An Introduction to Correlation Stress Testing

Defeng Sun

Department of Mathematics and Risk Management Institute

National University of Singapore

This is based on a joint work with GAO Yan at NUS

March 05, 2010

These problems can be modeled in the following way

$$
\begin{aligned}
\min \quad & \|H \circ (X - G)\|_F \\
\text{s.t.} \quad & X_{ii} = 1, \ i = 1, \ldots, n \\
& X_{ij} = e_{ij}, \ (i,j) \in \mathcal{B}_e \,, \\
& X_{ij} \geq l_{ij}, \ (i,j) \in \mathcal{B}_l \,, \\
& X_{ij} \leq u_{ij}, \ (i,j) \in \mathcal{B}_u \,, \\
& X \in \mathcal{S}_+^n \,,
\end{aligned}
\tag{1}
$$

where $\mathcal{B}_e$, $\mathcal{B}_l$, and $\mathcal{B}_u$ are three index subsets of $\{(i,j) \,|\, 1 \leq i < j \leq n\}$ satisfying $\mathcal{B}_e \cap \mathcal{B}_l = \emptyset$, $\mathcal{B}_e \cap \mathcal{B}_u = \emptyset$, and $l_{ij} < u_{ij}$ for any $(i,j) \in \mathcal{B}_l \cap \mathcal{B}_u$.

Here $\mathcal{S}^n$ and $\mathcal{S}^n_+$ are, respectively, the space of $n \times n$ symmetric matrices and the cone of positive semidefinite matrices in $\mathcal{S}^n$.

$\|\cdot\|_F$ is the Frobenius norm defined in $\mathcal{S}^n$.

$H \geq 0$ is a weight matrix.

- H_{ij} is larger if G_{ij} is better estimated.

- $H_{ij} = 0$ if G_{ij} is missing.

A matrix $X \in \mathcal{S}^n_+$ is called a correlation matrix if $X \succeq 0$ (i.e., $X \in \mathcal{S}^n_+$) and $X_{ii} = 1$, $i = 1, \ldots, n$.

$$\begin{aligned}
\min \quad & \|H \circ (X - G)\|_F \\
\text{s.t.} \quad & X_{ii} = 1\,, \ i = 1, \ldots, n \\
& X \succeq 0\,,
\end{aligned} \qquad (2)$$

$$
\begin{aligned}
\min \quad & \|(X - G)\|_F \\
\text{s.t.} \quad & X_{ii} = 1 \,, \ i = 1, \ldots, n \\
& X \succeq 0 \,,
\end{aligned}
\tag{3}
$$

In finance and statistics, correlation matrices are in many situations found to be inconsistent, i.e., $X \nsucceq 0$.

These include, but are not limited to,

- Structured statistical estimations; data come from different time frequencies

- Stress testing regulated by Basel II;

- Expert opinions in reinsurance, and etc.

Partial market data[1]

$$G = \begin{bmatrix} 1.0000 & 0.9872 & 0.9485 & 0.9216 & -0.0485 & -0.0424 \\ 0.9872 & 1.0000 & 0.9551 & 0.9272 & -0.0754 & -0.0612 \\ 0.9485 & 0.9551 & 1.0000 & 0.9583 & -0.0688 & -0.0536 \\ 0.9216 & 0.9272 & 0.9583 & 1.0000 & -0.1354 & -0.1229 \\ -0.0485 & -0.0754 & -0.0688 & -0.1354 & 1.0000 & 0.9869 \\ -0.0424 & -0.0612 & -0.0536 & -0.1229 & 0.9869 & 1.0000 \end{bmatrix}$$

The eigenvalues of G are: $0.0087, 0.0162, 0.0347, 0.1000, 1.9669$, and 3.8736.

---

[1]RiskMetrics (`www.riskmetrics.com/stddownload_edu.html`)

Let's change G to

[change $G(1,6) = G(6,1)$ from -0.0424 to -0.1000]

$$\begin{bmatrix} 1.0000 & 0.9872 & 0.9485 & 0.9216 & -0.0485 & \mathbf{-0.1000} \\ 0.9872 & 1.0000 & 0.9551 & 0.9272 & -0.0754 & -0.0612 \\ 0.9485 & 0.9551 & 1.0000 & 0.9583 & -0.0688 & -0.0536 \\ 0.9216 & 0.9272 & 0.9583 & 1.0000 & -0.1354 & -0.1229 \\ -0.0485 & -0.0754 & -0.0688 & -0.1354 & 1.0000 & 0.9869 \\ \mathbf{-0.1000} & -0.0612 & -0.0536 & -0.1229 & 0.9869 & 1.0000 \end{bmatrix}$$

The eigenvalues of G are: -0.0216, 0.0305, 0.0441, 0.1078, 1.9609, and 3.8783.

On the other hand, some correlations may not be reliable or even missing:

$$G = \begin{bmatrix} 1.0000 & 0.9872 & 0.9485 & 0.9216 & -0.0485 & --- \\ 0.9872 & 1.0000 & 0.9551 & 0.9272 & -0.0754 & -0.0612 \\ 0.9485 & 0.9551 & 1.0000 & 0.9583 & -0.0688 & -0.0536 \\ 0.9216 & 0.9272 & 0.9583 & 1.0000 & -0.1354 & -0.1229 \\ -0.0485 & -0.0754 & -0.0688 & -0.1354 & 1.0000 & 0.9869 \\ --- & -0.0612 & -0.0536 & -0.1229 & 0.9869 & 1.0000 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Let us rewrite the problem:

$$\min \quad \frac{1}{2} \| H \circ (X - G) \|_F^2$$
$$\text{s.t.} \quad X_{ii} = 1, \ i = 1, \dots, n \tag{4}$$
$$X \succeq 0.$$

When $H = E$, the matrix of ones, we get

$$\min \quad \frac{1}{2} \| X - G \|_F^2$$
$$\text{s.t.} \quad X_{ii} = 1, \ i = 1, \dots, n \tag{5}$$
$$X \succeq 0.$$

which is known as the nearest correlation matrix (NCM) problem, a terminology coined by Nick Higham (2002).

The NCM problem is a special case of the best approximation problem

$$
\begin{aligned}
\min \quad & \frac{1}{2}\|x - c\|^2 \\
\text{s.t.} \quad & \mathcal{A}x \in b + Q\,, \\
& x \in K\,,
\end{aligned}
$$

where $\mathcal{X}$ is a real Euclidean space equipped with a scalar product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$

$\mathcal{A} : \mathcal{X} \to \Re^m$ is a bounded linear operator

$Q = \{0\}^p \times \Re^q_+$ is a polyhedral convex cone, $1 \leq p \leq m$, $q = m - p$, and K is a closed convex cone in $\mathcal{X}$.

The Karush-Kuhn-Tucker conditions are

$$
\begin{cases}
(x + z) - c - \mathcal{A}^* y = 0 \\[2em]
Q^* \ni y \perp \mathcal{A}x - b \in Q \\[2em]
K^* \ni z \perp x \in K \,,
\end{cases}
\quad,
$$

where "$\perp$" means the orthogonality. $Q^* = \Re^p \times \Re^q_+$ is the dual cone of Q and K^{*2} is the dual cone of K.

$^2 K^* := \{ d \in \mathcal{X} \mid \langle d, x \rangle \geq 0 \ \forall x \in K \}.$

Equivalently,

$$
\begin{cases}
(x + z) - c - \mathcal{A}^* y = 0 \\[1.5em]
Q^* \ni y \perp \mathcal{A}x - b \in Q \\[1.5em]
x - \Pi_K(x + z) = 0
\end{cases} ,
$$

where $\Pi_K(x)$ is the unique optimal solution to

$$
\begin{aligned}
\min \quad & \frac{1}{2}\|u - x\|^2 \\
\text{s.t.} \quad & u \in K .
\end{aligned}
$$

Consequently, by first eliminating $(x + z)$ and then x, we get

$$Q^* \ni y \perp \mathcal{A}\Pi_K(c + \mathcal{A}^*y) - b \in Q \,,$$

which is equivalent to

$$F(y) := y - \Pi_{Q^*}[y - (\mathcal{A}\Pi_K(c + \mathcal{A}^*y) - b)] = 0, \quad y \in \Re^m \,.$$

The dual formulation

The above is nothing but the first order optimality condition to the convex dual problem

$$
\max \quad -\theta(y) := -\left[\frac{1}{2}\|\Pi_K(c + \mathcal{A}^*y)\|^2 - \langle b, y\rangle - \frac{1}{2}\|c\|^2\right]
$$
$$
\text{s.t.} \quad y \in Q^*.
$$

Then F can be written as

$$
F(y) = y - \Pi_{Q^*}(y - \nabla\theta(y)).
$$

Now, we only need to solve

$$F(y) = 0, \quad y \in \Re^m \,.$$

However, the difficulties are:

- F is not differentiable at y;

- F involves two metric projection operators;

- Even if F is differentiable at y, it is too costly to compute $F'(y)$.

For the nearest correlation matrix problem,

- $\mathcal{A}(X) = \mathrm{diag}(X)$, a vector consisting of all diagonal entries of X.

.

- $\mathcal{A}^*(y) = \mathrm{diag}(y)$, the diagonal matrix.

- $b = e$, the vector of all ones in $\Re^n$ and $K = \mathcal{S}_+^n$.

Consequently, F can be written as

$$F(y) = \mathcal{A}\Pi_{\mathcal{S}_+^n}(G + \mathcal{A}^*y) - b.$$

For $n = 1$, we have

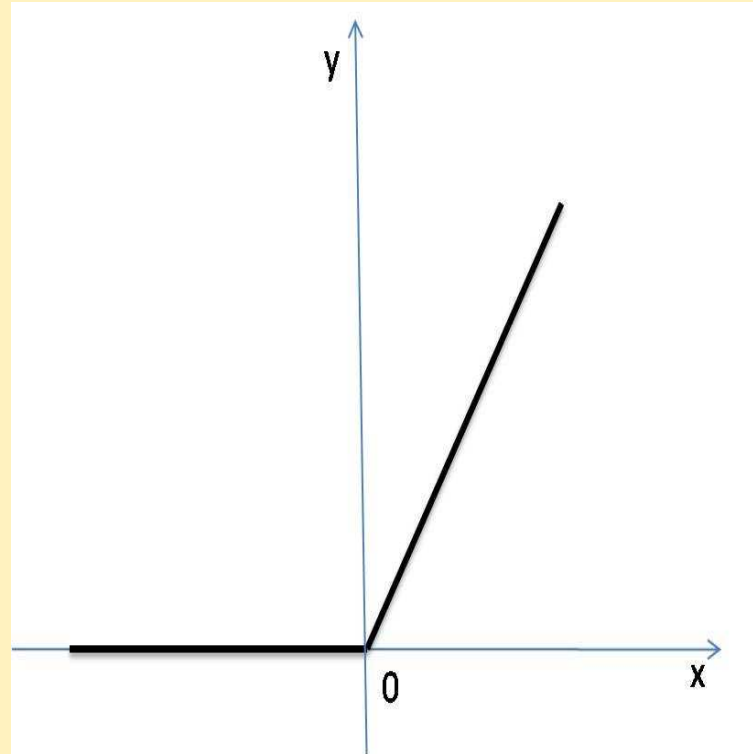$$x_+ := \Pi_{\mathcal{S}^1_+}(x) = \max(0, x).$$

Note that
- x_+ is only piecewise linear, but not smooth.
- $(x_+)^2$ is continuously differentiable with

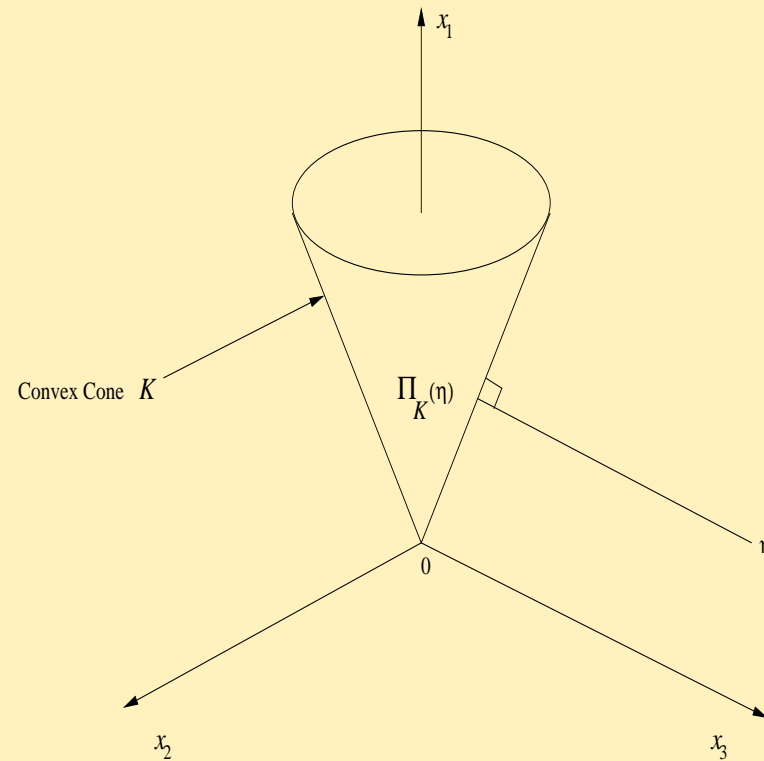$$\nabla \left\{ \frac{1}{2}(x_+)^2 \right\} = x_+,$$

but is not twice continuously differentiable.

The projector for $K = \mathcal{S}^n_+$:

Let $X \in \mathcal{S}^n$ have the following spectral decomposition

$$X = P\Lambda P^T,$$

where Λ is the diagonal matrix of eigenvalues of X and P is a corresponding orthogonal matrix of orthonormal eigenvectors.

Then

$$X_+ := P_{\mathcal{S}^n_+}(X) = P\Lambda_+ P^T.$$

We have

- $\|X_+\|^2$ is continuously differentiable with

$$\nabla\left(\frac{1}{2}\|X_+\|^2\right) = X_+,$$

but is not twice continuously differentiable.

- X_+ is not piecewise smooth, but strongly semismooth[3].

---

[3] D.F. Sun and J. Sun. Semismooth matrix valued functions. *Mathematics of Operations Research* 27 (2002) 150–169.

A quadratically convergent Newton's method is then designed by Qi and Sun[4] The written code is called CorNewton.m.

"This piece of research work is simply great and practical. I enjoyed reading your paper." – March 20, 2007, a home loan financial institution based in McLean, VA.

"It's very impressive work and I've also run the Matlab code found in Defeng's home page. It works very well."– August 31, 2007, a major investment bank based in New York city.

---

[4]H.D. QI AND D.F. SUN. A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications* 28 (2006) 360–385.

If we have lower and upper bounds on X, F takes the form

$$F(y) = y - \Pi_{Q^*} \left[y - \left(\mathcal{A} \Pi_{\mathcal{S}^n_+} (G + \mathcal{A}^* y) - b \right) \right],$$

which involves double layered projections over convex cones.

A quadratically convergent smoothing Newton method is designed by Gao and Sun[5].

Again, highly efficient.

---

[5]Y. GAO AND D.F. SUN. Calibrating least squares covariance matrix problems with equality and inequality constraints, SIAM Journal on Matrix Analysis and Applications 31 (2009), 1432–1457.

$$\min \quad \frac{1}{2}\|H \circ (X - G)\|_F^2$$

$$\text{s.t.} \quad \mathcal{A}(X) \in b + Q\,,$$

$$X \in \mathcal{S}_+^n\,,$$

Let $d \in \Re^n$ be a positive vector such that

$$H \circ H \leq dd^T.$$

For example, $d = \max(H_{ij})e$. Let $D^{1/2} = \mathrm{diag}(d_1^{0.5}, \ldots, d_n^{0.5})$.
Let

$$f(X) := \frac{1}{2}\|H \circ (X - G)\|_F^2.$$

Then g is majorized by

$$f^k(X) := f(X^k) + \langle H \circ H(X^k - G), X - X^k \rangle + \frac{1}{2}\|D^{1/2}(X - X^k)D^{1/2}\|_F^2,$$

i.e.,

$$f(X^k) = f^k(X^k) \quad \text{and} \quad f(X) \leq f^k(X).$$

The idea of the majorization is to solve, for given X^k, the following problem

$$
\begin{aligned}
\min \quad & f^k(X) \\
\text{s.t.} \quad & \mathcal{A}(X) \in b + Q\,, \\
& X \in \mathcal{S}_+^n\,,
\end{aligned}
$$

which is a diagonal weighted least squares correlation matrix problem

$$
\begin{aligned}
\min \quad & \frac{1}{2}\|D^{1/2}(X - \overline{X}^k)D^{1/2}\|_F^2 \\
\text{s.t.} \quad & \mathcal{A}(X) \in b + Q\,, \\
& X \in \mathcal{S}_+^n\,.
\end{aligned}
$$

Now, we can use the two Newton methods introduced earlier for the majorized subproblems!

$$f(X^{k+1}) < f(X^k) < \cdots < f(X^1).$$

$$G = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 0.5 \\ -1 & 1 & 0.5 & 1 \end{bmatrix}$$

Suppose that $G(1,2)$ and $G(2,1)$ are missing.

$$G = \begin{bmatrix} 1 & * & 1 & -1 \\ * & 1 & -1 & 1 \\ 1 & -1 & 1 & 0.5 \\ -1 & 1 & 0.5 & 1 \end{bmatrix}$$

We take

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

After 4 iterations, we get

$$X^* = \begin{bmatrix} 1.000 & -1.000 & 0.6894 & -0.6894 \\ -1.000 & 1.000 & -0.6894 & 0.6894 \\ 0.6894 & -0.6894 & 1.000 & 0.0495 \\ -0.6894 & 0.6894 & 0.0495 & 1.000 \end{bmatrix}$$

This is the same solution as the case with no-missing data.

Example 1	CorrMajor		AugLag	
n	*time*	*residue*	*time*	*residue*
100	0.9	2.9006e1	1.1	2.9006e1
200	1.8	6.6451e1	3.2	6.6451e1
500	9.7	1.8815e2	23.5	1.8815e2
1000	51.3	4.0108e2	223.4	4.0108e2

Table 1: Numerical results for Example 1

• A code named CorrMajor.m can efficiently solve correlation matrix problems with all sorts of bound constraints.

• The techniques may be used to solve many other problems, e.g., low rank matrix problems with sparsity.

• The limitation is that it cannot solve problems for matrices exceeding the dimension $5,000$ by $5,000$ on a PC due to memory constraints.

Thank you! :)