 can be used.

If $a_s \to a \neq 0$ and $a \neq \|Z^s\|$ foir some s, we know that Z^s converges to a diagonal matrix D^∞. The analysis above goes through 'module ε'.

When a matrix ρ has RBE structure, a different approach can be considered. Assume ρ_{ii} is a diagonal block in the GNF affected by rank reduction. Blocks like this, together with adjacent null blocks, can be cut out from ρ. The other eigenvalues of such a matrix are not affected by this operation and the matrix consisting of the remaining blocks is still a correlation matrix. Its rank can be reduced and then the ρ_{ii} block can be re-inserted in the resulting matrix. The re-insertion will increase the rank, but no other eigenvalues will be altered.

In the next section we address the different issue of the optimality of EZI. We analyze the properties of maps and sets involved in the algorithm. This shows how EZI relates to the theory of alternating projections, which is the foundation of the optimality of some classes of iterative algorithms for approximation problems.

[6] See Section 4.3.

4.3 EZI and alternating projections

The correlation rank reduction problem lacks explicit solutions. This appears related to the fact that rank-reduction is a basis-independent feature, depending on properties of the eigenvalues, whereas property 3 depends on the basis (Higham (2002) [12]).

The EZI algorithm can be formulated as an alternating projections scheme. Suppose H is a Hilbert space and suppose $T_k \subseteq H$, $k = 1, \ldots, K$ are subsets of H. A map $\Pi_k : H \to T_k$ is an orthogonal projection if $\Pi_k^2 = \Pi_k$ and $\langle a - \Pi_k(h),\ h - \Pi_k(h) \rangle = 0$ for all $a \in T_k$, $h \in H$. The point $\Pi_k(h) \in T_k$ is the point in T_k closest to $h \in H$. Let $T = \bigcap_{k=1}^{K} T_k$ and suppose $T \neq \phi$, and let Π_T be the orthogonal projection onto T.

Von Neumann (1950) [17] showed that if the T_k are closed subspaces of H then for $h \in H$

$$\lim_{n \to \infty} \left\| \left(\Pi_K \Pi_{K-1} \ldots \Pi_1 \right)^n (h) - \Pi_T (h) \right\| = 0. \tag{21}$$

so that the limit of alternating projections onto each T_k converges to the projection onto T.

Deutsch and Hundall (2002) [7] showed the same results applies when the T_k are affine sets[7] rather than subspaces.

This theory has been extended to sets T_k which are closed and convex. Han (1988) [10] shows that, in this case, the iteration of alternating projections does not ensure that a projection onto the intersection is reached. Convergence can be to non-optimal points. For closed and convex sets the optimal algorithm involves a correction to the projections. It reduces to the Von Neumann case when the T_k are closed subspaces or affine sets. See Boyle and Dykstra (1985) [1].

Maps in EZI are orthogonal projections. f_1 is an orthogonal projection into $L_r = S \cap P \cap U \cap K_r$ and f_2 is an orthogonal projection into C so EZI is consists of alternating projections.[8] Neither C nor L_r are subspaces, nor is L_r convex, so existing the theory of alternating projections does not apply. Nevertheless C and L_r enjoy particular properties that may help explain the good behaviour displayed by the EZI algorithm in empirical tests.

We now give some results for matrices in L_r. For $Y \in \mathbb{R}^{M \times M}$ write $\lambda_i(Y)$ for its ith largest eigenvalue.

Theorem 4 *Let A be an $M \times M$ symmetric matrix and B an $M \times M$ nonnegative definite matrix. Then for $i = 1, \ldots, M$, we have*

$$\lambda_i (A + B) \geq \lambda_i (A),$$

(See Schott (1996) [16]).

[7] Here, an affine set is defined to be a translatation of a subspace.

[8] We thank Igor Grubisic for recently signalling an application of alternating projections to correlation rank reduction in his master thesis "Interest rate theory. BGM model", Leiden University, 2002.

When both A,$B \in S \cap D$

$$\lambda_i (A+B) \geq \max (\lambda_i (A), \lambda_i (B)) \geq 0.$$

Suppose the initial correlation matrix is ρ. The rank reduction step does not affect the r highest eigenvalues, and Theorem 4 can be applied to the map f_2, so for any rank reduced matrix $\tilde{\rho}^s = f_1 (\rho^s)$ generated by the EZI algorithm

$$\lambda_i (\tilde{\rho}^s) \geq \lambda_i (\rho), \ i = 1, \ldots, r. \tag{22}$$

Define the set L_r^ρ as

$$L_r^\rho = \left\{ Y \in L_r \mid \lambda_i (Y) \geq \frac{1}{2} \lambda_i (\rho), \ i = 1, \ldots, r \right\}$$

(The reason for the factor of $\frac{1}{2}$ will become clear shoprtly). Then f_1 is the projection $\Pi_r^\rho : C \longrightarrow L_r^\rho$ giving the closest matrix in L_r^ρ to matrixes in C.

L_r^ρ is not convex, but we can investigate its distance from convexity. Suppose A,$B \in L_r^\rho$. From Theorem 4, for all $0 \leq \alpha \leq 1$, $i = 1, \ldots, r$,

$$\begin{aligned} \lambda_i (\alpha A + (1-\alpha) B) &\geq \max (\alpha \lambda_i (A), \ (1-\alpha) \lambda_i (B)) \\ &\geq \max (\alpha \lambda_i (\rho), \ (1-\alpha) \lambda_i (\rho)) \\ &\geq \frac{1}{2} \lambda_i (\rho). \end{aligned}$$

Since for any $Y \in L_r \subset U$ we have

$$\sum_{i=1}^{M} \lambda_i (Y) = tr (Y) \leq M,$$

it follows that for $0 \leq \alpha \leq 1$, $\sum_{i=1}^{M} \lambda_i (\alpha A + (1-\alpha) B) \leq M$ and hence

$$\begin{aligned} \sum_{i=r+1}^{M} \lambda_i (\alpha A + (1-\alpha) B) &\leq M - \sum_{i=1}^{r} \lambda_i (\alpha A + (1-\alpha) B) \\ &\leq M - \sum_{i=1}^{r} \frac{1}{2} \lambda_i (\rho). \end{aligned}$$

We define the distance of a matrix X from L_r^ρ as

$$\| X, \ L_r^\rho \| = \min_{Y \in L_r^\rho} \| X - Y \|.$$

Since the minimum is achieved for a matrix Y which is also the minimum in K_r we have

$$\| \alpha A + (1-\alpha) B, \ L_r^\rho \| \leq \left(M - \sum_{i=1}^{r} \frac{1}{2} \lambda_i (\rho) \right)^2.$$

This bound is related to the distance from convexity of the set L_r^ρ. It depends on the spectral properties of the initial matrix ρ, on M and on r. The bound is larger when the eigenvalues of ρ are of similar magnitude and when $M - r$ is high. It is narrower when some eigenvalues dominate the others, and when $M - r$ is lower.

These observations could help explain the empirical evidence that, although sets involved do not satisfy known theoretical requirements ensuring optimality, the EZI algorithm behaves very well in terms of both accuracy and speed. This happens for all data sets considered in the empirical tests in the next section, and is particularly notable when $M - r$ in small.[9]

Incorporating a projection correction? The above analysis suggests that results might plausibly be further improved by incorporating the projection correction given by Dykstra (1983) [15] and Boyle and Dykstra (1985) [1] for ensuring optimality when dealing with closed and convex sets. The correction is: at iteration s, before projecting onto a set T_k, subtract a vector normal to T_k so as to remove the increment associated with the projection onto T_k at iteration $s - 1$. Boyle and Dykstra point out the correction only applies when the target set is convex but not affine. Hence we apply the correction to f_1 but not to f_2.

A natural implementation of this correction into EZI algorithm is then

Algorithm 5

1. *Set the iteration number* $s = 1$ *and* $\rho^1 = \rho$ *and* $B^0 = 0$.
2. *Set* $\rho^s = \rho^s - B^{s-1}$.
3. *Reduce the rank:* $\tilde{\rho}^s = X^s \tilde{D}^s X^{s\prime}$. *Set* B^s $B^s = \tilde{\rho}^s - \rho^s$.
4. *If the stopping condition is true, stop.*
5. *Recovery unit diagonal:* $\rho^{s+1} = \tilde{\rho}^s + D^s$, *where* $D^s = diag\left\{1 - \tilde{\rho}_{ii}^s\right\}$. *Set* $s = s + 1$ *and go to 2).*

Unfortunately the algorithm (5) fails. In step 2) we subtract $B^{s-1} = \tilde{\rho}^{s-1} - \rho^{s-1}$from ρ^s. In general we obtain a matrix with diagonal values greater than unity, so $\tilde{\rho}^s$ computed at step 3) may also have this property, and f_2 may not project onto C.

Algorithm (5) was implemented and applied to the test matrixes of the next section. It performed badly on realistic matrixes, halting at approximations worse than those obtained by EZI, and returning matrixes that were not positive semidefinite. We do not report these results.

[9] In addition, when $M - r$ is large we can have, at some step of the EZI algorithm, a large number of eigenvalues all very close to 0. When λ_r is very close to λ_{r+1} conmputing the best r-rank approximation of a general symmetric matrix can encounter numerical problems in setting to zero the $M - r$ smallest eigenvalues.

5 Numerical results

We give numerical results to compare the performance EZI with that of EZN and OAP. Four test matrixes are used; a 3×3 matrix ρ_1 used by Rebonato and Jäckel (1999) [14] in a risk management example, a full rank 10×10 matrix ρ_2 used by Brigo (2002) [2] as a stylized forward rates correlation matrix, a variation ρ_3 on this also used by Brigo, and a 19×19 empirically estimated forward rate correlation matrix ρ_4.

The test matrixes have neither negative correlations, nor dependence structures with no dominant components. This is usual for financial applications where such correlations and structures would be atypical. Our test matrixes represent features commonly found in money market applications, for instance.

We describe the test matrixes and then present numerical results.

Example: ρ_1 The test matrix is

$$\rho_1 = \begin{pmatrix} 1 & 0.9 & 0.7 \\ 0.9 & 1 & 0.4 \\ 0.7 & 0.4 & 1 \end{pmatrix}$$

ρ_1 is not a correlation matrix.[10] The eigenvalues of ρ_1 are

$$(2.29673,\ 0.71063,\ \text{-}0.00733)\,.$$

The matrix has a negative eigenvalue, so applying correlation matrix rank reductions methods may be problematical. However the negative eigenvalue is also the smallest one in absolute value, so setting it to zero one simultaneously recovers a viable correlation matrix and reduces the rank. Hence EZI, OAP and EZN can be applied as usual.

Example: ρ_2 The test matrix is a full rank 10×10 matrix $\rho_2 = \left\{\rho_{2,ij}\right\}_{i,j=1,\ldots,10}$ with parametric form

$$\rho_{2,ij} = 0.5 + (1 - 0.5)\exp\left(-0.05\,|i-j|\right),\ i,j = 1,\ldots,10.$$

Brigo (2002) [2] uses ρ_2 to give comparisons between EZN and OAP. Eigenvalues of ρ_2 are given in table 1 which also shows the cumulative percentage variation accounted for by the leading eigenvalues.

Example: ρ_3 This 10×10 matrix $\rho_3 = \left\{\rho_{3,ij}\right\}_{i,j=1,\ldots,10}$ is given by the parametric form

$$\rho_{3,i,j} = \exp\left(-\,|i-j|\right),\ i,j = 1,\ldots,10$$

also used by Brigo (2002) [2]. As $|i-j|$ increases there is a sharp decrease in correlations. Eigenvalues of ρ_3 are given in table 1.

[10] Rebonato and Jäckel's goal is to recover a viable correlation matrix starting from ρ_1.

	Example ρ_2		Example ρ_3	
	Eigenvalue	Cum. Var.	Eigenvalue	Cum. Var.
1	9.27091	92.709%	2.03858	20.386%
2	0.42072	96.916%	1.73616	37.747%
3	0.12675	98.184%	1.39443	51.692%
4	0.05924	98.776%	1.10155	62.707%
5	0.03595	99.136%	0.88079	71.515%
6	0.02496	99.385%	0.72383	78.753%
7	0.01912	99.576%	0.61546	84.908%
8	0.01580	99.734%	0.54263	90.334%
9	0.01377	99.872%	0.49619	95.296%
10	0.01278	100.000%	0.47039	100.000%

Table 1: Eigenvalues for examples rho(2) and rho(3)

Example: ρ_4 The final test matrix ρ_4 is a 19×19 correlation matrix for annual discretely compounded forward rates. The matrix is estimated from market quoted Euro forward rates during the period from February 1, 2001 to February 1, 2002. Eigenvalues of ρ_4 are given in table 2. Reducing the rank of matrixes such as these is an essential step in calibrating Libor market models of interest rates.

$$\rho_4 = \begin{pmatrix}
1.00 & 0.82 & 0.69 & 0.65 & 0.58 & 0.47 & 0.29 & 0.23 & 0.43 & 0.47 & 0.33 & 0.43 & 0.29 & 0.23 & 0.26 & 0.21 & 0.23 & 0.29 & 0.25 \\
0.82 & 1.00 & 0.80 & 0.73 & 0.68 & 0.55 & 0.45 & 0.40 & 0.53 & 0.57 & 0.42 & 0.45 & 0.48 & 0.34 & 0.35 & 0.32 & 0.32 & 0.31 & 0.32 \\
0.69 & 0.80 & 1.00 & 0.76 & 0.72 & 0.63 & 0.47 & 0.56 & 0.67 & 0.61 & 0.48 & 0.52 & 0.48 & 0.54 & 0.46 & 0.42 & 0.45 & 0.42 & 0.35 \\
0.65 & 0.73 & 0.76 & 1.00 & 0.78 & 0.67 & 0.58 & 0.56 & 0.68 & 0.70 & 0.56 & 0.59 & 0.58 & 0.50 & 0.50 & 0.48 & 0.49 & 0.44 & 0.35 \\
0.58 & 0.68 & 0.72 & 0.78 & 1.00 & 0.84 & 0.66 & 0.67 & 0.71 & 0.73 & 0.70 & 0.67 & 0.64 & 0.59 & 0.58 & 0.65 & 0.65 & 0.53 & 0.42 \\
0.47 & 0.55 & 0.63 & 0.67 & 0.84 & 1.00 & 0.77 & 0.68 & 0.73 & 0.69 & 0.77 & 0.69 & 0.66 & 0.63 & 0.61 & 0.68 & 0.70 & 0.57 & 0.45 \\
0.29 & 0.45 & 0.47 & 0.58 & 0.66 & 0.77 & 1.00 & 0.72 & 0.71 & 0.65 & 0.65 & 0.62 & 0.71 & 0.62 & 0.63 & 0.66 & 0.64 & 0.52 & 0.38 \\
0.23 & 0.40 & 0.56 & 0.56 & 0.67 & 0.68 & 0.72 & 1.00 & 0.73 & 0.66 & 0.64 & 0.56 & 0.61 & 0.72 & 0.59 & 0.64 & 0.64 & 0.49 & 0.46 \\
0.43 & 0.53 & 0.67 & 0.68 & 0.71 & 0.73 & 0.71 & 0.73 & 1.00 & 0.75 & 0.59 & 0.66 & 0.69 & 0.69 & 0.69 & 0.63 & 0.64 & 0.52 & 0.40 \\
0.47 & 0.57 & 0.61 & 0.70 & 0.73 & 0.69 & 0.65 & 0.66 & 0.75 & 1.00 & 0.63 & 0.68 & 0.70 & 0.63 & 0.64 & 0.65 & 0.62 & 0.52 & 0.40 \\
0.33 & 0.42 & 0.48 & 0.56 & 0.70 & 0.77 & 0.65 & 0.64 & 0.59 & 0.63 & 1.00 & 0.83 & 0.72 & 0.64 & 0.58 & 0.68 & 0.73 & 0.57 & 0.45 \\
0.43 & 0.45 & 0.52 & 0.59 & 0.67 & 0.69 & 0.62 & 0.56 & 0.66 & 0.68 & 0.83 & 1.00 & 0.82 & 0.69 & 0.67 & 0.70 & 0.69 & 0.65 & 0.43 \\
0.29 & 0.48 & 0.48 & 0.58 & 0.64 & 0.66 & 0.71 & 0.61 & 0.69 & 0.70 & 0.72 & 0.82 & 1.00 & 0.79 & 0.78 & 0.79 & 0.72 & 0.59 & 0.42 \\
0.23 & 0.34 & 0.54 & 0.50 & 0.59 & 0.63 & 0.62 & 0.72 & 0.69 & 0.63 & 0.64 & 0.69 & 0.79 & 1.00 & 0.82 & 0.83 & 0.79 & 0.60 & 0.45 \\
0.26 & 0.35 & 0.46 & 0.50 & 0.58 & 0.61 & 0.63 & 0.59 & 0.69 & 0.64 & 0.58 & 0.67 & 0.78 & 0.82 & 1.00 & 0.90 & 0.80 & 0.50 & 0.22 \\
0.21 & 0.32 & 0.42 & 0.48 & 0.65 & 0.68 & 0.66 & 0.64 & 0.63 & 0.65 & 0.68 & 0.70 & 0.79 & 0.83 & 0.90 & 1.00 & 0.94 & 0.71 & 0.46 \\
0.23 & 0.32 & 0.45 & 0.49 & 0.65 & 0.70 & 0.64 & 0.64 & 0.64 & 0.62 & 0.73 & 0.69 & 0.72 & 0.79 & 0.80 & 0.94 & 1.00 & 0.82 & 0.66 \\
0.29 & 0.31 & 0.42 & 0.44 & 0.53 & 0.57 & 0.52 & 0.49 & 0.52 & 0.52 & 0.57 & 0.65 & 0.59 & 0.60 & 0.50 & 0.71 & 0.82 & 1.00 & 0.84 \\
0.25 & 0.32 & 0.35 & 0.35 & 0.42 & 0.45 & 0.38 & 0.46 & 0.40 & 0.40 & 0.45 & 0.43 & 0.42 & 0.45 & 0.22 & 0.46 & 0.66 & 0.84 & 1.00
\end{pmatrix} \tag{23}$$

Empirical matrix, ρ_4		
	Eigenvalue	Cum. Var.
1	11.69922	61.575%
2	2.14784	72.879%
3	1.18028	79.091%
4	0.71657	82.863%
5	0.64130	86.238%
6	0.42730	88.487%
7	0.38601	90.519%
8	0.33889	92.302%
9	0.28045	93.778%
10	0.25419	95.116%
11	0.19951	96.166%
12	0.16925	97.057%
13	0.16108	97.905%
14	0.15035	98.696%
15	0.08773	99.158%
16	0.06008	99.474%
17	0.05153	99.745%
18	0.03328	99.920%
19	0.01514	100.000%

Table 2: Eigenvalues for empirical example, rho(4)

Comparison of methods: errors and times						
Matrix:	ρ_1	ρ_2			ρ_3	
Target rank, r:	2	2	4	7	4	7
EZN:	1.004e-4 (0.02)	0.1134 (0.02)	0.0164 (0.05)	2.32e-3 (0.02)	6.14 (0.07)	1.20 (0.06)
OAP:	0.946e-4 (0.08)	0.0764 (0.34)	0.0069 (3.1)	0.916e-3 (15.2)	5.95 (12.7)	1.12 (27.6)
EZI:	0.946e-4 (0.02) [7]	0.0765 (0.29) [50]	0.0070 (0.08) [21]	0.918e-3 (0.05) [8]	5.96 (0.14) [13]	1.13 (0.10) [6]

Table 3: Comparison of rank reduction methods

Comparison of methods for the empirical matrix: errors and times							
Target rank, r:	2	4	6	8	10	12	14
EZN:	27.04 (0.03)	9.00 (0.04)	3.67 (0.04)	1.51 (0.03)	0.56 (0.03)	0.24 (0.03)	0.046 (0.03)
OAP:	19.11 (9.2)	4.54 (50.6)	1.51 (207)	0.60 (650)	0.23 (1608)	0.098 (2203)	0.022 (2523)
EZI:	19.59 (1.8) [157]	4.84 (0.86) [70]	1.57 (0.53) [38]	0.61 (0.33) [13]	0.23 (0.28) [9]	0.101 (0.27) [8]	0.022 (0.24) [5]

Table 4: Comparison of rank reduction methods

5.1 Comparison of results

The EZN, OAP and EZI methods were implemented in Matlab are run on a 1 Ghz Pentium III PC. Table 3 gives results for ρ_1, ρ_2 and ρ_3, table 4 for ρ_4. For each test matrix results are shown for reducing the rank down to a target rank r. For each method the error is shown[11] and the time taken in seconds (in round brackets). The EZI tolerance level was taken to be $\varepsilon_1 = 0.000001$. The algorithm stopped by rule 1 in every example. The number of iterations needed by the EZI method to achieve the tolerance level is given in square brackets. Matlab default stopping criteria with maximum function evaluations set to 10^7 were used to stop the OAP method.

Matrix ρ_1 is already almost a rank 2 correlation matrix and EZN performs almost as well as OAP. ρ_2 is harder to reduce. EZN performs progressively worse than OAP as r increases. ρ_3 is much further than ρ_2 from being a reduced rank correlation matrix. Errors here are significanty larger. However, EZN performs better here, in terms of relative error compared to OAP, than for ρ_2. For ρ_4 EZN is consistently less accurate than OAP, usually achieving an error twice as

[11] Using the criteria (6). Other error criteria were also computed, for instance the relative error, $SER = \sum_{i,j} \left(\frac{a_{ij} - b_{ij}}{a_{i,j}} \right)^2$, but very similar patterns of results were found so they are not reported.

great as the minimum achieved by OAP.

EZI consistently performs well. In the artificial examples in table 3 it is always within 1% of OAP, and often within 0.1%. For the more realistic case of ρ_4 given in table 4 its error is always within a few percent of that of OAP, often much closer. Its accuracy increases as r increases.

Note however the computation times. EZN is consistently fast, but EZI is usually not much slower. The worse relative performance for EZI compared to EZN is for ρ_4, particularly when r is small. However in these cases the accuracy of EZN is unacceptably low.

By contrast OAP is considerably slower than either EZN or EZI. Already for the toy example ρ_3 it is taking about one hundred times longer than EZI to achieve only slightly greater accuracy. For the empirical example ρ_4 it is taking over forty minutes to compute the $r = 14$ case, $10,000$ longer than EZI with the same accuracy. Computation time for OAP increases as r increases, whereas for EZI the number of interations, and hence the computation times, decreases. Even in the best case for OAP, $r = 2$, OAP is over five times slower but only a few percent more accurate.

EZI has clear advantages in speed over OAP and in accuracy over EZN. We conclude that for many typical applications where extreme accuracy is not required but rapid computation is, EZI may be the method of choice.

5.2 LMM calibration

From a financial point of view, computational speed and accuracy with respect to the initial matrix are not the only interesting features. We can also assess the behaviour of an algorithm when it is part of a pricing or calibration procedure, investigating if, and how, it affects final results. We consider an application of EZI to the calibration of a reduced rank Libor market model.

Brigo and Mercurio (2001) [3], (2002) [4] and Brigo, Mercurio and Morini (2002) [5] introduce a calibration method, the cascade algorithm, for a Libor market model, enabling the model to exactly recover swaption prices. Crucially their method involves the rank reduction of a correlation matrix ρ to a rank r correlation matrix $\hat{\rho}$. The method is sensitive to $\hat{\rho}$ so the choice of rank reduction method is important.

The EZI method can be readily used in this context. The output correlation matrix $\langle\tilde{\rho}^s\rangle$ is easily decomposed as $\langle\tilde{\rho}^s\rangle = JJ'$. With notation as in (11), set $J_i = \frac{B_i^s}{\sqrt{B_i^s \left(B_i^s\right)'}}$ where $B^s = X^s\sqrt{\tilde{D}^s}$. Then

$$\langle\tilde{\rho}^s\rangle_{ij} = \left(\frac{B_i^s}{\sqrt{B_i^s \left(B_i^s\right)'}}\right)\left(\frac{\left(B_j^s\right)'}{\sqrt{B_j^s \left(B_j^s\right)'}}\right) = J_i J_j'.$$

We use both EZI and OAP with the cascade method to reduced the rank of the correlation matrix ρ_4 and to calibrate to swaption market data for February

1, 2002.[12]

While both methods give results that are usually regular and significant, OAP can encounter numerical problems. When EZI was used no numerical problems were found with any version of the historical correlation matrix.

OAP is the computationally most burdensome step when it is used in cascade calibration. With EZI the entire calibration procedure is very fast, avoiding the use of an optimisation routine.

6 Conclusions

In this paper we introduce an iterative algorithm, EZI, for correlation matrix rank reduction. The algorithm is intuitive and simple to implement.

Although convergence to an optimal correlation matrix cannot be guaranteed, an analysis of conditions affecting its convergence to a viable correlation matrix of desired rank is presented. A correction to the algorithm is considered.

We present empirical tests using both stylized matrices and an historically estimated market forward rate correlation matrix. We compare EZI with two commonly used methods, the fast but inaccurate EZN method and slow but accurate OAP method. Results show that EZI is very close in accuracy to OAP, and is much more accurate than EZN. However the speed of EZI is comparable to EZN, making it much faster than OAP, particularly when required rank is large.

The EZI algorithm appears to provide a good balance of accuracy and speed, and is recommended for practical applications.

References

[1] J. P. Boyle and R. L. Dykstra. *A method for finding projections onto the intersection of convex sets in Hilbert spaces*, volume 37, Advances in Order Resticted Inference, of *Lecture Notes in Statistics*, pages 28–47. Springer, Berlin, 1985.

[2] D. Brigo. A note on Correlation and Rank Reduction. Working Paper, www.damianobrigo.it, 2002.

[3] D. Brigo and F. Mercurio. *Interest Rate Models Theory and Practice*. Springer Finance, 2001.

[4] D. Brigo and F. Mercurio. Calibrating LIBOR. *Risk Magazine*, January 2002.

[12] We used a modification of the cascade algorithm, as described in Morini and Brigo (2003) [13], including endogenous interpolate to recover values of missing data. The modified algorithm reduces or avoids certain numerical difficulties that might otherwise occur, such as negative volatility parameters.

[5] F. Mercurio D. Brigo and M. Morini. The Libor Model Dynamics: Approximations, Calibration and Diagnostics. *European Journal of Operation Research*, 2002. Forthcoming.

[6] F. Mercurio D. Brigo and F. Rapisarda. Parametrizing correlations: a geometric interpretation. Working paper, www.fabiomercurio.it, 2002.

[7] F. Deutsch and H. Hundal. The Rate of Convergence for the Method of Alternating Projections, II. *J.Math.Anal.and Appl.*, 205:381–405, 1997.

[8] M. E. De Giuli and U. Magnani. *Matematica per l'Economia*. Università degli Studi di Pavia, 1998.

[9] I. Grubisic and R. Pietersz. Efficient rank reduction of correlation matrices. Preprint, Utrecht University, 2003.

[10] S. P. Han. A successive projection method. *Math.Prog*, 40:1–14, 1988.

[11] D. A. Harville. *Matrix Algebra From a Statistician's Perspective*. Springer, 1999.

[12] J. Higham. Computing the nearest correlation matrix - A problem from finance. *IMA journal of Numerical Analysis*, 22:329–343, 2002.

[13] M. Morini and D. Brigo. An efficient analytical cascade calibration of the Libor market model to swaptions relying only on quoted data. Working Paper, extended abstract in *Amases XXVII* conference proceedings, 2003.

[14] R. Rebonato and P. Jäckel. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. QUARC preprint, 1999.

[15] R.L.Dykstra. An algorithm for restricted least squares regression. *J.Amer.Stat.Assoc.*, 78:837–842, 1983.

[16] J. R. Schott. *Matrix Analysis for Statistics*. Wiley, 1996.

[17] J. von Neumann. *Functional Operators. Vol.II. The geometry of orthogonal spaces*, volume 22 (reprint of 1933 notes) of *Annals of Math.Studies*. Princeton University Press, 1950.

[18] Z. Zhang and L. Wu. Optimal low-rank approximation to a correlation matrix. *Linear Algebra and its Applications*, (364 161187), 2003.