A Majorized Penalty Approach for Calibrating Rank Constrained Correlation Matrix Problems

Defeng Sun

Department of Mathematics and Risk Management Institute
National University of Singapore

This is a joint work with Yan Gao at NUS

July 19, 2010

$$\min \frac{1}{2} \|X - Z\|_F^2$$

s.t. \(\text{rank}(X) \leq r\)

admits an analytic solution for a given \(Z \in \mathbb{R}^{m \times n} (m \leq n \text{ without loss of generality})\):

$$X^* = \sum_{i=1}^{r} \sigma_i(Z) u_i v_i^T,$$

where \(Z\) has the following singular value decomposition (SVD):

$$Z = U[\text{diag}(\sigma(Z)) \ 0]V^T, \quad \sigma_1(Z) \geq \sigma_2(Z) \geq \ldots \geq \sigma_m(Z) \geq 0.$$
On October 27, 2009, I received this from Universiteit van Tilburg:

My thesis is about correlations in a pension fund pooling. It is important for economic capital calculations. For some risks such as operational risk, I don’t have data and hence I need to consult for an expert opinion. Then I might end up with not PSD matrices. Therefore, I need to calculate nearest correlation matrix.

In my given correlation matrix, I want to fix the correlations, which are data driven and I want the rest of the correlations not smaller than 0.1 from original matrix.

Your code is very convenient for my study. However, ...
On November 3, 2009:

Thank you for your valuable time, comments and helping me about solving my problem.

I gave no chance that my fixed constraints could be non-PSD before. Your advice solves the problem. I will modify my study in the light of it.
In this talk, we are interested in the following rank constrained covariance matrix problem

$$\begin{align*}
\min & \quad \| H \circ (X - G) \|_F^2 \\
\text{s.t.} & \quad X_{ii} = 1, \ i = 1, \ldots, n \\
& \quad X_{ij} = e_{ij}, \ (i, j) \in \mathcal{B}_e, \\
& \quad X_{ij} \geq l_{ij}, \ (i, j) \in \mathcal{B}_l, \\
& \quad X_{ij} \leq u_{ij}, \ (i, j) \in \mathcal{B}_u, \\
& \quad X \in \mathcal{S}_+^n, \\
& \quad \text{rank}(X) \leq r, 
\end{align*}$$

(1)

where $\mathcal{B}_e$, $\mathcal{B}_l$, and $\mathcal{B}_u$ are three index subsets of $\{(i, j) \mid 1 \leq i < j \leq n\}$ 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 $S^n$ and $S^+_n$ are, respectively, the space of $n \times n$ symmetric matrices and the cone of positive semidefinite matrices in $S^n$.

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

$H \geq 0$ is a weight matrix.

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

A matrix $X \in S^n$ is called a correlation matrix if $X \succeq 0$ (i.e., $X \in S^+_n$) and $X_{ii} = 1$, $i = 1, \ldots, n$. 
\[
\begin{align*}
\text{min} & \quad \|H \odot (X - G)\|_F^2 \\
\text{s.t.} & \quad X_{ii} = 1, \ i = 1, \ldots, n \\
& \quad X \succeq 0, \\
& \quad \text{rank}(X) \leq r.
\end{align*}
\]
The simplest corr. matrix model

\[
\begin{align*}
\min & \quad \|(X - G)\|_F^2 \\
\text{s.t.} & \quad X_{ii} = 1, \; i = 1, \ldots, n \\
& \quad X \succeq 0, \\
& \quad \text{rank}(X) \leq r.
\end{align*}
\]
In finance and statistics, correlation matrices are in many situations found to be inconsistent, i.e., $X \not\succeq 0$.

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

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

The eigenvalues of \( G \) are: 0.0087, 0.0162, 0.0347, 0.1000, 1.9669, and 3.8736.

\(^1\text{RiskMetrics (www.riskmetrics.com/stddownload_edu.html)}\)
Let's change $G$ to

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

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

The eigenvalues of $G$ are: $-0.0216, 0.0305, 0.0441, 0.1078, 1.9609, \text{ and } 3.8783.$
On the other hand, some correlations may not be reliable or even missing:

$$G = \begin{bmatrix}
1.0000 & 0.9872 & 0.9485 & 0.9216 & -0.0485 & - & - & - \\
0.9872 & 1.0000 & 0.9551 & 0.9272 & -0.0754 & -0.0612 & - & - & - \\
0.9485 & 0.9551 & 1.0000 & 0.9583 & -0.0688 & -0.0536 & - & - & - \\
0.9216 & 0.9272 & 0.9583 & 1.0000 & -0.1354 & -0.1229 & - & - & - \\
-0.0485 & -0.0754 & -0.0688 & -0.1354 & 1.0000 & 0.9869 & - & - & - \\
- & - & - & -0.0612 & -0.0536 & -0.1229 & 0.9869 & 1.0000 & - \\
- & - & - & - & - & - & - & - & -
\end{bmatrix}$$
Let us first consider the problem without the rank constraint:

\[
\begin{align*}
\text{min } & \quad \frac{1}{2} \| H \circ (X - G) \|_F^2 \\
\text{s.t. } & \quad X_{ii} = 1, \ i = 1, \ldots, n \\
& \quad X \succeq 0.
\end{align*}
\] (4)

When \( H = E \), the matrix of ones, we get

\[
\begin{align*}
\text{min } & \quad \frac{1}{2} \| X - G \|_F^2 \\
\text{s.t. } & \quad X_{ii} = 1, \ i = 1, \ldots, n \\
& \quad X \succeq 0.
\end{align*}
\] (5)

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

$$\min \quad \frac{1}{2}\|x - c\|^2$$

s.t. \quad Ax \in b + Q, \quad x \in K,$$

where \( \mathcal{X} \) is a real Hilbert space equipped with a scalar product \( \langle \cdot, \cdot \rangle \) and its induced norm \( \| \cdot \| \), \( A : \mathcal{X} \to \mathbb{R}^m \) is a bounded linear operator, \( Q = \{0\}^p \times \mathbb{R}_+^q \) is a polyhedral convex cone, \( 1 \leq p \leq m, \quad q = m - p \), and \( K \) is a closed convex cone in \( \mathcal{X} \).
The Karush-Kuhn-Tucker conditions are

\[
\begin{align*}
\left\{ \begin{align*}
(x - z) - c - A^*y &= 0 \\
Q^* &\ni y \perp Ax - b \in Q, \\
K^* &\ni z \perp x \in K,
\end{align*} \right. 
\end{align*}
\]

where “\(\perp\)” means the orthogonality. \(Q^*\) is the dual cone of \(Q\) and \(K^*\) is the dual cone of \(K\).
Equivalently,

\[
\begin{cases}
(x - z) - c - A^*y = 0 \\
Q^* \ni y \perp Ax - b \in Q \\
x - \Pi_K(x - z) = 0
\end{cases}
\]

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

\[
\min \frac{1}{2} \|u - x\|^2 \\
s.t. \quad u \in K
\]
Consequently, by first eliminating \((x - z)\) and then \(x\), we get

\[ Q^* \ni y \perp A \Pi_K (c + A^* y) - b \in Q , \]

which is equivalent to

\[ F(y) := y - \Pi_{Q^*} [y - (A \Pi_K (c + A^* y) - b)] = 0, \quad y \in \mathbb{R}^m . \]
The above is nothing but the first order optimality condition to the convex dual problem

\[
\max -\theta(y) := - \left[ \frac{1}{2} \| \Pi_K(c + A^* y) \|^2 - \langle b, y \rangle - \frac{1}{2} \| c \|^2 \right]
\]

s.t. \( y \in Q^* \).

Then \( F \) can be written as

\[
F(y) = y - \Pi_{Q^*} (y - \nabla \theta(y)) .
\]
Now, we only need to solve

\[ F(y) = 0, \quad y \in \mathbb{R}^m. \]

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,

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

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

- \( b = e \), the vector of all ones in \( \mathbb{R}^n \) and \( K = S_+^n \).

Consequently, \( F \) can be written as

\[
F(y) = \mathcal{A} \Pi S_+^n (G + \mathcal{A}^* y) - b.
\]
For $n = 1$, we have

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

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 one dimensional case
The projector for $K = S^n_+$:
Let $X \in \mathcal{S}^n$ have the following spectral decomposition

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

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

Then

$$X_+ := \Pi \mathcal{S}^n_+ (X) = P \Lambda_+ P^T.$$
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\textsuperscript{2}.

A quadratically convergent Newton’s method is then designed by Qi and Sun\(^3\). 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.

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

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

which involves double layered projections over convex cones.

A quadratically convergent inexact smoothing Newton-BICGStab method is designed by Gao and Sun$^4$.

Again, highly efficient.

Back to the rank constraint

\[
\begin{align*}
\min & \quad \frac{1}{2} \| H \circ (X - G) \|_F^2 \\
\text{s.t.} & \quad AX \in b + Q, \\
& \quad X \in \mathcal{S}^n_+, \\
& \quad \text{rank}(X) \leq k,
\end{align*}
\]

equivalently,

\[
\begin{align*}
\min & \quad \frac{1}{2} \| H \circ (X - G) \|_F^2 \\
\text{s.t.} & \quad AX \in b + Q, \\
& \quad X \in \mathcal{S}^n_+, \\
& \quad \lambda_i(X) = 0, \ i = k + 1, \ldots, n.
\end{align*}
\]
Given $c > 0$, we consider a penalized version

$$
\min \frac{1}{2} \| H \circ (X - G) \|_F^2 + c \sum_{i=k+1}^{n} \lambda_i(X)
$$

s.t. \quad AX \in \mathcal{B} + Q, \quad X \in \mathcal{S}^n_+,

or equivalently

$$
\min \ f_c(X) := \frac{1}{2} \| H \circ (X - G) \|_F^2 + c\langle I, X \rangle - c \sum_{i=1}^{k} \lambda_i(X)
$$

s.t. \quad AX \in \mathcal{B} + Q, \quad X \in \mathcal{S}^n_+.$$

The penalty approach
Let $h(X) := \sum_{i=1}^{k} \lambda_i(X) − \langle I, X \rangle$. Since $h$ is a convex function, for given $X^k$, we have

$$h(X) \geq h^k(X) := h(X^k) + \langle V^k, X - X^k \rangle,$$

where $V^k \in \partial h(X^k)$. Thus, $-h$ is majorized by $-h^k$.

Let $d \in \mathbb{R}^n$ be a positive vector such that

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

For example, $d = \max(H_{ij})e$. Let $D^{1/2} = \text{diag}(d_1^{0.5}, \ldots, d_n^{0.5})$. 
Let
\[ g(X) := \frac{1}{2} \| H \circ (X - G) \|_F^2. \]

Then \( g \) is majorized by
\[ g^k(X) := g(X^k) + \langle H \circ H(X^k - G), X - X^k \rangle + \frac{1}{2} \| D^{1/2}(X - X^k) D^{1/2} \|_F^2. \]

Thus, at \( X^k \), \( f_c \) is majorized by
\[ f_c(X) \leq f^k_c(X) := g^k(X) - c h^k(X) \]
and \( f_c(X^k) = f^k_c(X^k) \).
Instead of solving the penalized problem, the idea of the majorization is to solve, for given $X^k$, the following problem

\[
\min f_c^k(X) = g^k(X) - ch^k(X)
\]

s.t. $AX \in b + Q$, 

$X \in S^n_+$, 

which is a diagonal weighted least squares correlation matrix problem

\[
\min \frac{1}{2} \| D^{1/2}(X - X^k)D^{1/2} \|^2_F \\
\text{s.t. } AX \in b + Q, \\
X \in S^n_+.
\]
Now, we can use the two Newton methods introduced earlier for the majorized subproblems!

\[ f_c(X^{k+1}) < f_c(X^k) < \cdots < f_c(X^1). \]
Where is the rank condition?

Looks good? But how can one guarantee that we can get a final $X^*$ such that its rank is less or equal to $k$?

The answer is: increase $c$. That is, to have a sequence of $\{c_k\}$ with $c_{k+1} \geq c_k$.

Will it work? Numerical stability? Does not need a large $c_k$ in numerical computations.

There are no known methods that can solve the general rank constrained problem. For the $H$-normed correlation matrix problems (without constraints on the off diagonal entries), the major.m of R. Pietersz and J.F. Groenen (2004) is the most efficient one so far [write $X = YY^T$ for $Y \in \mathbb{R}^{n \times k}$ and apply component-by-component majorization.].
Let $Y \in S^n$ be arbitrarily chosen. Suppose that $Y$ has the spectral decomposition

$$Y = U \Sigma(Y) U^T,$$

(6)

where $U \in O_n$ is a corresponding orthogonal matrix of orthonormal eigenvectors of $Y$ and $\Sigma(Y) := \text{diag}(\sigma(Y))$ where $\sigma(Y) = (\sigma_1(Y), \ldots, \sigma_n(Y))^T$ is the column vector containing all the eigenvalues of $Y$ being arranged in the non-increasing order in terms of their absolute values, i.e.,

$$|\sigma_1(Y)| \geq \cdots \geq |\sigma_n(Y)|,$$

and whenever the equality holds, the larger one comes first, i.e.,

if $|\sigma_i(Y)| = |\sigma_j(Y)|$ and $\sigma_i(Y) > \sigma_j(Y)$, then $i < j$. 

Define

$$\bar{\alpha} := \{ i \mid |\sigma_i(Y)| > |\sigma_r(Y)| \}$$,
$$\bar{\beta} := \{ i \mid |\sigma_i(Y)| = |\sigma_r(Y)| \}$$,
$$\bar{\gamma} := \{ i \mid |\sigma_i(Y)| < |\sigma_r(Y)| \}$$,

and

$$\bar{\beta}^+ := \{ i \mid \sigma_i(Y) = |\sigma_r(Y)| \}$$,
$$\bar{\beta}^- := \{ i \mid \sigma_i(Y) = -|\sigma_r(Y)| \}$$.

Denote

$$\Psi_r(Y) := \min_{Z} \frac{1}{2} \| Z - Y \|^2$$

s.t. $Z \in S^n_r(Y)$. \hspace{1cm} (7)

Denote the set of optimal solutions to (7) by $\Pi_{S^n_r(Y)}(Y)$. 
Lemma 1. Let $Y \in S^n$ have the spectral decomposition as in (6). Then the solution set $\Pi_{S^n(r)}(Y)$ to problem (7) can be characterized as follows

$$\Pi_{S^n(r)}(Y) = \left\{ \begin{bmatrix} U_{\bar{\alpha}} & U_{\bar{\beta}} \bar{Q}_{\bar{\beta}} & U_{\bar{\gamma}} \end{bmatrix} \text{diag}(v) \begin{bmatrix} U_{\bar{\alpha}} & U_{\bar{\beta}} \bar{Q}_{\bar{\beta}} & U_{\bar{\gamma}} \end{bmatrix}^T \right|$$

$$v \in \mathcal{V}, Q_{\bar{\beta}} = \begin{bmatrix} Q_{\bar{\beta}^+} & 0 \\ 0 & Q_{\bar{\beta}^-} \end{bmatrix}, Q_{\bar{\beta}^+} \in \mathcal{O}_{|\bar{\beta}^+|}, Q_{\bar{\beta}^-} \in \mathcal{O}_{|\bar{\beta}^-|}$$

where

$$\mathcal{V} := \left\{ v \in \mathbb{R}^n \mid v_i = \sigma_i(Y) \text{ for } i \in \bar{\alpha} \cup \bar{\beta}_1, \ v_i = 0 \text{ for } i \in (\bar{\beta} \setminus \bar{\beta}_1) \cup \bar{\gamma}, \right.$$  

where $\bar{\beta}_1 \subseteq \bar{\beta}$ and $|\bar{\beta}_1| = r - |\bar{\alpha}|$.

(8)
Theorem 1. The optimal solution \((\bar{y}, \overline{Y}) \in Q^* \times S^n\) to the the dual problem satisfies

\[
b - A\Pi_{S^n_+}(C + A^*\bar{y} + \overline{Y}) \in N_{Q^*}(\bar{y})
\]

and

\[
\Pi_{S^n_+}(C + A^*\bar{y} + \overline{Y}) \in \text{conv} \left\{ \Pi_{S^n(r)}(C - \overline{Y}) \right\},
\]

where \(\Pi_{S^n(r)}(\cdot)\) is defined as in Lemma 1. Furthermore, if there exists a matrix \(\overline{X} \in \Pi_{S^n(r)}(C - \overline{Y})\) such that \(\overline{X} = \Pi_{S^n_+}(C + A^*\bar{y} + \overline{Y})\), then \(\overline{X}\) and \((\bar{y}, \overline{Y})\) globally solve the primal problem with \(H = E\) and the corresponding dual problem, respectively and there is no duality gap between the primal and dual problems.

---

\(^5\)Y. Gao and D.F. Sun, A majorized penalty approach for calibrating rank constrained correlation matrix problems, manuscript, March 2010.
The testing examples to be reported are given below.

**Example 1.** Let $n = 500$ and the weight matrix $H = E$. For $i, j = 1, \ldots, n$, $C_{ij} = 0.5 + 0.5e^{-0.05|i-j|}$. The index sets are $B_e = B_l = B_u = \emptyset$.

**Example 2.** Let $n = 500$ and the weight matrix $H = E$. The matrix $C$ is extracted from the correlation matrix which is based on a 10,000 gene micro-array data set obtained from 256 drugs treated rat livers. The index sets are $B_e = B_l = B_u = \emptyset$.

**Example 3.** Let $n = 500$. The matrix $C$ is the same as in Example 1, i.e., $C = 0.5 + 0.5e^{-0.05|i-j|}$ for $i, j = 1, \ldots, n$. The index sets are $B_e = B_l = B_u = \emptyset$. The weight matrix $H$ is generated in the way such that all its entries are uniformly distributed in $[0.1, 10]$ except for $2 \times 100$ entries in $[0.01, 100]$. 
Example 4. Let $n = 500$. The matrix $C$ is the same as in Example 2. The index sets are $B_e = B_l = B_u = \emptyset$. The weight matrix $H$ is generated in the same way as in Example 3.

Example 5. The matrix $C$ is obtained from the gene data sets with dimension $n = 1,000$ as in Example 2. The weight matrix $H$ is the same as in Example 3. The index sets $B_e, B_l, \text{ and } B_u \subset \{(i, j) \mid 1 \leq i < j \leq n\}$ consist of the indices of $\min(\hat{n}_r, n - i)$ randomly generated elements at the $i$th row of $X, i = 1, \ldots, n$ with $\hat{n}_r = 5$ for $B_e$ and $\hat{n}_r = 10$ for $B_l$ and $B_u$. We take $e_{ij} = 0$ for $(i, j) \in B_e$, $l_{ij} = -0.1$ for $(i, j) \in B_l$ and $u_{ij} = 0.1$ for $(i, j) \in B_u$. 
Example 5.1: n=500, H=E

PenCorr | Major | SemiNewton | Dual-BFGS

- [Graph showing numerical results for different algorithms with respect to rank and time (secs) for Example 5.1: n=500, H=E.]

- [Graph showing relative gap for different algorithms with respect to rank for Example 5.1: n=500, H=E.]
<table>
<thead>
<tr>
<th>Eg1</th>
<th>Major</th>
<th>SemiNewton</th>
<th>Dual-BFGS</th>
<th>PenCorr</th>
</tr>
</thead>
<tbody>
<tr>
<td>rank</td>
<td>time</td>
<td>residue</td>
<td>relgap</td>
<td>time</td>
</tr>
<tr>
<td>2</td>
<td>1.9</td>
<td>1.564e2</td>
<td>3.4e-3</td>
<td>63.0</td>
</tr>
<tr>
<td>5</td>
<td>2.2</td>
<td>7.883e1</td>
<td>6.5e-5</td>
<td>23.5</td>
</tr>
<tr>
<td>10</td>
<td>2.7</td>
<td>3.869e1</td>
<td>6.9e-5</td>
<td>19.0</td>
</tr>
<tr>
<td>15</td>
<td>4.2</td>
<td>2.325e1</td>
<td>8.3e-5</td>
<td>18.5</td>
</tr>
<tr>
<td>20</td>
<td>7.5</td>
<td>1.571e1</td>
<td>8.8e-5</td>
<td>15.3</td>
</tr>
<tr>
<td>25</td>
<td>12.8</td>
<td>1.145e1</td>
<td>1.1e-4</td>
<td>14.4</td>
</tr>
<tr>
<td>30</td>
<td>19.4</td>
<td>8.797e0</td>
<td>1.3e-4</td>
<td>14.0</td>
</tr>
<tr>
<td>35</td>
<td>34.4</td>
<td>7.020e0</td>
<td>1.7e-4</td>
<td>14.0</td>
</tr>
<tr>
<td>40</td>
<td>43.4</td>
<td>5.766e0</td>
<td>2.2e-4</td>
<td>1.3</td>
</tr>
<tr>
<td>45</td>
<td>63.6</td>
<td>4.843e0</td>
<td>3.0e-4</td>
<td>1.3</td>
</tr>
<tr>
<td>50</td>
<td>80.1</td>
<td>4.141e0</td>
<td>4.0e-4</td>
<td>1.4</td>
</tr>
<tr>
<td>60</td>
<td>145.0</td>
<td>3.156e0</td>
<td>6.7e-4</td>
<td>1.4</td>
</tr>
<tr>
<td>70</td>
<td>243.0</td>
<td>2.507e0</td>
<td>1.1e-3</td>
<td>1.4</td>
</tr>
<tr>
<td>80</td>
<td>333.0</td>
<td>2.053e0</td>
<td>1.6e-3</td>
<td>1.5</td>
</tr>
<tr>
<td>90</td>
<td>452.0</td>
<td>1.722e0</td>
<td>2.4e-3</td>
<td>1.6</td>
</tr>
<tr>
<td>100</td>
<td>620.0</td>
<td>1.471e0</td>
<td>3.3e-3</td>
<td>1.5</td>
</tr>
<tr>
<td>125</td>
<td>1180.0</td>
<td>1.055e0</td>
<td>6.8e-3</td>
<td>1.7</td>
</tr>
</tbody>
</table>

**Table 1: Numerical results for Example 1**
Example 5.2: $n=500$, $H=E$

- **PenCorr**
- **Major**
- **SemiNewton**
- **Dual-BFGS**
<table>
<thead>
<tr>
<th>rank</th>
<th>time</th>
<th>residue</th>
<th>relgap</th>
<th>time</th>
<th>residue</th>
<th>relgap</th>
<th>time</th>
<th>residue</th>
<th>relgap</th>
<th>time</th>
<th>residue</th>
<th>relgap</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.6</td>
<td>2.858e2</td>
<td>6.5e-4</td>
<td>54.4</td>
<td>2.860e2</td>
<td>1.5e-3</td>
<td>304.5</td>
<td>2.862e2</td>
<td>2.1e-3</td>
<td>37.2</td>
<td>2.859e2</td>
<td>8.2e-4</td>
</tr>
<tr>
<td>5</td>
<td>6.0</td>
<td>1.350e2</td>
<td>2.0e-3</td>
<td>38.2</td>
<td>1.358e2</td>
<td>8.1e-3</td>
<td>78.8</td>
<td>1.367e2</td>
<td>1.5e-2</td>
<td>99.2</td>
<td>1.351e2</td>
<td>2.4e-3</td>
</tr>
<tr>
<td>10</td>
<td>9.3</td>
<td>6.716e1</td>
<td>4.4e-4</td>
<td>32.7</td>
<td>6.735e1</td>
<td>3.2e-3</td>
<td>58.3</td>
<td>6.802e1</td>
<td>1.3e-2</td>
<td>32.1</td>
<td>6.719e1</td>
<td>9.7e-4</td>
</tr>
<tr>
<td>15</td>
<td>8.8</td>
<td>4.097e1</td>
<td>3.4e-4</td>
<td>26.8</td>
<td>4.100e1</td>
<td>1.0e-3</td>
<td>44.6</td>
<td>4.096e1</td>
<td>1.0e-4</td>
<td>18.4</td>
<td>4.099e1</td>
<td>7.5e-4</td>
</tr>
<tr>
<td>20</td>
<td>13.0</td>
<td>2.842e1</td>
<td>7.3e-4</td>
<td>18.8</td>
<td>2.844e1</td>
<td>1.4e-3</td>
<td>40.4</td>
<td>2.842e1</td>
<td>8.9e-4</td>
<td>16.6</td>
<td>2.843e1</td>
<td>1.1e-3</td>
</tr>
<tr>
<td>25</td>
<td>34.9</td>
<td>2.149e1</td>
<td>1.2e-3</td>
<td>18.0</td>
<td>2.152e1</td>
<td>2.6e-3</td>
<td>26.6</td>
<td>2.149e1</td>
<td>1.2e-3</td>
<td>16.4</td>
<td>2.151e1</td>
<td>2.2e-3</td>
</tr>
<tr>
<td>30</td>
<td>33.7</td>
<td>1.693e1</td>
<td>4.3e-4</td>
<td>17.3</td>
<td>1.695e1</td>
<td>1.7e-3</td>
<td>23.0</td>
<td>1.694e1</td>
<td>7.8e-4</td>
<td>14.5</td>
<td>1.694e1</td>
<td>1.2e-3</td>
</tr>
<tr>
<td>35</td>
<td>71.8</td>
<td>1.379e1</td>
<td>1.3e-3</td>
<td>18.1</td>
<td>1.381e1</td>
<td>2.6e-3</td>
<td>19.7</td>
<td>1.378e1</td>
<td>7.1e-4</td>
<td>11.9</td>
<td>1.379e1</td>
<td>1.6e-3</td>
</tr>
<tr>
<td>40</td>
<td>50.0</td>
<td>1.151e1</td>
<td>1.5e-3</td>
<td>12.5</td>
<td>1.152e1</td>
<td>2.1e-3</td>
<td>34.7</td>
<td>1.145e1</td>
<td>3.2e-4</td>
<td>7.7</td>
<td>1.151e1</td>
<td>1.6e-3</td>
</tr>
<tr>
<td>45</td>
<td>43.3</td>
<td>9.733e0</td>
<td>9.6e-4</td>
<td>10.6</td>
<td>9.736e0</td>
<td>1.3e-3</td>
<td>23.1</td>
<td>9.733e0</td>
<td>9.2e-4</td>
<td>6.3</td>
<td>9.733e0</td>
<td>1.0e-3</td>
</tr>
<tr>
<td>50</td>
<td>44.5</td>
<td>8.318e0</td>
<td>4.1e-4</td>
<td>10.7</td>
<td>8.319e0</td>
<td>4.8e-3</td>
<td>19.7</td>
<td>8.315e0</td>
<td>5.1e-6</td>
<td>5.7</td>
<td>8.318e0</td>
<td>4.5e-3</td>
</tr>
<tr>
<td>60</td>
<td>66.5</td>
<td>6.214e0</td>
<td>8.1e-4</td>
<td>10.9</td>
<td>6.214e0</td>
<td>7.4e-4</td>
<td>6.1</td>
<td>6.209e0</td>
<td>1.4e-13</td>
<td>6.9</td>
<td>6.213e0</td>
<td>5.9e-4</td>
</tr>
<tr>
<td>70</td>
<td>91.2</td>
<td>4.733e0</td>
<td>1.1e-3</td>
<td>11.0</td>
<td>4.731e0</td>
<td>8.2e-4</td>
<td>23.1</td>
<td>4.728e0</td>
<td>1.9e-4</td>
<td>4.6</td>
<td>4.731e0</td>
<td>7.2e-4</td>
</tr>
<tr>
<td>80</td>
<td>93.0</td>
<td>3.663e0</td>
<td>8.7e-4</td>
<td>2.2</td>
<td>3.800e0</td>
<td>3.8e-2</td>
<td>5.2</td>
<td>3.660e0</td>
<td>4.0e-13</td>
<td>2.9</td>
<td>3.662e0</td>
<td>4.5e-4</td>
</tr>
<tr>
<td>90</td>
<td>125.0</td>
<td>2.865e0</td>
<td>1.2e-3</td>
<td>2.0</td>
<td>2.962e0</td>
<td>3.5e-2</td>
<td>5.0</td>
<td>2.862e0</td>
<td>5.1e-13</td>
<td>3.0</td>
<td>2.864e0</td>
<td>7.0e-4</td>
</tr>
<tr>
<td>100</td>
<td>150.0</td>
<td>2.255e0</td>
<td>1.4e-3</td>
<td>1.7</td>
<td>2.323e0</td>
<td>3.2e-2</td>
<td>15.1</td>
<td>2.254e0</td>
<td>7.8e-4</td>
<td>2.9</td>
<td>2.254e0</td>
<td>8.3e-4</td>
</tr>
<tr>
<td>125</td>
<td>288.6</td>
<td>1.269e0</td>
<td>2.4e-3</td>
<td>1.4</td>
<td>1.304e0</td>
<td>3.0e-2</td>
<td>17.1</td>
<td>1.266e0</td>
<td>1.6e-4</td>
<td>2.7</td>
<td>1.268e0</td>
<td>1.4e-3</td>
</tr>
</tbody>
</table>

Table 2: Numerical results for Example 2
<table>
<thead>
<tr>
<th>rank</th>
<th>Example 3</th>
<th></th>
<th>Example 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Majorw</td>
<td></td>
<td>PenCorr</td>
<td></td>
</tr>
<tr>
<td></td>
<td>time residue</td>
<td></td>
<td>time residue</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8.8 1.805e2</td>
<td></td>
<td>2.9 3.274e2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>27.0 8.984e1</td>
<td></td>
<td>34.4 1.523e2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>38.7 4.382e1</td>
<td></td>
<td>48.5 7.423e1</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>55.5 2.616e1</td>
<td></td>
<td>70.5 4.442e1</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>84.4 1.751e1</td>
<td></td>
<td>101.4 2.985e1</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>117.0 1.265e1</td>
<td></td>
<td>289.6 2.197e1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>171.8 9.657e0</td>
<td></td>
<td>335.6 1.694e1</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>250.6 7.639e0</td>
<td></td>
<td>436.7 1.345e1</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>324.7 6.213e0</td>
<td></td>
<td>470.7 1.098e1</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>408.4 5.169e0</td>
<td></td>
<td>498.7 9.104e0</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>502.2 4.391e0</td>
<td></td>
<td>639.5 7.625e0</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>654.1 3.290e0</td>
<td></td>
<td>837.6 5.552e0</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>972.5 2.579e0</td>
<td></td>
<td>987.5 4.135e0</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1274.9 2.090e0</td>
<td></td>
<td>1212.0 3.127e0</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>1526.9 1.740e0</td>
<td></td>
<td>1417.0 2.393e0</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1713.7 1.478e0</td>
<td></td>
<td>1612.0 1.865e0</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>2438.1 1.052e0</td>
<td></td>
<td>1873.0 1.030e0</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Numerical results for Example 3 and 4
### Example 5

<table>
<thead>
<tr>
<th>Rank</th>
<th>Time</th>
<th>Residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>11640.0</td>
<td>1.872e2</td>
</tr>
<tr>
<td>50</td>
<td>1570.0</td>
<td>1.011e2</td>
</tr>
<tr>
<td>100</td>
<td>899.0</td>
<td>8.068e1</td>
</tr>
<tr>
<td>250</td>
<td>318.3</td>
<td>7.574e1</td>
</tr>
<tr>
<td>500</td>
<td>326.3</td>
<td>7.574e1</td>
</tr>
</tbody>
</table>

Table 4: Numerical results for Example 5
Final remarks

- A code named PenCorr.m can efficiently solve all sorts of rank constrained correlation matrix problems. Faster when rank is larger.

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

- The limitation is that it cannot solve problems for matrices exceeding the dimension 4,000 by 4,000 on a PC due to memory constraints.
Thank you! :}