# Computing the Nearest Correlation Matrix— A Problem from Finance

Nick Higham
Department of Mathematics
University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

THE UNIVERSITY
of MANCHESTER

## The MathWorks, Solution Number 23117

**Product:** Financial Toolbox

**Problem:**
Why do I receive error messages stating that my covariance matrix is not positive semi-definite when it should be? Sometimes this happens when I've created it with `cov` or `ewstats`.

**Solution:**
These situations can occur when you give `cov` or `ewstats` a small number of effective samples, increasing the effects of machine imprecision and causing small negative eigenvalues in the covariance matrix ...

# MakeVC

Financial Engineering Associates, Inc.,
www.fea.com

. . . . . .

MakeVC is a stand-alone application that reads historical price and interest-rate time series data and uses exponentially-weighted moving averages to forecast volatilities and correlations.

. . . . . .

## Features

. . . . . .

*Matrix Adjustment* You can check if a volatility-correlation matrix is positive-definite and, if not, automatically shift its eigenvalues to make it positive-definite.

## Stock Research

- Predictions based on sample correlation matrices constructed from vectors of stock returns.
- On a given day, data rarely available for all stocks of interest.
- Can compute sample correlations of pairs of stocks based on days on which both stocks have data available.
- Resulting matrix of correlations is *approximate*, since built from inconsistent data sets.
- Relatively few vectors of observations available, so approximate correlation matrix has *low rank*.

## Constructing the Stock Correlations

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Stock 1 | $\surd$ | – | $\surd$ | – | $\surd$ |
| Stock 2 | $\surd$ | $\surd$ | – | $\surd$ | $\surd$ |
| Stock 3 | – | – | $\surd$ | $\surd$ | $\surd$ |
| Stock 4 | $\surd$ | $\surd$ | $\surd$ | $\surd$ | – |
| Stock 5 | – | $\surd$ | $\surd$ | – | $\surd$ |
| Stock 6 | – | $\surd$ | $\surd$ | – | $\surd$ |

Denoting rows above by $x_i^T$,

$$a_{ij} = \frac{(x_i - \overline{x}_i)^T (x_j - \overline{x}_j)}{\|x_i - \overline{x}_i\|_2 \|x_j - \overline{x}_j\|_2}.$$

# How to Proceed

Either

Plug the gaps in the missing data, then compute an exact correlation matrix.

or

Compute and work with the *nearest correlation matrix*.

## Properties of Correlation Matrices

**Theorem 1 (Schur, Horn)** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. A necessary and sufficient condition for $A$ to have eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ and diagonal elements $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_n$ (in any order along the diagonal) is that*

$$\sum_{i=1}^{k} \lambda_i \le \sum_{i=1}^{k} \alpha_i, \qquad k = 1\colon n,$$

*with equality for $k = n$.*

**Conclusion**

For a correlation matrix any set of nonnegative $\lambda_i$ summing to $n$ is possible.

- An efficient algorithm of Bendel & Mickey (1978) [improved by Davies & Higham, 2000] transforms a given spsd matrix with $\sum_i \lambda_i = n$ into a correlation matrix.

# Problem

Compute distance

$\gamma(A) = \min\{\, \|A - X\| : X \text{ is a correlation matrix}\,\}$

and a matrix achieving the distance.

Use a weighted Frobenius norm:

★ $\|A\|_W = \|W^{1/2} A W^{1/2}\|_F \quad (W \text{ positive definite}),$

★ $\|A\|_H = \|H \circ A\|_F \quad (h_{ij} > 0),$

where $\|A\|_F^2 = \sum_{i,j} a_{ij}^2$.

**Note:** $\|A\|_H \equiv \|A\|_W$ for $W = \operatorname{diag}(w_i)$,
$h_{ij} = (w_i w_j)^{1/2}$.

We describe results only for $W$.

## Theory

Correlation matrix lies in intersection of convex sets

$$\begin{aligned} S &= \{\, Y = Y^T \in \mathbb{R}^{n\times n} : Y \geq 0 \,\}, \\ U &= \{\, Y = Y^T \in \mathbb{R}^{n\times n} : y_{ii} = 1,\ i = 1\colon n \,\}. \end{aligned}$$

Define

$$\langle A, B\rangle = \mathrm{trace}(A^T W B W).$$

Convex optimization theory says that $X$ is characterized by

$$\langle Z - X, A - X\rangle \leq 0 \quad \text{for all } Z \in S \cap U,$$

that is,

$$\begin{aligned} A - X &\in \partial(S \cap U)(X) \\ &= \partial S(X) + \partial U(X). \end{aligned}$$

# Characterization

**Theorem 2** *The correlation matrix* $X$ *solves the problem iff*

$$X = A + W^{-1}\big(VDV^T + \mathrm{diag}(\theta_i)\big)W^{-1}, \qquad (1)$$

*where* $V \in \mathbb{R}^{n\times p}$ *has orthonormal cols spanning* $\mathrm{null}(X)$, $D = \mathrm{diag}(d_i) \geq 0$, *and the* $\theta_i$ *are arbitrary.*

**Theorem 3** *Let* $a_{ii} \geq 1$ *and* $W$ *be diagonal. Then, in* (1), $\theta_i \leq 0$ *for all* $i$. *Moreover, if* $A$ *has* $t$ *nonpositive eigenvalues then* $X$ *has at least* $t$ *zero eigenvalues.*

## Example

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad \min \|A - X\|_F.$$

$\lambda(A) = \{\, 1 \pm \sqrt{2}, 1 \,\}$.

Is $X = ee^T$ the nearest correlation matrix?

$\mathrm{Null}(X)$ spanned by cols of

$$V = \begin{bmatrix} 1 & -1 \\ -1 & 0 \\ 0 & 1 \end{bmatrix},$$

so from Theorem 2

$$X = A + \begin{bmatrix} d_1 + d_2 & -d_1 & -d_2 \\ -d_1 & d_1 & 0 \\ -d_2 & 0 & d_2 \end{bmatrix} + \mathrm{diag}(\theta_i)$$

with $d_i \geq 0$. This equation implies $d_2 = -1$ and so $X$ cannot be a solution.

Solution is

$$X = \begin{bmatrix} 1.0000 & 0.7607 & 0.1573 \\ 0.7607 & 1.0000 & 0.7607 \\ 0.1573 & 0.7607 & 1.0000 \end{bmatrix}.$$

## Previous Work

**Knol & ten Berge, 1989:** write the correlation matrix as $X^TX$ and let $x_k$ denote $k$th column of $X$. Write

$$\|A - X^TX\|_F^2 = f_k(x_k) + \rho_k,$$

with $\rho$ indep. of $x_k$ and

$$f_k(x_k) = \|Cx_k - b\|_2^2,$$

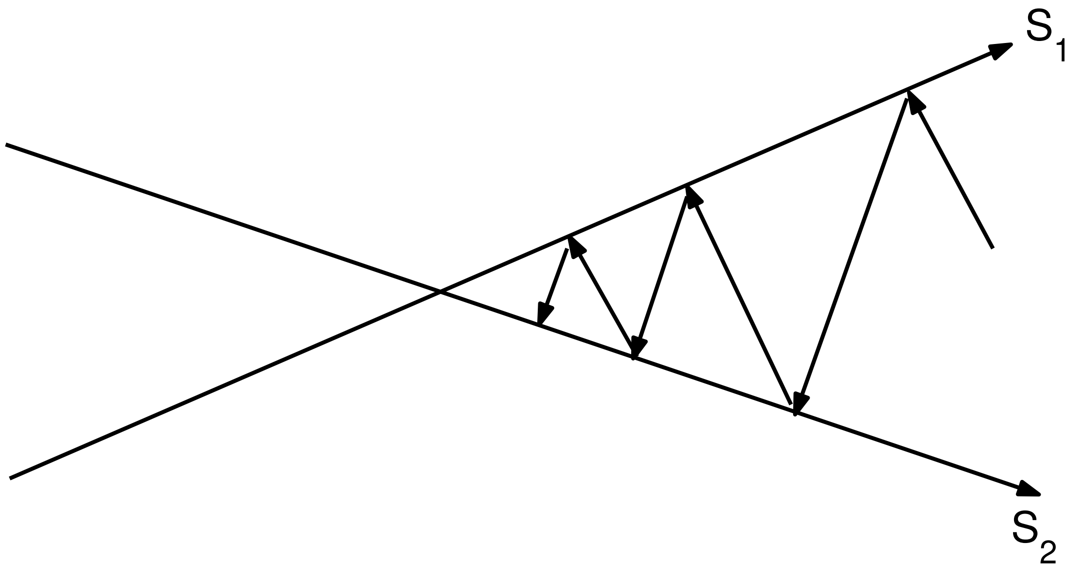
$C \in \mathbb{R}^{(n-1)\times n}$ and $b$ indep. of $x_k$. Repeatedly minimize over each $k$ in turn: generates decreasing sequence of norms.

**Lurie & Goldberg, 1998:** use Gauss–Newton to minimize $\|A - R^TR\|_F^2$, where $R$ is upper triangular with cols of unit 2-norm.

# Alternating Projections

To find nearest point in intersection of sets $S_1$ and $S_2$, iteratively project (von Neumann, 1933):

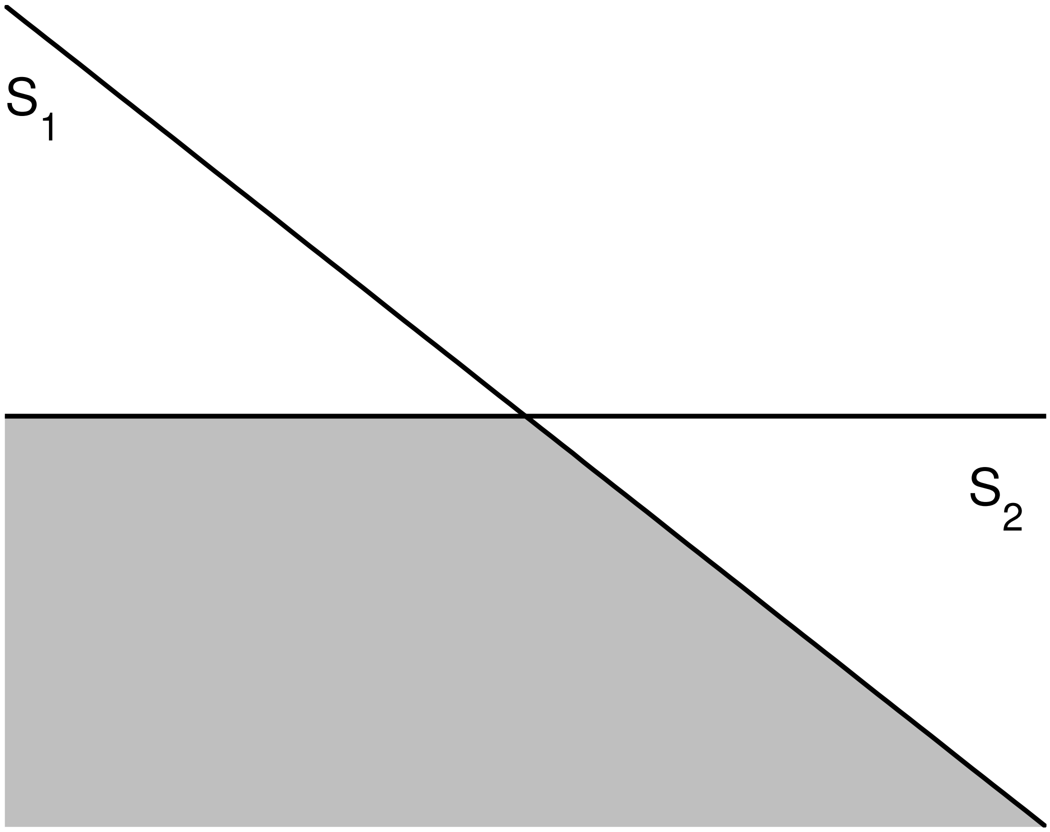
$$X \leftarrow (P_2 P_1)^k A.$$




Converges if $S_2$ and $S_2$ are subspaces.

## Modified Alternating Projections

Dykstra (1983) incorporated corrections for the case of closed convex sets.



## Projecting onto $U$ and $S$

- Recall, $\|A\|_W = \|W^{1/2}AW^{1/2}\|_F$.
- For $A = Q\,\mathrm{diag}(\lambda_i)Q^T$ let

$$A_+ := Q\,\mathrm{diag}(\max(\lambda_i, 0))Q^T.$$

**Theorem 4**

$$P_U(A) = A - W^{-1}\,\mathrm{diag}(\theta_i)W^{-1},$$

*where* $\theta = [\theta_1, \ldots, \theta_n]^T$ *is the solution of the linear system*

$$(W^{-1} \circ W^{-1})\theta = \mathrm{diag}(A - I).$$

**Theorem 5**

$$P_S(A) = W^{-1/2}\big((W^{1/2}AW^{1/2})_+\big)W^{-1/2}.$$

*Moreover,*

$$\mathrm{diag}(P_S(A)) \geq \mathrm{diag}(A).$$

## Algorithm

Given symmetric $A \in \mathbb{R}^{n\times n}$ this algorithm computes nearest correlation matrix:

$$
\begin{aligned}
&\Delta S_0 = 0,\ Y_0 = A\\
&\text{for } k = 1, 2, \ldots\\
&\qquad R_k = Y_{k-1} - \Delta S_{k-1} \qquad \%\ \text{Dykstra's correction.}\\
&\qquad X_k = P_S(R_k)\\
&\qquad \Delta S_k = X_k - R_k\\
&\qquad Y_k = P_U(X_k)\\
&\text{end}
\end{aligned}
$$

- $X_k$ and $Y_k$ both converge to solution.
- $O(n^3)$ operations per step.
- Linear convergence.

## Property of Iterates

Assume now $W$ is diagonal and $a_{ii} \geq 1$, $i = 1\colon n$.

**Theorem 6** *If $A$ has $t$ nonpositive eigenvalues then $R_k$ has at least $t$ nonpositive eigenvalues and $X_k$ has at least $t$ zero eigenvalues, for all $k$.*

Note that

$$P_s(R_k) = (R_k)_+ = \sum_{\lambda_i > 0} \lambda_i q_i q_i^T.$$

So suffices to compute largest $n - t$ eigenpairs of $R_k$. Big saving if $t$ is large.

## Semidefinite Programming

The positive semidefinite program in primal standard form is

$$\begin{aligned} &\text{minimize } \langle C, X\rangle \\ &\text{subject to } \langle A_i, X\rangle = b_i, \quad i = 1\colon m, \\ &X = \operatorname{diag}(X_1, X_2, \dots, X_r) \ge 0 \end{aligned}$$

where $C$, $A_i$ are given $n \times n$ symmetric matrices.

Our problem: minimize

$$\|A - X\|_F^2 = \|A\|_F^2 + \langle X, X\rangle - 2\langle A, X\rangle$$

subject to $X$ a correlation matrix.

- Write as SDP, but have $O(n^4)$ constraints!
- Write with a quadratic cone constraint $\alpha \ge \|\operatorname{vec}(A - X)\|_2$, but still $O(n^2)$ variables.
- Johnson, Kroschel & Wolkowicz (1998): interior-point method for PSD completion problems. But $O(n^{12})$ flops, can't handle $x_{ii} \equiv 1$.
- Anjos, Higham & Wolkowicz: work in progress.

# Numerical Example 1

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

With tol $= 10^{-8}$, alg converges in 19 iterations to

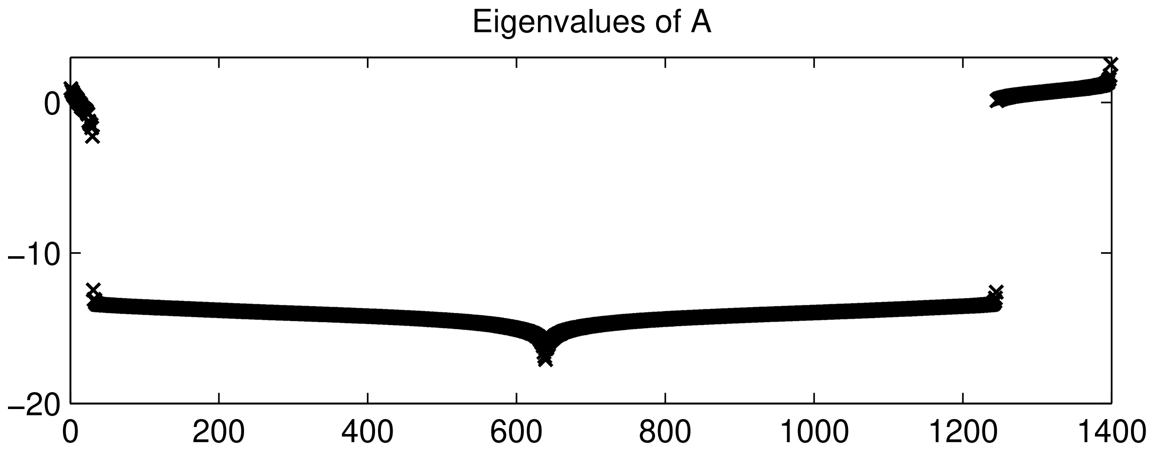
$$X = \begin{bmatrix} 1.0000 & -0.8084 & 0.1916 & 0.1068 \\ -0.8084 & 1.0000 & -0.6562 & 0.1916 \\ 0.1916 & -0.6562 & 1.0000 & -0.8084 \\ 0.1068 & 0.1916 & -0.8084 & 1.000 \end{bmatrix}.$$

$\|A - X\|_F = 2.13$ and $X$ has rank 3.

## Numerical Example 2, from Finance

Matrix $A$ of stock data, dimension 1399.
$a_{ii} \equiv 1$, $|a_{ij}| \leq 1$, but not psd.

$A$ highly rank deficient with 1245 nonpositive ei'vals
$\Rightarrow \operatorname{rank}(X) \leq 154$.




tol $= 10^{-4}$, since data accurate to 2–3 sig figs only.
67 iterations, $\|A - X\|_F = 20.96$.

On a 1Ghz Pentium III:

| Eigensolver | ei'pairs | time |
|---|---|---|
| `eig` | 1399 | 2hrs 45 mins |
| `dsyevr` (MEX) | 154 | 37 mins |

## Conclusions

- ★ Nearest correlation matrix needed in finance applications, where $n$ is large!
- ★ Modified alternating projections easy to implement, guaranteed to find global minimum, and can fully exploit low rank solutions.
- ★ Linearly convergent, but only $O(n^3)$ flops per iteration and $O(n^2)$ storage.
- ★ Semidefinite programming could potentially provide an alternative method with some advantages.
- ★ Now looking more closely at how $A$ is generated in the first place.