A Riemannian Inexact Newton-CG Method for Constructing Nonnegative Matrix with Prescribed Realizable Spectrum

Xiao-Qing Jin

Joint with Zhi Zhao and Zheng-Jian Bai

Department of Mathematics, University of Macau, Macao. Email: xqjin@umac.mo.

Outline

An $m \times n$ real matrix C is called a nonnegative matrix if

$$(C)_{ij} \geq 0, \;\; i, = 1, \ldots, m, \; j = 1, \ldots, n,$$

where $(C)_{ij}$ denotes the (i, j)-entry of C. Nonnegative matrices arise in various applications such as Markov chains, linear complementarity problems, probabilistic algorithms, discrete distributions, group theory, and economics.

A spectrum which occurs as the spectrum of some nonnegative matrix is called a realizable spectrum for nonnegative matrices, and a nonnegative matrix with the realizable spectrum is called a realization or realizing matrix.

Nonnegative inverse eigenvalue problem (NIEP) is a special kind of inverse eigenvalue problems, which concerns whether or not a self-conjugate collections of n complex numbers (counting multiplicities) is a realizable spectrum for nonnegative matrices. The NIEP is a classical unsolved problem in linear algebra.

The NIEP sparks the interest in constructing algorithms to find realizing matrices for prescribed realizable spectrum.

In this talk, we consider the inverse eigenvalue problem of constructing a nonnegative matrix with prescribed realizable spectrum.

Problem I. Given an n-tuple $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, which is a realizable spectrum for nonnegative matrices, find an n-by-n nonnegative matrix C such that its eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_n$.

There are a few numerical methods for solving **Problem I** such as the constructive method [9], the alternating projection method [8], isospectral gradient flow methods [2, 3, 4, 5], and a fast recursive algorithm [6] for the case where the prescribed eigenvalues are all real and satisfy an additional inequality.

The deficiency of the existing numerical methods is that these methods are either restricted to solve small- or medium-scale problems or can only be used to solve some special subproblems.

Outline

We would like to reformulate **Problem I** as an equivalent nonlinear matrix equation between a Riemannian manifold and a Euclidean space. First we need to introduce some special matrix sets.

The matrix set $\mathbb{R}_+^{n\times n}$

$$\mathbb{R}_+^{n\times n} = \{S \odot S \mid S \in \mathbb{R}^{n\times n}\},$$

where $\odot$ denotes the Hadmard product.

Since we assume the n-tuple $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ is a realizable spectrum, it is closed under complex conjugation. Without loss of generality, we can assume

$$\lambda_{2j-1} = a_j + b_j\sqrt{-1}, \quad \lambda_{2j} = a_j - b_j\sqrt{-1}, \quad j = 1, \ldots, s;$$
$$\lambda_j \in \mathbb{R}, \quad j = 2s+1, \ldots, n,$$

where $a_j, b_j \in \mathbb{R}$ for $j = 1, \ldots, s$. Then we define a block diagonal matrix by

$$\Lambda := \mathrm{blkdiag}\left(\lambda_1^{[2]}, \ldots, \lambda_s^{[2]}, \lambda_{2s+1}, \ldots, \lambda_n\right)$$

with diagonal blocks $\lambda_1^{[2]}, \ldots, \lambda_s^{[2]}, \lambda_{2s+1}, \ldots, \lambda_n$, where

$$\lambda_j^{[2]} = \begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}, \quad j = 1, \ldots, s.$$

The set $\mathcal{V}$

$$\mathcal{V} := \{V \in \mathbb{R}^{n\times n} \mid V_{ij} = 0,\ (i,j) \in \mathcal{I}\},$$

where $\mathcal{I}$ is the index subset:

$$\mathcal{I} := \{(i,j) \mid i \geq j \text{ or } \Lambda_{ij} \neq 0,\ i,j = 1,\ldots,n\}.$$

The orthogonal group

$$\mathcal{O}(n) := \{Q \in \mathbb{R}^{n\times n} \mid Q^T Q = I_n\},$$

where I_n is the identity matrix of order n.

The matrix set $\mathcal{M}(\Lambda)$

$$\mathcal{M}(\Lambda) := \{X \in \mathbb{R}^{n\times n} \mid X = Q(\Lambda + V)Q^T,\ Q \in \mathcal{O}(n),\ V \in \mathcal{V}\}.$$

The matrix set $\mathcal{M}(\Lambda)$ just consists of all the real $n \times n$ matrices with the same spectrum $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$.

Solvability of the NIEP

The NIEP has a solution if and only if $\mathcal{M}(\Lambda) \cap \mathbb{R}_+^{n\times n} \neq \emptyset$.

Since the prescribed n-tuple $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ is realizable, the **Problem I** aims to solve the following constrained nonlinear matrix equation:

$$G(S, Q, V) = \mathbf{0}_{n\times n}. \tag{1}$$

The smooth mapping $G : \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V} \to \mathbb{R}^{n\times n}$ is defined by

$$G(S, Q, V) := S \odot S - Q(\Lambda + V)Q^T, \quad (S, Q, V) \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}.$$

Once we find a solution $(\overline{S}, \overline{Q}, \overline{V}) \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ to the nonlinear equation (1), then the matrix $\overline{C} := \overline{S} \odot \overline{S}$ is a solution to **Problem I**.

The dimension of $\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ is given by

$$\dim(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) = n^2 + \frac{n(n-1)}{2} + |\mathcal{J}|,$$

where $\mathcal{J}$ is the complementary index set of $\mathcal{I}$ with respect to the index set $\mathcal{N} := \{(i,j) \mid i,j = 1,\ldots,n\}$, and $|\mathcal{J}|$ is the cardinality of $\mathcal{J}$. Thus

$$\dim(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) > \dim \mathbb{R}^{n\times n} \quad \text{for } n \geq 2.$$

Hence, the nonlinear matrix equation (1) is under-determined for all $n \geq 2$.

The tangent space of the product manifold $\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ at a point $(S,Q,V)\in\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ is given by

$$T_{(S,Q,V)}\big(\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}\big)=\big\{(A,Q\Omega,B)\mid\Omega^T=-\Omega,\ A,\Omega\in\mathbb{R}^{n\times n},B\in\mathcal{V}\big\}.$$

Let $\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ be endowed with the induced Riemannian metric:

$$g_{(S,Q,V)}\big((\xi_1,\zeta_1,\eta_1),(\xi_2,\zeta_2,\eta_2)\big):=\mathrm{tr}(\xi_1^T\xi_2)+\mathrm{tr}(\zeta_1^T\zeta_2)+\mathrm{tr}(\eta_1^T\eta_2),\qquad(2)$$

for all $(S,Q,V)\in\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ and $(\xi_1,\zeta_1,\eta_1),(\xi_2,\zeta_2,\eta_2)\in T_{(S,Q,V)}\big(\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}\big)$. Then $\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ is an embedded Riemannian submanifold of $\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times n}$ with the induced Riemannian metric.

The differential $\mathrm{D}G(S,Q,V) : T_{(S,Q,V)}\big(\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}\big)\to T_{G(S,Q,V)}\mathbb{R}^{n\times n}\simeq\mathbb{R}^{n\times n}$ of G at $(S,Q,V)\in\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V}$ is determined by

$$\mathrm{D}G(S,Q,V)[(\Delta S,\Delta Q,\Delta V)] = 2S\odot\Delta S + [Q(\Lambda+V)Q^T, \Delta QQ^T] - Q\Delta VQ^T$$

for all $(\Delta S,\Delta Q,\Delta V)\in T_{(S,Q,V)}(\mathbb{R}^{n\times n}\times\mathcal{O}(n)\times\mathcal{V})$, where $[A,B] := AB-BA$ means the Lie Bracket of two matrices.

With respect to the induced Riemannian metric given by (2), the adjoint operator $(\mathrm{D}G(S,Q,V))^* : T_{G(S,Q,V)}\mathbb{R}^{n\times n} \to T_{(S,Q,V)}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V})$ of $\mathrm{D}G(S,Q,V)$ is determined by

$$
\begin{aligned}
& (\mathrm{D}G(S,Q,V))^*[\Delta Z] \\
= \;& \left(2S \odot \Delta Z, \frac{1}{2}\left([Q(\Lambda+V)Q^T, (\Delta Z)^T], -W \odot (Q^T \Delta Z Q)\right)\right)
\end{aligned}
$$

for all $\Delta Z \in T_{G(S,Q,V)}\mathbb{R}^{n\times n}$, where $W \in \mathbb{R}^{n\times n}$ is defined by

$$
W_{ij} = \begin{cases} 0, & (i,j) \in \mathcal{I}, \\ 1, & (i,j) \in \mathcal{J}. \end{cases}
$$

To solve **Problem I**, one may propose the following geometric Newton method: Given the current iterate $X^k := (S^k, Q^k, V^k) \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$, solve the Newton equation:

$$\mathrm{D}G(X^k)[\Delta X^k] = -G(X^k) \tag{3}$$

for $\Delta X^k := (\Delta S^k, \Delta Q^k, \Delta V^k) \in T_{X^k}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V})$ and set

$$X^{k+1} := R_{X^k}(\Delta X^k),$$

where $R : T(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ is a retraction and the set $T(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V})$ denotes the tangent bundle of $\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$.

The Newton equation (3) has a solution if and only if it holds

$$\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^{\dagger}[-G(X^k)] = -G(X^k), \tag{4}$$

where $(\mathrm{D}G(X^k))^{\dagger}$ means the pseudoinverse of the linear operator $\mathrm{D}G(X^k)$. The Newton equation (3) is an under-determined linear system. If it is solvable, then it has infinite solutions. The minimum norm solution of (3) is given by:

$$\Delta X^k = -(\mathrm{D}G(X^k))^{\dagger} G(X^k).$$

In particular, if the linear operator $\mathrm{D}G(X^k)$ is surjective, then we have [7, p.165]:

$$(\mathrm{D}G(X^k))^\dagger = (\mathrm{D}G(X^k))^* \circ \left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right)^{-1}. \tag{5}$$

Based on (5), the solvability condition (4) is satisfied if $\mathrm{D}G(X^k)$ is surjective. In this case, one may solve the following normal equation:

$$\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*[\Delta Z^k] = -G(X^k), \quad \text{s.t.} \quad \Delta Z^k \in T_{G(X^k)}\mathbb{R}^{n\times n} \tag{6}$$

for the minimum norm solution

$$\Delta X^k = (\mathrm{D}G(X^k))^*[\Delta Z^k] \in T_{X^k}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}).$$

Thus, the conjugate gradient (CG) method can be used to solve the self-adjoint and positive definite equation (6).

If $\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*$ is singular, then the equality in (5) does not hold. For the pseudoinverse $(\mathrm{D}G(X^k))^\dagger$, we have

$$(\mathrm{D}G(X^k))^\dagger = \lim_{\sigma \to 0+} (\mathrm{D}G(X^k))^* \circ \left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^* + \sigma \mathrm{id}_{T_{G(X^k)}\mathbb{R}^{n\times n}}\right)^{-1},$$

where $\mathrm{id}_{T_{G(X^k)}\mathbb{R}^{n\times n}}$ denotes the identity operator on $T_{G(X^k)}\mathbb{R}^{n\times n}$. Instead of (6), one may solve the following perturbed normal equation:

$$\left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^* + \overline{\sigma} \mathrm{id}_{T_{G(X^k)}\mathbb{R}^{n\times n}}\right)[\Delta Z^k] = -G(X^k)$$

for $\Delta Z^k \in T_{G(X^k)}\mathbb{R}^{n\times n} \simeq \mathbb{R}^{n\times n}$, where $\overline{\sigma} > 0$ is a prescribed constant.

Algorithm 2.1 (**Riemannian inexact Newton-CG method)**

Step 0. Choose an initial point $X^0 \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$, $\overline{\sigma}_{\max}, \overline{\eta}_{\max}, \widehat{\eta}_{\max} \in [0,1)$, $t \in (0,1)$, $0 < \theta_{\min} < \theta_{\max} < 1$. Let $k := 0$.

Step 1. Apply the CG method to solving

$$\big(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^* + \overline{\sigma}_k \mathrm{id}_{T_{G(X^k)}\mathbb{R}^{n\times n}}\big)[\Delta Z^k] = -G(X^k), \tag{7}$$

for $\Delta Z^k \in T_{G(X^k)}\mathbb{R}^{n\times n}$ such that

$$\big\|\big(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^* + \overline{\sigma}_k \mathrm{id}_{T_{G(X^k)}\mathbb{R}^{n\times n}}\big)[\Delta Z^k] + G(X^k)\big\|_F \le \overline{\eta}_k \|G(X^k)\|_F, \tag{8}$$

and

$$\|\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*[\Delta Z^k] + G(X^k)\|_F \le \widehat{\eta}_{\max} \|G(X^k)\|_F, \tag{9}$$

where $\overline{\sigma}_k := \min\{\overline{\sigma}_{\max}, \|G(X^k)\|_F\}$, $\overline{\eta}_k := \min\{\overline{\eta}_{\max}, \|G(X^k)\|_F\}$.

Then let

$$\widehat{\Delta X}^k = (\mathrm{D}G(X^k))^*[\Delta Z^k], \quad \widehat{\eta}_k := \frac{\|\mathrm{D}G(X^k)[\widehat{\Delta X}^k] + G(X^k)\|_F}{\|G(X^k)\|_F}. \tag{10}$$

Step 2. Evaluate $G\big(R^k_X(\widehat{\Delta X}^k)\big)$. Set $\eta_k = \widehat{\eta}_k$ and $\Delta X^k = \widehat{\Delta X}^k$.

Repeat until $\|G\big(R_{X^k}(\Delta X^k)\big)\|_F \le (1 - t(1-\eta_k))\|G(X^k)\|_F$.

Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.

Replace ΔX^k by $\theta \Delta X^k$ and η_k by $1 - \theta(1-\eta_k)$.

end (Repeat)

Set

$$X^{k+1} := R_{X^k}(\Delta X^k).$$

Step 3. Replace k by $k+1$ and go to **Step 1**.

A retraction $R : T(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ is given by

$$R_{(S,Q,V)}(\xi_S, \zeta_Q, \eta_V) = \big(S + \xi_S, \mathrm{qf}(Q + \zeta_Q), V + \eta_V\big)$$

for all $(S, Q, V) \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ and $(\xi_S, \eta_Q, \gamma_V) \in T_{(S,Q,V)}\big(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}\big)$. Here, $\mathrm{qf}(A)$ means the Q factor of the QR decomposition of a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ in the form of $A = Q\widetilde{R}$ with $Q \in \mathcal{O}(n)$ and $\widetilde{R}$ being an upper triangular matrix with strictly positive diagonal entries.

Outline

On the iterate $\widehat{\Delta X}^k$ generated by Algorithm 2.1, we have the following estimate.

Lemma 3.1

Assume that $\mathrm{D}G(X^k) : T_{X^k}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to T_{G(X^k)}\mathbb{R}^{n\times n}$ is surjective for all k. If the linear matrix equation (7) is solvable such that conditions (8) and (9) are satisfied for all k, then one has for all k,

$$\|\widehat{\Delta X}^k\| \leq (1+\overline{\eta}_k)|||(\mathrm{D}G(X^k))^\dagger||| \cdot \|G(X^k)\|_F.$$

Proof

Let

$$\mathrm{id} := \mathrm{id}_{T_{G(X^k)}},\ J(X^k) := \mathrm{D}G(X^k)\circ(\mathrm{D}G(X^k))^* + \overline{\sigma}_k \mathrm{id},\ V(X^k) := G(X^k) + J(X^k)[\Delta Z^k].$$

We get by (8),

$$\|V(X^k)\|_F \leq \overline{\eta}_k \|G(X^k)\|_F. \tag{11}$$

By the assumption that $\mathrm{D}G(X^k)$ is surjective for all k, we have by (8) and (11),

$$
\begin{aligned}
& \|\widehat{\Delta X}^k\| = \|(\mathrm{D}G(X^k))^*[\Delta Z^k]\| \\
\leq \quad & |||(\mathrm{D}G(X^k))^* \circ \big(J(X^k)\big)^{-1}||| \cdot \|J(X^k)[\Delta Z^k]\|_F \\
= \quad & |||(\mathrm{D}G(X^k))^* \circ \big(J(X^k)\big)^{-1}||| \cdot \|V(X^k) - G(X^k)\|_F \\
\leq \quad & |||(\mathrm{D}G(X^k))^* \circ \big(J(X^k)\big)^{-1}||| \cdot \big(\|V(X^k)\|_F + \|G(X^k)\|_F\big) \\
\leq \quad & (1+\overline{\eta}_k)|||(\mathrm{D}G(X^k))^* \circ \big(J(X^k)\big)^{-1}||| \cdot \|G(X^k)\|_F \\
\leq \quad & (1+\overline{\eta}_k)|||(\mathrm{D}G(X^k))^* \circ \big(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\big)^{-1}||| \cdot \|G(X^k)\|_F \\
= \quad & (1+\overline{\eta}_k)|||(\mathrm{D}G(X^k))^\dagger||| \cdot \|G(X^k)\|_F.
\end{aligned}
$$

On the upper bound of the iterate $\widehat{\eta}_k$ generated by Algorithm 2.1, we have the following result.

Lemma 3.2

Assume that $\mathrm{D}G(X^k) : T_{X^k}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to T_{G(X^k)}\mathbb{R}^{n\times n}$ is surjective for all k. If the linear matrix equation (7) is solvable such that conditions (8) and (9) are satisfied for all k, then one has for all k,

$$\widehat{\eta}_k \leq \min\left\{ \frac{\overline{\sigma}_k}{\lambda_{\min}\left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right) + \overline{\sigma}_k} + \overline{\eta}_k, \widehat{\eta}_{\max} \right\}, \tag{12}$$

where $\lambda_{\min}(\cdot)$ means the smallest eigenvalue of a positive definite linear operator.

Proof

Let id, $J(X^k)$ and $V(X^k)$ be defined as in Lemma 3.1. By assumption, $\mathrm{D}G(X^k)$ is surjective for all k. It follows from (10) and (11) that for all k,

$$
\begin{aligned}
& \|G(X^k) + \mathrm{D}G(X^k)[\widehat{\Delta X}^k]\|_F \\
= \ & \|G(X^k) + \mathrm{D}G(X^k)\left[(\mathrm{D}G(X^k))^*[\Delta Z_k]\right]\|_F \\
= \ & \|G(X^k) + \left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right) \circ \left(J(X^k)\right)^{-1}[V(X^k) - G(X^k)]\|_F \\
\leq \ & |\!|\!|\mathrm{id} - \left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right) \circ \left(J(X^k)\right)^{-1}|\!|\!| \cdot \|G(X^k)\|_F \\
& + |\!|\!|\left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right) \circ \left(J(X^k)\right)^{-1}|\!|\!| \cdot \|V(X^k)\|_F \\
\leq \ & \left(\frac{\overline{\sigma}_k}{\lambda_{\min}\left(\mathrm{D}G(X^k) \circ (\mathrm{D}G(X^k))^*\right) + \overline{\sigma}_k} + \overline{\eta}_k\right)\|G(X^k)\|_F.
\end{aligned}
$$

This, together with (9), yields (12).

On the repeat-loop of Algorithm 2.1, we have the following lemma.

Lemma 3.3

Assume that the differential $\mathrm{D}G(X^k) : T_{X^k}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to T_{G(X^k)}\mathbb{R}^{n\times n}$ is surjective and the linear matrix equation (7) is solvable such that conditions (8) and (9) are satisfied in the k-th iteration of Algorithm 2.1. Then the repeat-loop terminates in finite steps with ΔX^k and η_k satisfying

$$\begin{cases} \|G(X^k) + \mathrm{D}G(X^k)[\Delta X^k]\|_F \leq \eta_k \|G(X^k)\|_F, \\ \|G(X^{k+1})\|_F \leq \big(1 - t(1-\eta_k)\big)\|G(X^k)\|_F. \end{cases} \tag{13}$$

Assumption 3.4

The linear operator $\mathrm{D}G(\overline{X}) : T_{\overline{X}}(\mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}) \to T_{G(\overline{X})}\mathbb{R}^{n\times n}$ is surjective, where $\overline{X} \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}$ is an accumulation point of the sequence $\{X^k\}$ generated by Algorithm 2.1.

Under the above assumption, we have the following theorem on the global convergence of Algorithm 2.1.

Theorem 3.5

Let $\overline{X}$ be an accumulation point of the sequence $\{X^k\}$ generated by Algorithm 2.1. Suppose that Assumption 3.4 is satisfied. Then the sequence $\{X^k\}$ converges to $\overline{X}$ and $G(\overline{X}) = \mathbf{0}_{n\times n}$.

Quadratic convergence

We have the following result on the backtracking line search procedure.

Lemma 3.6

Let $\overline{X}$ be an accumulation point of the sequence $\{X^k\}$ generated by Algorithm 2.1. Suppose that Assumption 3.4 is satisfied. Then $\eta_k = \widehat{\eta}_k$ and $\Delta X^k = \widehat{\Delta X}^k$ for all k sufficiently large.

We now establish the quadratic convergence of Algorithm 2.1.

Theorem 3.7

Let $\overline{X}$ be an accumulation point of the sequence $\{X^k\}$ generated by Algorithm 2.1. Suppose that Assumption 3.4 is satisfied. Then the sequence $\{X^k\}$ converges to $\overline{X}$ quadratically.

Denote

$$\begin{cases} J_{\overline{S}} := \mathrm{Diag}\big(\mathrm{vec}(\overline{S})\big) \in \mathbb{R}^{n^2\times n^2}, \\ J_{\overline{Q}} := (I_{n^2} - \widehat{P})\Big((\overline{S}\odot\overline{S})\otimes I_n - I_n\otimes(\overline{S}\odot\overline{S})^T\Big) \in \mathbb{R}^{n^2\times n^2}, \\ J_{\overline{V}} := \mathrm{Diag}\big(\mathrm{vec}(W)\big)(\overline{Q}\otimes\overline{Q})^T \in \mathbb{R}^{n^2\times n^2} \end{cases}$$

and

$$J_{\overline{X}} := \begin{bmatrix} J_{\overline{S}} \\ J_{\overline{Q}} \\ J_{\overline{V}} \end{bmatrix} \in \mathbb{R}^{3n^2\times n^2}.$$

The operator $\mathrm{D}G(\overline{X})$ is surjective $\iff$ The matrix $J_{\overline{X}}$ is of full column rank.

Outline

Example 4.1

We consider **Problem I** with varying n. Let $\widehat{C}$ be a random $n \times n$ nonnegative matrix with each entry generated from the uniform distribution on the interval $[0, 1]$. We choose the eigenvalues of $\widehat{C}$ as prescribed spectrum.

To illustrate the efficiency of our algorithm, we compare Algorithm 2.1 with the alternating projection method (ALP) [8], the Riemannian Fletcher-Reeves conjugate gradient method (RFR) [10], and the geometric Polak-Ribière-Polyak-based nonlinear conjugate gradient method (GPRP) [11].

The starting points by the built-in functions `rand`, `schur`, and `svd`:

$$S \odot S = \texttt{rand}\,(n,n), \quad S^0 = S \in \mathbb{R}^{n\times n}, \quad C^0 = S^0 \odot S^0,$$

$$\left[Q^0, V\right] = \texttt{schur}\,(S^0 \odot S^0, '\mathrm{real}'), \quad V^0 = W \odot V.$$

The stopping criteria are set to be

$$\|G(X^k)\|_F < 10^{-8}.$$

In our numerical tests, we set $\overline{\sigma}_{\max} = 0.01$, $\overline{\eta}_{\max} = 0.1$, $\widehat{\eta}_{\max} = 0.9$, $\theta_{\min} = 0.1$, $\theta_{\max} = 0.9$, and $t = 10^{-4}$. The largest number of iterations in ALP is set to be 100000. The largest number of outer iterations in Algorithm 2.1 is set to be 100 and the largest number of iterations in the CG method is set to be n^2.

For comparison purposes, we repeat our experiments over 10 different starting points. In our numerical tests, we use the following notations.

- '`CT.`' : the averaged total computing time in seconds;
- `IT.`' : the averaged number of iterations;
- '`NF.`': the averaged number of function evaluations;
- '`NCG.`': the averaged number of inner CG iterations;
- '`Res.`' : the averaged residual $\|G(X^k)\|_F$;
- '`grad.`': the averaged residual $\|\mathrm{grad}\ \phi(X^k)\|$, where

$$\phi(S,Q,V) := \frac{1}{2}\|G(S,Q,V)\|_F^2, \quad (S,Q,V) \in \mathbb{R}^{n\times n} \times \mathcal{O}(n) \times \mathcal{V}.$$

Table: Numerical results of Example 4.1.

Alg.	n	CT.	IT.	NF.	NCG.	Res.	grad.
GPRP	50	0.3731 s	352.4	357.7		9.7×10^{-9}	6.0×10^{-8}
	100	2.4659 s	753.7	760.1		9.8×10^{-9}	9.1×10^{-8}
	150	8.3070 s	1225.9	1232.9		9.9×10^{-9}	1.1×10^{-7}
	200	17.208 s	1492.9	1500.9		9.9×10^{-9}	1.2×10^{-7}
RFR	50	0.3781 s	308.6	313.6		9.4×10^{-9}	7.0×10^{-8}
	100	1.7944 s	523.0	529.0		9.7×10^{-9}	1.1×10^{-7}
	150	6.7529 s	951.2	958.2		9.8×10^{-9}	1.1×10^{-7}
	200	14.094 s	1163.9	1170.9		9.9×10^{-9}	1.4×10^{-7}
ALP	50	2.1249 s	388.5			7.4×10^{-9}	
	100	17.053 s	913.6			4.2×10^{-9}	
	150	132.36 s	3292.5			5.1×10^{-9}	
	200	1349.5 s	19111			0.1681^{*}	
Alg. 2.1	50	0.0550 s	6.0	7.0	52.5	1.8×10^{-11}	2.3×10^{-10}
	100	0.3634 s	6.8	7.8	80.6	1.2×10^{-9}	3.2×10^{-8}
	150	0.9421 s	7.0	8.0	98.6	3.9×10^{-13}	1.2×10^{-11}
	200	1.7102 s	7.0	8.0	105.3	1.8×10^{-11}	6.1×10^{-10}

Table: Numerical results of Example 4.1.

Alg.	n	CT.	IT.	NF.	NCG.	Res.	grad.
GPRP	400	04 m 17 s	3074	3083		9.9×10^{-9}	9.0×10^{-8}
	600	14 m 11 s	3641	3650		9.9×10^{-9}	2.5×10^{-7}
	800	41 m 47 s	5108	5118		9.9×10^{-9}	4.2×10^{-7}
	1000	01 h 14 m 23 s	5420	5430		9.9×10^{-9}	5.0×10^{-7}
RFR	400	04 m 36 s	3324	3332		9.9×10^{-9}	1.9×10^{-7}
	600	24 m 23 s	5840	5849		9.9×10^{-9}	1.9×10^{-7}
	800	01 h 07 m 53 s	8087	8096		9.9×10^{-9}	2.5×10^{-7}
	1000	02 h 38 m 59 s	11157	11167		9.9×10^{-9}	2.6×10^{-7}
Alg. 2.1	400	23.7 s	8.0	9.0	166.9	3.2×10^{-13}	7.1×10^{-12}
	600	01 m 05 s	8.0	9.0	169.5	3.4×10^{-12}	2.7×10^{-10}
	800	02 m 23 s	8.0	9.0	162.4	8.3×10^{-9}	1.7×10^{-6}
	1000	07 m 09 s	9.0	10.0	229.3	1.2×10^{-12}	4.2×10^{-11}

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, 2008.

[2] X. Chen and D. L. Liu, *Isospectral flow method for nonnegative inverse eigenvalue problem with prescribed structure*, J. Comput. Appl. Math., 235 (2011), pp. 3990–4002.

[3] M. T. Chu, F. Diele, and I. Sgura, *Gradient flow method for matrix completion with prescribed eigenvalues*, Linear Algebra Appl., 379 (2004), pp. 85–112.

[4] M. T. Chu and K. R. Driessel, *Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations*, SIAM J. Math. Anal., 22 (1991), pp. 1372–1387.

[5] M. T. Chu and Q. Guo, *A numerical method for the inverse stochastic spectrum problem*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 1027–1039.

[6] M. M. LIN, *Fast recursive algorithm for constructing nonnegative matrices with prescribed real eigenvalues*, Appl. Math. Comput., 256 (2015), pp. 582–590.

[7] D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, 1969.

[8] R. ORSI, *Numerical methods for solving inverse eigenvalue problems for nonnegative matrices*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 190–212.

[9] G. W. SOULES, *Constructing symmetric nonnegative matrices*, Linear and Multilinear Alg., 13 (1983), pp. 241–251.

[10] T. T. YAO, Z. J. BAI, Z. ZHAO, AND W. K. CHING, *A Riemannian Fletcher–Reeves conjugate gradient method for doubly stochastic inverse eigenvalue problems*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 215–234,

[11] Z. ZHAO, X. Q. JIN, AND Z. J. BAI, *A geometric nonlinear conjugate gradient method for stochastic inverse eigenvalue problems*, SIAM J. Numer. Anal., 54 (2016), 2015–2035.

Thank You for Your Attention!