An inexact accelerated proximal gradient method for large scale linearly constrained convex SDP

Defeng Sun

Department of Mathematics and Risk Management Institute
National University of Singapore

Based on a joint work with Kaifeng Jiang and Kim Chuan Toh at NUS

Convex semidefinite programming (SDP)

Consider the following convex (SDP) problem:

$$(P) \quad \min\left\{ f(x) \;:\; \mathcal{A}(x) = b,\; x \succeq 0,\; x \in \mathcal{S}^n \right\},$$

where f is a smooth convex function on $\mathcal{S}^n$, $\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$ is a linear map, $b \in \mathbb{R}^m$, and $\mathcal{S}^n$ is the space of $n \times n$ symmetric matrices. The notation $x \succeq 0$ means that x is positive semidefinite.

Let $\mathcal{A}^*$ be the adjoint of $\mathcal{A}$. The dual problem associated with (P)

$$(D) \; \max\left\{ f(x) - \langle \nabla f(x), x\rangle + \langle b, p\rangle \;:\; \nabla f(x) - \mathcal{A}^* p - z = 0, \; \begin{array}{c} z \succeq 0, \\ x \succeq 0 \end{array} \right\}.$$

Assume that the linear map $\mathcal{A}$ is surjective, and that strong duality holds for (P) and (D).
Let x_* be an optimal solution of (P) and (x_*, p_*, z_*) be an optimal solution of (D). Then, they must satisfy the following KKT conditions:

$$\mathcal{A}(x) = b, \quad \nabla f(x) - \mathcal{A}^* p - z = 0, \quad \langle x, z \rangle = 0, \quad x \succeq 0, \; z \succeq 0.$$

The problem (P) contains the following important special case of convex quadratic semidefinite programming (QSDP):

$$\min \left\{ \frac{1}{2} \langle x, \mathcal{Q}(x) \rangle + \langle c, x \rangle \; : \; \mathcal{A}(x) = b, x \succeq 0 \right\}, \tag{1}$$

where $\mathcal{Q} : \mathcal{S}^n \to \mathcal{S}^n$ is a given self-adjoint positive semidefinite linear operator and $c \in \mathcal{S}^n$.

The nearest correlation matrix problem

A typical example of QSDP is the nearest correlation matrix problem [Higham 2002].
Given a symmetric matrix $u \in \mathcal{S}^n$, we want to solve

$$\min\left\{\frac{1}{2}\|\mathcal{L}(x-u)\|^2 \ : \ \mathrm{Diag}(x)=e,\ x\succeq 0\right\}, \qquad (2)$$

where $\mathcal{L}:\mathcal{S}^n \to \mathbb{R}^{n\times n}$ is a linear map and $e \in \mathbb{R}^n$ is the vector of all ones. Here $\mathcal{Q}=\mathcal{L}^*\mathcal{L}$ and $c=-\mathcal{L}^*\mathcal{L}(u)$. A well studied special case of (2) is the W-weighted nearest correlation matrix problem

$$\min\left\{\frac{1}{2}\|W^{1/2}(x-u)W^{1/2}\|^2 \ : \ \mathrm{Diag}(x)=e,\ x\succeq 0\right\},$$

where $W \in \mathcal{S}^n$ is a given positive definite matrix.

For the W-weighted nearest correlation matrix problem, we have

- The alternating projection method [Higham 2002]
- The quasi-Newton method [Malick 2004]
- An inexact semismooth Newton-CG method [Qi and Sun 2006]
- An inexact interior-point method [Toh, Tütüncü and Todd 2007]

The second case is the H-weighted case of (2)

$$\min\left\{\frac{1}{2}\|H\circ(x-u)\|^2 \;:\; \mathrm{Diag}(x)=b,\; x\succeq 0\right\}, \tag{3}$$

where $H\in\mathcal{S}^n$ with nonnegative entries and"$\circ$" denoting the Hardamard product of two matrices defined by $(A\circ B)_{ij}=A_{ij}B_{ij}$.

The weight matrix H represents one's confidence levels on the estimated matrix on a component by component basis.

The corresponding entries of H are zeros for missing entries of u.

For the H-weighted nearest correlation matrix problem, we have

- An inexact interior-point method for a general convex QSDP [Toh 2008].
- an augmented Lagrangian dual method [Qi and Sun 2010]

  If the weight matrix H is very sparse or ill-conditioned, the conjugate gradient (CG) method would have great difficulty in solving the linear system of equations.

- A semismooth Newton-CG augmented Lagrangian method for convex quadratic programming over symmetric cones [Zhao 2009].
- A modified alternating direction method for convex quadratically constrained QSDPs [Sun and Zhang 2010].

Motivation: Choice of Algorithms

Assume that we are interested in solving the unconstrained problem

$$\min \ f(x)$$

with highly ill conditioned Hessian $\nabla f^2(x)$. Then

- Newton's method including inexact ones is certainly not feasible.
- Quasi Newton methods are out of touch due to high dimension.
- Gradient type methods are very few possible choices.

Why the APG method?

The accelerated proximal gradient (APG) method was first proposed by [Nesterov 1983] for minimizing smooth convex functions, later extended by [Beck and Teboulle 2009] to composite convex objective functions, and studied in a unifying manner by [Tseng 2008]. The algorithm we propose is based on the APG method (FISTA) [Beck and Teboulle 2009], where in the kth iteration with iterate $\overline{x}_k$, a subproblem of the following form must be solved:

$$\min_{x\in\mathcal{S}^n}\Big\{\langle\nabla f(\overline{x}_k), x-\overline{x}_k\rangle+\frac{1}{2}\langle x-\overline{x}_k, \mathcal{H}_k(x-\overline{x}_k)\rangle : \mathcal{A}(x) = b, x\succeq 0\Big\}, \tag{4}$$

where $\mathcal{H}_k : \mathcal{S}^n \to \mathcal{S}^n$ is a given self-adjoint positive definite linear operator.

Assume that $\nabla f(\cdot)$ is globally Lipschitz continuous. That is, there exists $L > 0$ such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y.$$

Attractive iteration complexity: $O(\sqrt{L/\varepsilon})$ for APG vs $O(L/\varepsilon)$ for proximal gradient (PG) method.

Limitations of FISTA:

1 $\mathcal{H}_k$ is restricted to $L\mathcal{I}$, where $\mathcal{I} : \mathcal{S}^n \to \mathcal{S}^n$ denotes the identity map and L is the Lipschitz constant of ∇f. (L could be very large)

2 The subproblem (4) must be solved exactly to generate the next iterate x_{k+1}.

An inexact APG method

For more generality, we consider the following minimization problem

$$\min\{F(x) := f(x) + g(x) : x \in \mathcal{X}\} \tag{5}$$

where $\mathcal{X}$ is a finite-dimensional Euclidean space. The functions $f : \mathcal{X} \to \mathbb{R}$, $g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ are proper, lower semi-continuous convex functions (possibly nonsmooth). We assume that dom$(g) := \{x \in \mathcal{X} : g(x) < \infty\}$ is closed, f is continuously differentiable on $\mathcal{X}$ and its gradient ∇f is Lipschitz continuous with modulus L on $\mathcal{X}$. It is a well known property that

$$f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{L}{2}\|x - y\|^2 \quad \forall\, x, y \in \mathcal{X}.$$

We also assume that the problem (5) is solvable with an optimal solution $x_* \in$ dom(g).

Algorithm 1: An inexact APG for (5)

Algorithm 1. Given a tolerance $\varepsilon > 0$. Input $y_1 = x_0 \in \text{dom}(g)$, $t_1 = 1$. Set $k = 1$. Iterate the following steps.

Step 1. Find an approximate minimizer

$$x_k \approx \arg\min_{y \in \mathcal{X}} \Big\{ f(y_k) + \langle \nabla f(y_k), y - y_k \rangle + \frac{1}{2} \langle y - y_k, \mathcal{H}_k (y - y_k) \rangle + g(y) \Big\},$$

where $\mathcal{H}_k$ is a self-adjoint positive definite linear operator that is chosen by the user.

Step 2. Compute $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$.

Step 3. Compute $y_{k+1} = x_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x_k - x_{k-1})$.

Given any positive definite linear operator $\mathcal{H}_j : \mathcal{X} \to \mathcal{X}$, and $y_j \in \mathcal{X}$, we define $q_j(\cdot) : \mathcal{X} \to \mathbb{R}$ by

$$q_j(x) = f(y_j) + \langle \nabla f(y_j), x - y_j \rangle + \frac{1}{2}\langle x - y_j, \mathcal{H}_j(x - y_j)\rangle.$$

Let $\{\xi_k\}, \{\epsilon_k\}$ be given convergent sequences of nonnegative numbers such that $\sum_{k=1}^{\infty} \xi_k < \infty$ and $\sum_{k=1}^{\infty} \epsilon_k < \infty$. In the j-th iteration of Algorithm 1, we assume the approximation minimizer x_j satisfies the following conditions

$$F(x_j) \;\leq\; q_j(x_j) + g(x_j) + \tfrac{\xi_j}{2t_j^2}, \tag{6}$$

$$\nabla f(y_j) + \mathcal{H}_j(x_j - y_j) + \gamma_j = \delta_j \text{ with } \|\mathcal{H}_j^{-1/2}\delta_j\| \leq \epsilon_j/(\sqrt{2}t_j) \tag{7}$$

where $\gamma_j \in \partial g(x_j; \frac{\xi_j}{2t_j^2})$ (the set of $\frac{\xi_j}{2t_j^2}$-subgradients of g at x_j).

Note that for x_j to be an approximate minimizer, we must have $x_j \in \text{dom}(g)$. We let $\tau = \frac{1}{2}\langle x_0 - x_*, \mathcal{H}_1(x_0 - x_*)\rangle,\ \bar{\epsilon}_k = \sum_{j=1}^{k} \epsilon_j,\ \bar{\xi}_k = \sum_{j=1}^{k}(\xi_j + \epsilon_j^2)$.

Theorem 1 *Suppose the conditions (6) and (7) hold, and $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then*

$$0 \le F(x_k) - F(x_*) \le \frac{4}{(k+1)^2}\Big((\sqrt{\tau} + \bar{\varepsilon}_k)^2 + 2\bar{\xi}_k\Big).$$

Specialized to $g(\cdot) = \delta(\cdot \,|\, \Omega)$

For the problem (P), we have $g(\cdot) = \delta(\cdot \,|\, \Omega)$ where $\Omega = \{x \in \mathcal{S}^n \ : \ \mathcal{A}(x) = b, \ x \succeq 0\}$ is the feasible set of (P). We need to solve the following constrained minimization problem:

$$\min \Big\{ \langle \nabla f(y_k), x - y_k \rangle + \frac{1}{2} \langle x - y_k, \mathcal{H}_k (x - y_k) \rangle \ : \ \mathcal{A}(x) = b, x \succeq 0 \Big\}. \quad (8)$$

Suppose we have an approximate solution (x_k, p_k, z_k) to the KKT optimality conditions for (8):

$$\begin{aligned} \nabla f(y_k) + \mathcal{H}_k(x_k - y_k) - \mathcal{A}^* p_k - z_k =: \delta_k &\approx 0 \\ \mathcal{A}(x_k) - b &= 0 \\ \langle x_k, z_k \rangle =: \varepsilon_k &\approx 0, \quad x_k, z_k \succeq 0. \end{aligned} \quad (9)$$

Let $\gamma_k = -\mathcal{A}^* p_k - z_k$. Then γ_k is an ε_k-subgradient of g at $x_k \in \Omega$ if $z_k \succeq 0$. However, in practice, we have $x_k \succeq 0$ but $r_k := \mathcal{A}(x_k) - b \neq 0$.

Suppose that there exists $\bar{x} \succ 0$ such that $\mathcal{A}(\bar{x}) = b$. Since $\mathcal{A}$ is surjective, $\mathcal{A}\mathcal{A}^*$ is nonsingular. Let $\omega_k = -\mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}(r_k)$. Note that $\mathcal{A}(x_k + \omega_k) = b$. However, $x_k + \omega_k$ may not be positive semidefinite. Thus we consider the following iterate:

$$\tilde{x}_k = \lambda(x_k + \omega_k) + (1-\lambda)\bar{x} = \lambda x_k + (\lambda\omega_k + (1-\lambda)\bar{x}), \ \lambda \in [0,1].$$

It is clear that $\mathcal{A}\tilde{x}_k = b$. By choosing $\lambda = 1 - \|\omega_k\|_2/(\|\omega_k\|_2 + \lambda_{\min}(\bar{x}))$, we can have that $\tilde{x}_k$ is positive semidefinite. We can also have that $(\tilde{x}_k, p_k, z_k)$ satisfies the condition (6) if

$$\|\omega_k\|_2 \le \min\left\{ \frac{\xi_k}{4t_k^2\sqrt{n}\|z_k\|}\left(1 + \frac{\lambda_{\max}(\bar{x})}{\lambda_{\min}(\bar{x})}\right)^{-1}, \ \frac{\varepsilon_k}{2\sqrt{2n\lambda_{\max}(\mathcal{H}_1)}\,t_k}\left(1 + \frac{\|\bar{x}-x_k\|_2}{\lambda_{\min}(\bar{x})}\right)^{-1} \right\}.$$

An inexact APG for (P)

Let $q_k(x) = f(y_k) + \langle \nabla f(y_k), x - y_k\rangle + \frac{1}{2}\langle x - y_k, \mathcal{H}_k(x - y_k)\rangle,\ x \in \mathcal{S}^n$.
Algorithm 2. Given a tolerance $\varepsilon > 0$. Input $y_1 = x_0 \in \mathcal{S}^n$, $t_1 = 1$. Set $k = 1$. Iterate the following steps.

Step 1. Find an approximate minimizer

$$x_k \approx \arg\min_{x\in\mathcal{X}} \Big\{ q_k(x) \ :\ x \in \Omega \Big\}, \qquad (10)$$

where $x_k \in \Omega_k \supseteq \Omega$.

Step 2. Compute $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$.

Step 3. Compute $y_{k+1} = x_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x_k - x_{k-1})$.

When $\Omega_k = \Omega$, the dual problem of (10) is given by

$$\max\Big\{ q_k(x) - \langle \nabla q_k(x), x\rangle + \langle b, p\rangle \mid \nabla q_k(x) - \mathcal{A}^* p - z = 0,\ z \succeq 0, x \succeq 0 \Big\}. \qquad (11)$$

Let $\{\xi_k\}, \{\epsilon_k\}, \{\mu_k\}$ be given convergent sequences of nonnegative numbers such that $\sum_{k=1}^{\infty} \xi_k < \infty$, $\sum_{k=1}^{\infty} \epsilon_k < \infty$, and $\sum_{k=1}^{\infty} \mu_k < \infty$, and Δ be a given positive number. We assume that (x_k, p_k, z_k) satisfies the following conditions:

$$
\begin{aligned}
f(x_k) &\le q_k(x_k) + \xi_k/(2t_k^2) \\
|\langle \nabla q_k(x_k), x_k\rangle - \langle b, p_k\rangle| &\le \Delta \\
\nabla q_k(x_k) - \mathcal{A}^* p_k - z_k &= \delta_k, \ \text{with } \|H_k^{-1/2}\delta_k\| \le \epsilon_k/(\sqrt{2}t_k) \\
\|\mathcal{A}(x_k) - b\| &\le \mu_k/t_k^2 \\
\langle x_k, z_k\rangle &\le \xi_k/(2t_k^2), \quad x_k \succeq 0, \ z_k \succeq 0.
\end{aligned}
$$

We also assume that $\mu_k/t_k^2 \geq \mu_{k+1}/t_{k+1}^2$ and $\epsilon_k/t_k \geq \epsilon_{k+1}/t_{k+1}$ for all k. We can have that (x_k, p_k, z_k) is an approximate optimal solution of (10) and (11). Note that $\Omega_k := \left\{ x \in \mathcal{S}^n : \|\mathcal{A}(x) - b\| \leq \mu_k/t_k^2,\ x \succeq 0 \right\}$ and $\Omega_{k+1} \subseteq \Omega_k$.

We let (x_*, p_*, z_*) be an optimal solution of (P) and (D),

$$\tau = \frac{1}{2}\langle x_0 - x_*, \mathcal{H}_1(x_0 - x_*)\rangle,\ \chi_k = \|p_{k-1} - p_k\|\mu_{k-1},\ \text{with } \chi_1 = 0,$$

$$\bar{\epsilon}_k = \sum_{j=1}^{k} \epsilon_j, \quad \bar{\chi}_k = \sum_{j=1}^{k} (\xi_j + \epsilon_j^2), \quad \bar{\xi}_k = \sum_{j=1}^{k} \chi_j.$$

Theorem 2 *Suppose* $M_k = \max_{1 \le j \le k}\{\sqrt{(\|p_*\| + \|p_j\|)\mu_j}\}$. *Then we have*

$$-\frac{4\|p_*\|\mu_k}{(k+1)^2} \le f(x_k) - f(x_*)$$
$$\le \frac{4}{(k+1)^2}\Big((\sqrt{\tau} + \bar{\epsilon}_k)^2 + \|p_k\|\mu_k + 2\bar{\epsilon}_k M_k + 2(\bar{\xi}_k + \bar{\chi}_k)\Big).$$

$\{\|p_k\|\}$ bounded (?) $\Longrightarrow \{M_k\}$ and $\{\bar{\chi}_k\}$ bounded $\Longrightarrow O(1/k^2)$.

Boundedness of $\{p_k\}$

Lemma 1 *Suppose that there exists* $(\bar{x}, \bar{p}, \bar{z})$ *such that*

$$\mathcal{A}(\bar{x}) = b, \; \bar{x} \succeq 0, \quad \nabla f(\bar{x}) = \mathcal{A}^* \bar{p} + \bar{z}, \; \bar{z} \succ 0.$$

If the sequence $\{f(x_k)\}$ *is bounded from above, then the sequence* $\{x_k\}$ *is bounded.*

Lemma 2 *Suppose that* $\{x_k\}$ *is bounded and there exists* $\hat{x}$ *such that*

$$\mathcal{A}(\hat{x}) = b, \; \hat{x} \succ 0.$$

Then the sequence $\{z_k\}$ *is bounded. In addition, the sequence* $\{p_k\}$ *is also bounded.*

In many cases, such as the nearest correlation matrix problem (2), the condition that $\{f(x_k)\}$ is bounded above or that $\{x_k\}$ is bounded can be ensured since Ω_1 is bounded.

A semismooth Newton-CG method for inner subproblems

Suppose that at each iteration we are able to choose the self-adjoint positive definite linear operator $\mathcal{H}_k$ of the form:

$$\mathcal{H}_k(x) := w_k \circledast w_k(x) = w_k x w_k, \quad \text{where } w_k \in \mathcal{S}^n \text{ positive definite,}$$

such that $f(x) \le q_k(x)$ for all $x \in \Omega$ (A simple choice: $w_k = \sqrt{L}I$). Then $q_k(\cdot)$ in (10) can equivalently be written as

$$q_k(x) = \frac{1}{2}\|w_k^{1/2}(x - u_k)w_k^{1/2}\|^2 + f(y_k) - \frac{1}{2}\|w_k^{-1/2}\nabla f(y_k)w_k^{-1/2}\|^2,$$

where $u_k = y_k - w_k^{-1}\nabla f(y_k)w_k^{-1}$.

Then (10) can be equivalently written as the following well-studied W-weighted semidefinite least squares problem

$$\min\left\{\frac{1}{2}\|w_k^{1/2}(x-u_k)w_k^{1/2}\|^2 \,:\, \mathcal{A}(x)=b,\, x\succeq 0\right\}, \tag{12}$$

which can be efficiently solved by the SSNCG method in [Qi and Sun 2006].

The availability of the SSNCG is vital for our inexact APG to work.

For example, for a 2000 by 2000 weighted nearest correlation matrix problem, SSNCG needs 23 seconds to get error less than 10^{-9} while the APG needs more than 4980 seconds to get gradient error as 0.68.

Symmetrized Kronecker product approximation of $\mathcal{Q}$

- For the H-weighted NCM problem where $\mathcal{Q}(x) = (H \circ H) \circ x$, let $w = \mathrm{diag}(d)$, where the vector $d \in \mathbb{R}^n$ can be chosen as

$$d_j = \max\left\{\epsilon, \max_{1 \le i \le n}\{H_{ij}\}\right\}, \; j = 1, \ldots, n,$$

where $\epsilon > 0$ is a small positive number.

- If $\mathcal{Q}(x) = B \circledast I(x) = (Bx + xB)/2$, $B \in \mathcal{S}^n_+$. Suppose we have the eigenvalue decomposition $B = P\Lambda P^T$, where $\Lambda = \mathrm{diag}(\lambda)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)^T$ is the vector of eigenvalues of B. Let $M = \frac{1}{2}(\lambda e^T + e\lambda^T)$ with $e \in \mathbb{R}^n$ being the vector of all ones. We consider the following nonconvex minimization problem:

$$\min \left\{ \sum_{i=1}^n \sum_{j=1}^n h_i h_j \mid h_i h_j - M_{ij} \geq 0 \;\forall\, i, j = 1, \ldots, n,\; h \in \mathbb{R}^n_+ \right\}.$$

If $\hat{h}$ is a feasible solution to the above problem, let $w_k = P\mathrm{diag}(\hat{h})P^T$.

Numerical results for convex QSDP problems

In our numerical experiments, we stop the inexact APG algorithm when

$$\max\{R_P, R_D\} \leq 10^{-6}.$$

Example 1 We consider the following H-weighted nearest correlation matrix problem

$$\min\left\{\frac{1}{2}\|H \circ (x-u)\|^2 \mid \operatorname{Diag}(x) = e, x \succeq 0\right\}.$$

We compare the performance of our inexact APG (IAPG) method and the augmented Lagrangian dual method (AL) studied by [Qi and Sun 2010]. We set the tolerance $\texttt{Tol1} = 10^{-4}$ in AL. Given the correlation matrices $\widehat{u}$, we perturb $\widehat{u}$ to

$$u := (1-\alpha)\widehat{u} + \alpha E,$$

where $\alpha \in (0,1)$ and E is a randomly generated symmetric matrix with entries in $[-1,1]$.

The weight matrix H is a sparse random symmetric matrix with about 50% nonzero entries.

Algo.	problem (n)	α	iter/newt	R_P	R_D	pobj	time	rank
IAPG	ER (692)	0.1	167/172	2.27e-10	9.92e-7	1.26095534e+1	3:30	189
		0.05	187/207	3.93e-11	9.54e-7	1.14555927e+0	3:40	220
AL	ER (692)	0.1	12	3.73e-7	4.63e-7	1.26095561e+1	9:28	189
		0.05	12	3.21e-7	1.02e-6	1.14555886e+0	14:14	220
IAPG	Arabidopsis (834)	0.1	125/133	3.28e-10	9.36e-7	3.46252363e+1	4:01	191
		0.05	131/148	2.41e-10	9.75e-7	5.50148194e+0	4:09	220
AL	Arabidopsis (834)	0.1	13	2.28e-7	7.54e-7	3.46252429e+1	12:35	191
		0.05	12	2.96e-8	1.01e-6	5.50148169e+0	22:49	220
IAPG	Leukemia (1255)	0.1	104/111	5.35e-10	7.97e-7	1.08939600e+2	9:24	254
		0.05	96/104	4.81e-10	9.31e-7	2.20789464e+1	8:35	276
AL	Leukemia (1255)	0.1	12	3.06e-7	2.74e-7	1.08939601e+2	22:04	254
		0.05	11	2.90e-7	8.57e-7	2.20789454e+1	28:37	276
IAPG	hereditarybc (1869)	0.1	67/87	2.96e-10	8.68e-7	4.57244497e+2	17:56	233
		0.05	64/85	9.58e-10	7.04e-7	1.13171325e+2	17:32	236
AL	hereditarybc (1869)	0.1	13	2.31e-7	3.55e-7	4.57244525e+2	38:35	233
		0.05	11	2.51e-7	6.29e-7	1.13171335e+2	36:31	236

Example 2: "bad" weight matrix H

We consider the same problem as in Example 1, but the weight matrix H is generated from a weight matrix H_0 used by a hedge fund company. The matrix H_0 is a 93×93 symmetric matrix with all positive entries.

- It has 24% of the entries equal to 10^{-5} and the rest are in the interval $[2, 1.28 \times 10^3]$.
- It has 28 eigenvalues in $[-520, -0.04]$, 11 eigenvalues in $[-5 \times 10^{-13}, 2 \times 10^{-13}]$, and the rest 54 eigenvalues in $[10^{-4}, 2 \times 10^4]$.

We set the tolerance `Tol1` $= 10^{-2}$ in AL. ("$*$" means "$>$ 24 hours")

Algo.	problem (n)	α	iter/newt	R_P	R_D	pobj	time	rank
IAPG	ER (692)	0.1	62/156	2.48e-9	9.72e-7	1.51144194e+7	2:33	254
		0.05	56/145	3.58e-9	9.55e-7	3.01128282e+6	2:22	295
AL	ER (692)	0.1	16	1.22e-5	5.80e-6	1.51144456e+7	2:05:38	288
		0.05	12	3.11e-5	6.29e-6	3.01123631e+6	53:15	309
IAPG	Arabidopsis (834)	0.1	61/159	6.75e-9	9.98e-7	2.69548461e+7	4:01	254
		0.05	54/145	1.06e-8	9.82e-7	5.87047119e+6	3:41	286
AL	Arabidopsis (834)	0.1	19	3.04e-6	3.94e-6	2.69548769e+7	4:49:00	308
		0.05	13	1.69e-5	6.76e-6	5.87044318e+6	1:28:59	328
IAPG	Leukemia (1255)	0.1	65/158	8.43e-9	9.86e-7	7.17192454e+7	11:32	321
		0.05	55/143	1.19e-7	9.80e-7	1.70092540e+7	10:18	340
AL	Leukemia (1255)	0.1	*	*	*	*	*	*
		0.05	13	3.19e-5	5.15e-6	1.70091646e+7	5:55:21	432
IAPG	hereditarybc (1869)	0.1	48/156	2.08e-8	9.16e-7	2.05907938e+8	29:07	294
		0.05	49/136	6.39e-8	9.61e-7	5.13121563e+7	26:16	297
AL	hereditarybc (1869)	0.1	*	*	*	*	*	*
		0.05	*	*	*	*	*	*

Example 3 We consider the linearly constrained convex QSDP problem, where $\mathcal{Q}(x) = \frac{1}{2}(Bx + xB)$ for a given $B \succ 0$ and $\mathcal{A}(x) = \mathrm{Diag}(x)$.

n; m	cond(B)	iter/newt	R_P	R_D	pobj	dobj	time
500; 500	9.21e+0	9/9	3.24e-10	9.70e-7	-4.09219187e+4	-4.09219188e+4	13
1000; 1000	9.43e+0	9/9	3.68e-10	9.28e-7	-8.41240999e+4	-8.41241006e+4	1:13
2000; 2000	9.28e+0	9/9	3.16e-10	8.53e-7	-1.65502323e+5	-1.65502325e+5	8:49
2500; 2500	9.34e+0	9/9	3.32e-10	8.57e-7	-2.07906307e+5	-2.07906309e+5	16:15
3000; 3000	9.34e+0	9/9	2.98e-10	8.13e-7	-2.49907743e+5	-2.49907745e+5	29:02

Example 4 We consider the same problem as Example 3, but the linear map $\mathcal{A}$ is generated by using the first generator in [Malick, Povh, Rendl, and Wiegele 2009] with order $p = 3$. The positive definite matrix B is generated by using MATLAB's built-in function: `B = gallery('lehmer',n)` with $\mathrm{cond}(B) \in [n, 4n^2]$.

n; m	cond(B)	iter/newt	R_P	R_D	pobj	dobj	time
500; 10000	2.67e+5	51/102	3.02e-8	9.79e-7	-9.19583895e+3	-9.19584894e+3	1:29
1000; 50000	1.07e+6	62/115	2.43e-8	9.71e-7	-1.74777588e+4	-1.74776690e+4	11:46
2000; 100000	4.32e+6	76/94	5.24e-9	5.28e-7	-3.78101950e+4	-3.78101705e+4	1:14:04
2500; 100000	6.76e+6	80/96	4.62e-9	5.64e-7	-4.79637904e+4	-4.79637879e+4	2:11:01

We need more creative ideas!