Factorization of completely positive matrices using iterative projected gradient steps

Radu Ioan Boţ
(the talk is based on a joint work with D.-K. Nguyen)

University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
1090 Vienna
Austria
www.mat.univie.ac.at/~rabot

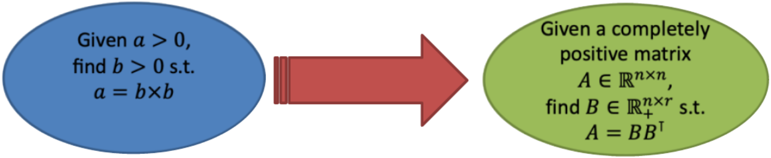
Optimization: Theory, Algorithms, Applications Lecture Series
The Fields Institute and University of Waterloo
September 9, 2021

universität wien

FWF Der Wissenschaftsfonds.

Contents

Problem formulation



A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called completely positive if there exists $B \in \mathbb{R}_+^{n \times r}$, an entrywise nonnegative matrix in $\mathbb{R}^{n \times r}$, such that

$$A = BB^T.$$

Let

$$\mathcal{CP}_n := \left\{ A \in \mathbb{R}^{n \times n} : A = BB^T \text{ where } B \in \mathbb{R}_+^{n \times r} \text{ and } r \in \mathbb{N} \right\}$$

denote the set of $n \times n$ completely positive matrices.

In this talk we will ...

- address the nonnegative facrotization of a completely positive matrix by formulating it as an optimization problem;
- propose a first-order optimization algorithm for solving the resulting optimization problem and investigate its convergence behaviour;
- validate and test the theoretical findings in various numerical experiments.

The value of r

- The factorization of a completely positive matrix is never unique (one can "enlarge" the factor B by adding zero columns).
- Dickinson (EJLA, 2010): For the matrix

$$A := \begin{pmatrix} 18 & 9 & 9 \\ 9 & 18 & 9 \\ 9 & 9 & 18 \end{pmatrix}$$

one has $A = B_i B_i^T$, $i = 1, \dots, 4$, for

$$B_1 := \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}, \qquad B_2 := \begin{pmatrix} 3 & 3 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix},$$

$$B_3 := \begin{pmatrix} 3 & 3 & 0 \\ 3 & 0 & 3 \\ 0 & 3 & 3 \end{pmatrix}, \qquad B_4 := \begin{pmatrix} -1.2030 & 2.1337 & 3.4641 \\ 2.4494 & 0.0250 & 3.4641 \\ -1.2463 & -2.1087 & 3.4641 \end{pmatrix}.$$

cp-rank and cp$^+$-rank

Let $A \in \mathbb{R}^{n\times n}$.

- The cp-rank of A: $\operatorname{cpr}(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}_+^{n\times r}, A = BB^T\right\}$.
- The cp$^+$-rank of A: $\operatorname{cpr}^+(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}_{++}^{n\times r}, A = BB^T\right\}$, where $\mathbb{R}_{++}^{n\times r}$ denotes the set of matrices in $\mathbb{R}_+^{n\times r}$ with at least one column with positive entries.

- We consider on $\mathbb{R}^{n\times n}$ the Frobenius inner product and the Frobenius norm defined for $X, Y \in \mathbb{R}^{n\times n}$ by

$$\langle X, Y\rangle := \operatorname{trace}\left(X^T Y\right) \text{ and } \|X\|_{\mathsf{F}} := \sqrt{\langle X, X\rangle} = \sqrt{\operatorname{trace}\left(X^T X\right)}, \text{ respectively.}$$

The interior of $\mathcal{CP}_n$ (Dickinson (EJLA, 2010))

- $\operatorname{int}(\mathcal{CP}_n) = \left\{A \in \mathbb{R}^{n\times n} : \operatorname{rank}(A) = n, A = BB^T, B \in \mathbb{R}_{++}^{n\times r}\right\}$

Upper bounds for the cp-rank and the cp$^+$-rank (Bomze, Dickinson, Still (LAA, 2015))

- If $A \in \mathcal{CP}_n$, then $\operatorname{cpr}(A) \le \operatorname{cp}_n := \begin{cases} n & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 4 & \text{for } n \ge 5. \end{cases}$
- If $A \in \operatorname{int}(\mathcal{CP}_n)$, then $\operatorname{cpr}^+(A) \le \operatorname{cp}_n^+ := \begin{cases} n+1 & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 3 & \text{for } n \ge 5. \end{cases}$

cp-rank and cp$^+$-rank

Let $A \in \mathbb{R}^{n\times n}$.

- The cp-rank of A: $\operatorname{cpr}(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}^{n\times r}_+, A = BB^T\right\}$.

- The cp$^+$-rank of A: $\operatorname{cpr}^+(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}^{n\times r}_{++}, A = BB^T\right\}$, where $\mathbb{R}^{n\times r}_{++}$ denotes the set of matrices in $\mathbb{R}^{n\times r}_+$ with at least one column with positive entries.

- We consider on $\mathbb{R}^{n\times n}$ the Frobenius inner product and the Frobenius norm defined for $X, Y \in \mathbb{R}^{n\times n}$ by

$$\langle X, Y\rangle := \operatorname{trace}\left(X^T Y\right) \text{ and } \|X\|_F := \sqrt{\langle X, X\rangle} = \sqrt{\operatorname{trace}\left(X^T X\right)}, \text{ respectively.}$$

The interior of $\mathcal{CP}_n$ (Dickinson (EJLA, 2010))

- $\operatorname{int}(\mathcal{CP}_n) = \left\{A \in \mathbb{R}^{n\times n} : \operatorname{rank}(A) = n, A = BB^T, B \in \mathbb{R}^{n\times r}_{++}\right\}$

Upper bounds for the cp-rank and the cp$^+$-rank (Bomze, Dickinson, Still (LAA, 2015))

- If $A \in \mathcal{CP}_n$, then $\operatorname{cpr}(A) \le \operatorname{cp}_n := \begin{cases} n & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 4 & \text{for } n \ge 5. \end{cases}$

- If $A \in \operatorname{int}(\mathcal{CP}_n)$, then $\operatorname{cpr}^+(A) \le \operatorname{cp}^+_n := \begin{cases} n+1 & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 3 & \text{for } n \ge 5. \end{cases}$

cp-rank and cp$^+$-rank

Let $A \in \mathbb{R}^{n\times n}$.

- The cp-rank of A: $\operatorname{cpr}(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}_+^{n\times r}, A = BB^T\right\}$.
- The cp$^+$-rank of A: $\operatorname{cpr}^+(A) := \inf\left\{r \in \mathbb{N} : \exists B \in \mathbb{R}_{++}^{n\times r}, A = BB^T\right\}$, where $\mathbb{R}_{++}^{n\times r}$ denotes the set of matrices in $\mathbb{R}_+^{n\times r}$ with at least one column with positive entries.

- We consider on $\mathbb{R}^{n\times n}$ the Frobenius inner product and the Frobenius norm defined for $X, Y \in \mathbb{R}^{n\times n}$ by

$$\langle X, Y\rangle := \operatorname{trace}\left(X^TY\right) \text{ and } \|X\|_{\mathsf{F}} := \sqrt{\langle X, X\rangle} = \sqrt{\operatorname{trace}\left(X^TX\right)}, \text{ respectively.}$$

The interior of $\mathcal{CP}_n$ (Dickinson (EJLA, 2010))

- $\operatorname{int}(\mathcal{CP}_n) = \left\{A \in \mathbb{R}^{n\times n} : \operatorname{rank}(A) = n, A = BB^T, B \in \mathbb{R}_{++}^{n\times r}\right\}$

Upper bounds for the cp-rank and the cp$^+$-rank (Bomze, Dickinson, Still (LAA, 2015))

- If $A \in \mathcal{CP}_n$, then $\operatorname{cpr}(A) \leq \operatorname{cp}_n := \begin{cases} n & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 4 & \text{for } n \geq 5. \end{cases}$
- If $A \in \operatorname{int}(\mathcal{CP}_n)$, then $\operatorname{cpr}^+(A) \leq \operatorname{cp}_n^+ := \begin{cases} n+1 & \text{for } n \in \{2,3,4\}, \\ \frac{1}{2}n(n+1) - 3 & \text{for } n \geq 5. \end{cases}$

The nonnegative factorization of completely positive matrices via projection onto the orthogonal set $\mathbb{O}_r$

- In (Groetzner, Dür (LAA, 2020)) the factorization problem has been formulated as a feasibility problem:
- For a given matrix $A \in \mathbb{R}^{n\times n}$, let $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$.
- The aim is

$$\text{to find a } r \times r \text{ square matrix } Q \text{ such that } Q \in \mathcal{P}(B) \cap \mathbb{O}_r,$$

where

- ◇ $\mathcal{P}(B) := \{X \in \mathbb{R}^{r\times r} : BX \in \mathbb{R}^{n\times r}_+\}$ is the polyhedral cone associated to B;
- ◇ $\mathbb{O}_r := \left\{X \in \mathbb{R}^{r\times r} : XX^T = X^TX = \mathbb{I}_r\right\}$ is the set of $r \times r$ orthogonal matrices.

- Notice that, for $B_1, B_2 \in \mathbb{R}^{n\times r}$ it holds $B_1B_1^T = B_2B_2^T$ if and only if there exists $Q \in \mathbb{O}_r$ such that $B_1Q = B_2$.

The Method of Alternating Projections (Groetzner, Dür (LAA, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} P_k & := \text{Pr}_{\mathcal{P}(B)}(Q_k), \\ Q_{k+1} & \in \text{Pr}_{\mathbb{O}_r}(P_k). \end{cases} \qquad \text{(MAP)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

The nonnegative factorization of completely positive matrices via projection onto the orthogonal set $\mathbb{O}_r$

- In (Groetzner, Dür (LAA, 2020)) the factorization problem has been formulated as a feasibility problem:
- For a given matrix $A \in \mathbb{R}^{n\times n}$, let $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$.
- The aim is

to find a $r \times r$ square matrix Q such that $Q \in \mathcal{P}(B) \cap \mathbb{O}_r$,

where

◇ $\mathcal{P}(B) := \{X \in \mathbb{R}^{r\times r} : BX \in \mathbb{R}_+^{n\times r}\}$ is the polyhedral cone associated to B;

◇ $\mathbb{O}_r := \left\{X \in \mathbb{R}^{r\times r} : XX^T = X^T X = \mathbb{I}_r\right\}$ is the set of $r \times r$ orthogonal matrices.

- Notice that, for $B_1, B_2 \in \mathbb{R}^{n\times r}$ it holds $B_1 B_1^T = B_2 B_2^T$ if and only if there exists $Q \in \mathbb{O}_r$ such that $B_1 Q = B_2$.

The Method of Alternating Projections (Groetzner, Dür (LAA, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} P_k & := \text{Pr}_{\mathcal{P}(B)}(Q_k), \\ Q_{k+1} & \in \text{Pr}_{\mathbb{O}_r}(P_k). \end{cases} \qquad \text{(MAP)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

The Modified Method of Alternating Projections - avoids the calculation of the projection on $\mathcal{P}(B)$ (Groetzner, Dür (LAA, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n \times r}$ such that $A = BB^T$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} R_k & := \mathrm{Pr}_{\mathbb{R}_+^{n \times r}}(BQ_k), \\ \widehat{P}_k & := B^+ R_k + (\mathbb{I}_r - B^+ B) Q_k, \\ Q_{k+1} & \in \mathrm{Pr}_{\mathbb{O}_r}\left(\widehat{P}_k\right). \end{cases} \qquad \text{(ModMAP)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

A difference-of-convex approach (Chen, Pong, Tan, Zeng (JOGO, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n \times r}$ such that $A = BB^T$, a fixed stepsize $L_B > \lambda_{\max}(B^T B)$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} W_k & := \mathrm{Pr}_{\mathbb{R}_+^{n \times r}}(BQ_k), \\ Q_{k+1} & \in \mathrm{Pr}_{\mathbb{O}_r}\left(Q_k - \frac{1}{L_B} B^T (BQ_k - W_k)\right). \end{cases} \qquad \text{(SpFeasDC)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

The Modified Method of Alternating Projections - avoids the calculation of the projection on $\mathcal{P}(B)$ (Groetzner, Dür (LAA, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} R_k & := \mathrm{Pr}_{\mathbb{R}_+^{n\times r}}(BQ_k), \\ \widehat{P}_k & := B^+ R_k + (\mathbb{I}_r - B^+ B) Q_k, \\ Q_{k+1} & \in \mathrm{Pr}_{\mathbb{O}_r}(\widehat{P}_k). \end{cases} \qquad \text{(ModMAP)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

A difference-of-convex approach (Chen, Pong, Tan, Zeng (JOGO, 2020))

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.
Input: a given $B \in \mathbb{R}^{n\times r}$ such that $A = BB^T$, a fixed stepsize $L_B > \lambda_{\max}(B^T B)$ and a starting point $Q_0 \in \mathbb{O}_r$.

Main iterate:

$$(\forall k \geq 0) \begin{cases} W_k & := \mathrm{Pr}_{\mathbb{R}_+^{n\times r}}(BQ_k), \\ Q_{k+1} & \in \mathrm{Pr}_{\mathbb{O}_r}\left(Q_k - \frac{1}{L_B} B^T (BQ_k - W_k)\right). \end{cases} \qquad \text{(SpFeasDC)}$$

Output: $Q_{k+1} \in \mathbb{O}_r$ such that $A = (BQ_{k+1})(BQ_{k+1})^T$.

- One can find a matrix $B \in \mathbb{R}^{n \times r}$ such that $A = BB^T$:
  - by Cholesky decomposition, in this case B is a lower triangular matrix;
  - by spectral decomposition $A = V \Sigma V^T$, and then by setting $B := V \Sigma^{\frac{1}{2}}$.
- The projection of a matrix $P \in \mathbb{R}^{r \times r}$ onto the set $\mathbb{O}_r$ can be computed via singular value decomposition

$$P = U \Sigma V^T,$$

  in a subroutine that needs $\mathcal{O}\left(r^3\right)$ steps. Then

$$UV^T \in \mathrm{Pr}_{\mathbb{O}_r}(P).$$

The optimization model

Given a nonzero completely positive matrix $A \in \mathbb{R}^{n \times n}$, we consider the optimization problem

$$\min_{X \in \mathbb{R}^{n \times r}} \mathcal{E}(X) := \frac{1}{2} \left\| A - XX^T \right\|_{\mathsf{F}}^2 . \qquad \text{s.t. } X \in \mathcal{D} := \mathbb{R}_+^{n \times r} \cap \mathbb{B}_{\mathsf{F}}\left(0, \sqrt{\operatorname{trace}(A)}\right) \tag{P}$$

- The critical points of the objective function $\mathcal{E} + \delta_{\mathcal{D}}$ are those $X_* \in \mathbb{R}^{n \times r}$ such that

$$-\nabla \mathcal{E}(X_*) \in \mathcal{N}_{\mathcal{D}}(X_*),$$

where $\mathcal{N}_{\mathcal{D}}(X_*)$ denots the normal cone to the convex set $\mathcal{D}$ at X_*.

- The additional constraint does not restrict the generality of the problem, since, for $A \in \mathcal{CP}_n$ and $X \in \mathbb{R}^{n \times r}$ such that $A = XX^T$, it holds

$$\|X\|_{\mathsf{F}} \leq \sqrt{\operatorname{trace}(A)}.$$

- Moreover,

$$A = X_* X_*^T \text{ with } X_* \in \mathbb{R}_+^{n \times r} \iff \left[X_* \text{ solves (P) and } \min_{X \in \mathcal{D}} \mathcal{E}(X) = 0 \right].$$

The optimization model

Given a nonzero completely positive matrix $A \in \mathbb{R}^{n \times n}$, we consider the optimization problem

$$\min_{X \in \mathbb{R}^{n \times r}} \mathcal{E}(X) := \frac{1}{2} \left\| A - XX^T \right\|_{\mathsf{F}}^2 .$$
$$\text{s.t. } X \in \mathcal{D} := \mathbb{R}_+^{n \times r} \cap \mathbb{B}_{\mathsf{F}} \left(0, \sqrt{\operatorname{trace}(A)}\right) \qquad \text{(P)}$$

- The critical points of the objective function $\mathcal{E} + \delta_{\mathcal{D}}$ are those $X_* \in \mathbb{R}^{n \times r}$ such that

$$-\nabla \mathcal{E}(X_*) \in \mathcal{N}_{\mathcal{D}}(X_*),$$

where $\mathcal{N}_{\mathcal{D}}(X_*)$ denots the normal cone to the convex set $\mathcal{D}$ at X_*.

- The additional constraint does not restrict the generality of the problem, since, for $A \in \mathcal{CP}_n$ and $X \in \mathbb{R}^{n \times r}$ such that $A = XX^T$, it holds

$$\|X\|_{\mathsf{F}} \leq \sqrt{\operatorname{trace}(A)}.$$

- Moreover,

$$A = X_* X_*^T \text{ with } X_* \in \mathbb{R}_+^{n \times r} \Longleftrightarrow \left[X_* \text{ solves (P) and } \min_{X \in \mathcal{D}} \mathcal{E}(X) = 0 \right].$$

A projected gradient algorithm with relaxation and inertial parameters

RIB, D.-K. Nguyen (2021): *Factorization of completely positive matrices using iterative projected gradient steps*, Numerical Linear Algebra with Applications, DOI: 10.1002/nla.2391

Let $A \in \mathcal{CP}_n$ and r be a positive integer value.

Input:

- starting points $X_1 := X_0 \in \mathcal{D}$;
- a sequence $\{\alpha_k\}_{k\geq 1} \subseteq [0,1]$, for which we set $\alpha_+ := \sup_{k\geq 0} \alpha_k$ and
$$L_{\mathsf{F}}(\alpha_+) := 2\left[\left(3 + 8\alpha_+ + 6\alpha_+^2\right)\operatorname{trace}(A) - \lambda_{\min}(A)\right] > 0;$$
- a relaxation parameter $\rho \in (0,1]$ chosen such that
$$0 < \frac{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2} + \sqrt{L_{\mathsf{F}}(\alpha_+)}} < \rho < \frac{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}{(1+\alpha_+)\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2} - \sqrt{L_{\mathsf{F}}(\alpha_+)}}. \quad \text{(RelaxInertial)}$$

Main iterate:

$$(\forall k \geq 1) \quad \begin{cases} Y_k & := X_k + \alpha_k (X_k - X_{k-1}), \\ Z_{k+1} & := \mathsf{Pr}_{\mathcal{D}}\left(Y_k - \frac{1}{L_{\mathsf{F}}(\alpha_+)} \nabla \mathcal{E}(Y_k)\right), \\ X_{k+1} & := (1-\rho) X_k + \rho Z_{k+1}. \end{cases} \quad \text{(RIPG)}$$

Output: $X_{k+1} \in \mathcal{D}$, which provides a factorization $A = X_{k+1} X_{k+1}^T$.

- Other works addressing the interplay between relaxation and inertial parameters for convex optimization and monotone inclusions: RIB, Csetnek (SICON, 2016), Attouch, Peypouquet (MathProg, 2019), RIB, Sedlmayer, Vuong (ArXiv, 2020)

Useful facts

- For $X \in \mathbb{R}^{n \times r}$, it holds (Bauschke, Bui, Wang, (SIOPT, 2018))

$$\mathrm{Pr}_{\mathcal{D}}(X) := \frac{\sqrt{\operatorname{trace}(A)}}{\max\left\{\left\|[X]_+\right\|_{\mathsf{F}}, \sqrt{\operatorname{trace}(A)}\right\}}[X]_+,$$

where $[X]_+ := \max\{X, 0\}$ and the $\max$ operator is understood entrywise.

- For $X, Y \in \mathbb{R}^{n \times r}$, it holds

$$-\|A\|_2 \cdot \|X - Y\|_{\mathsf{F}}^2 \leq \mathcal{E}(X) - \mathcal{E}(Y) - \langle \nabla \mathcal{E}(Y), X - Y \rangle \leq \frac{L(X,Y)}{2}\|X - Y\|_{\mathsf{F}}^2,$$

where

$$L(X,Y) := 2\left(\|Y\|_2^2 - \lambda_{\min}(A)\right) + \left(\|X\|_2 + \|Y\|_2\right)^2.$$

- For every $k \geq 1$, we have
  - $X_{k+1} \in \mathcal{D}$ and $\|Y_k\|_{\mathsf{F}} \leq (1 + 2\alpha_+)\sqrt{\operatorname{trace}(A)}$;
  - $L(Z_{k+1}, Y_k) \leq L_{\mathsf{F}}(\alpha_+) = 2\left[\left(3 + 8\alpha_+ + 6\alpha_+^2\right)\operatorname{trace}(A) - \lambda_{\min}(A)\right]$.

Useful facts

- For $X \in \mathbb{R}^{n \times r}$, it holds (Bauschke, Bui, Wang, (SIOPT, 2018))

$$\mathsf{Pr}_{\mathcal{D}}(X) := \frac{\sqrt{\operatorname{trace}(A)}}{\max\left\{\left\|[X]_+\right\|_{\mathsf{F}}, \sqrt{\operatorname{trace}(A)}\right\}}[X]_+,$$

where $[X]_+ := \max\{X, 0\}$ and the $\max$ operator is understood entrywise.

- For $X, Y \in \mathbb{R}^{n \times r}$, it holds

$$-\|A\|_2 \cdot \|X - Y\|_{\mathsf{F}}^2 \leq \mathcal{E}(X) - \mathcal{E}(Y) - \langle \nabla \mathcal{E}(Y), X - Y \rangle \leq \frac{L(X,Y)}{2}\|X - Y\|_{\mathsf{F}}^2,$$

where

$$L(X,Y) := 2\left(\|Y\|_2^2 - \lambda_{\min}(A)\right) + \left(\|X\|_2 + \|Y\|_2\right)^2.$$

- For every $k \geq 1$, we have
  - $X_{k+1} \in \mathcal{D}$ and $\|Y_k\|_{\mathsf{F}} \leq (1 + 2\alpha_+)\sqrt{\operatorname{trace}(A)}$;
  - $L(Z_{k+1}, Y_k) \leq L_{\mathsf{F}}(\alpha_+) = 2\left[\left(3 + 8\alpha_+ + 6\alpha_+^2\right)\operatorname{trace}(A) - \lambda_{\min}(A)\right]$.

Useful facts

- For $X \in \mathbb{R}^{n\times r}$, it holds (Bauschke, Bui, Wang, (SIOPT, 2018))

$$\mathsf{Pr}_{\mathcal{D}}(X) := \frac{\sqrt{\operatorname{trace}(A)}}{\max\left\{\left\|[X]_+\right\|_{\mathsf{F}}, \sqrt{\operatorname{trace}(A)}\right\}}[X]_+,$$

where $[X]_+ := \max\{X, 0\}$ and the $\max$ operator is understood entrywise.

- For $X, Y \in \mathbb{R}^{n\times r}$, it holds

$$-\|A\|_2 \cdot \|X - Y\|_{\mathsf{F}}^2 \le \mathcal{E}(X) - \mathcal{E}(Y) - \langle \nabla\mathcal{E}(Y), X - Y\rangle \le \frac{L(X,Y)}{2}\|X - Y\|_{\mathsf{F}}^2,$$

where

$$L(X,Y) := 2\left(\|Y\|_2^2 - \lambda_{\min}(A)\right) + \left(\|X\|_2 + \|Y\|_2\right)^2.$$

- For every $k \ge 1$, we have
  - $X_{k+1} \in \mathcal{D}$ and $\|Y_k\|_{\mathsf{F}} \le (1 + 2\alpha_+)\sqrt{\operatorname{trace}(A)}$;
  - $L(Z_{k+1}, Y_k) \le L_{\mathsf{F}}(\alpha_+) = 2\left[\left(3 + 8\alpha_+ + 6\alpha_+^2\right)\operatorname{trace}(A) - \lambda_{\min}(A)\right]$.

The decreasing property

For every $k \geq 1$, it holds

$$(\mathcal{E} + \delta_{\mathcal{D}})(Z_{k+1}) + \left(\frac{L_{\mathsf{F}}(\alpha_+) - \left(L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2\right)\gamma}{2} + \frac{\tau}{2} \right) \|X_{k+1} - X_k\|_{\mathsf{F}}^2$$
$$\leq (\mathcal{E} + \delta_{\mathcal{D}})(Z_k) + \frac{\tau}{2} \|X_k - X_{k-1}\|_{\mathsf{F}}^2,$$

where

$\gamma := \max\left\{ \left(\frac{1}{\rho} - 1\right)^2, \left(1 + \alpha_+ - \frac{1}{\rho}\right)^2 \right\}$ and $\tau := \frac{(1-\rho) L_{\mathsf{F}}(\alpha_+)}{\rho} + \left(L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2\right)\gamma$.

- It holds $L_{\mathsf{F}}(\alpha_+) - \left(L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2\right)\gamma > 0$.

The energy function

For a given $\tau \geq 0$, we consider the following energy function

$$\Psi_\tau : \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} \to \mathbb{R} \cup \{+\infty\}, \Psi_\tau(Z, X) := (\mathcal{E} + \delta_{\mathcal{D}})(Z) + \frac{\rho^2\tau}{2} \|Z - X\|_{\mathsf{F}}^2.$$

- For every $k \geq 2$ it holds

$$\Psi_\tau(Z_{k+1}, X_k) + \frac{L_{\mathsf{F}}(\alpha_+) - \left(L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2\right)\gamma}{2} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 \leq \Psi_\tau(Z_k, X_{k-1})$$

  - If $\tau = 0$, which corresponds to the case when $\rho = 1$ and $\alpha_+ = 0$, in which case RIPG becomes the projected gradient algorithm, then

$$\Psi_\tau(Z, X) = (\mathcal{E} + \delta_{\mathcal{D}})(Z) \quad \forall (Z, X) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r}.$$

    Thus

$$Z_* \in \operatorname{crit}(\mathcal{E} + \delta_{\mathcal{D}}) \text{ and } X_* \in \mathbb{R}^{n\times r} \Leftrightarrow (Z_*, X_*) \in \operatorname{crit}\Psi_\tau.$$

  - If $\tau > 0$, then

$$X_* \in \operatorname{crit}(\mathcal{E} + \delta_{\mathcal{D}}) \text{ and } \Leftrightarrow (X_*, X_*) \in \operatorname{crit}\Psi_\tau.$$

The energy function

For a given $\tau \geq 0$, we consider the following energy function

$$\Psi_\tau : \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} \to \mathbb{R} \cup \{+\infty\}, \Psi_\tau(Z, X) := (\mathcal{E} + \delta_\mathcal{D})(Z) + \frac{\rho^2\tau}{2}\|Z - X\|_\mathsf{F}^2.$$

- For every $k \geq 2$ it holds

$$\Psi_\tau(Z_{k+1}, X_k) + \frac{L_\mathsf{F}(\alpha_+) - \left(L_\mathsf{F}(\alpha_+) + 2\|A\|_2\right)\gamma}{2}\|X_{k+1} - X_k\|_\mathsf{F}^2 \leq \Psi_\tau(Z_k, X_{k-1})$$

- If $\tau = 0$, which corresponds to the case when $\rho = 1$ and $\alpha_+ = 0$, in which case RIPG becomes the projected gradient algorithm, then

$$\Psi_\tau(Z, X) = (\mathcal{E} + \delta_\mathcal{D})(Z) \quad \forall (Z, X) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r}.$$

Thus

$$Z_* \in \operatorname{crit}(\mathcal{E} + \delta_\mathcal{D}) \text{ and } X_* \in \mathbb{R}^{n\times r} \Leftrightarrow (Z_*, X_*) \in \operatorname{crit}\Psi_\tau.$$

- If $\tau > 0$, then

$$X_* \in \operatorname{crit}(\mathcal{E} + \delta_\mathcal{D}) \text{ and } \Leftrightarrow (X_*, X_*) \in \operatorname{crit}\Psi_\tau.$$

The energy function

For a given $\tau \geq 0$, we consider the following energy function

$$\Psi_\tau : \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} \to \mathbb{R} \cup \{+\infty\}, \Psi_\tau(Z, X) := (\mathcal{E} + \delta_{\mathcal{D}})(Z) + \frac{\rho^2 \tau}{2} \|Z - X\|_{\mathsf{F}}^2 .$$

- For every $k \geq 2$ it holds

$$\Psi_\tau(Z_{k+1}, X_k) + \frac{L_{\mathsf{F}}(\alpha_+) - \left(L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2\right)\gamma}{2} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 \leq \Psi_\tau(Z_k, X_{k-1})$$

  - If $\tau = 0$, which corresponds to the case when $\rho = 1$ and $\alpha_+ = 0$, in which case RIPG becomes the projected gradient algorithm, then

$$\Psi_\tau(Z, X) = (\mathcal{E} + \delta_{\mathcal{D}})(Z) \quad \forall (Z, X) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r}.$$

    Thus

$$Z_* \in \operatorname{crit}(\mathcal{E} + \delta_{\mathcal{D}}) \text{ and } X_* \in \mathbb{R}^{n\times r} \Leftrightarrow (Z_*, X_*) \in \operatorname{crit}\Psi_\tau .$$

  - If $\tau > 0$, then

$$X_* \in \operatorname{crit}(\mathcal{E} + \delta_{\mathcal{D}}) \text{ and } \Leftrightarrow (X_*, X_*) \in \operatorname{crit}\Psi_\tau .$$

The energy function

For a given $\tau \geq 0$, we consider the following energy function

$$\Psi_\tau : \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} \to \mathbb{R} \cup \{+\infty\}, \Psi_\tau(Z, X) := (\mathcal{E} + \delta_\mathcal{D})(Z) + \frac{\rho^2 \tau}{2} \|Z - X\|_\mathsf{F}^2.$$

- For every $k \geq 2$ it holds

$$\Psi_\tau(Z_{k+1}, X_k) + \frac{L_\mathsf{F}(\alpha_+) - \left(L_\mathsf{F}(\alpha_+) + 2\|A\|_2\right)\gamma}{2} \|X_{k+1} - X_k\|_\mathsf{F}^2 \leq \Psi_\tau(Z_k, X_{k-1})$$

  - If $\tau = 0$, which corresponds to the case when $\rho = 1$ and $\alpha_+ = 0$, in which case RIPG becomes the projected gradient algorithm, then

$$\Psi_\tau(Z, X) = (\mathcal{E} + \delta_\mathcal{D})(Z) \quad \forall (Z, X) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r}.$$

    Thus

$$Z_* \in \mathrm{crit}(\mathcal{E} + \delta_\mathcal{D}) \text{ and } X_* \in \mathbb{R}^{n\times r} \Leftrightarrow (Z_*, X_*) \in \mathrm{crit}\Psi_\tau.$$

  - If $\tau > 0$, then

$$X_* \in \mathrm{crit}(\mathcal{E} + \delta_\mathcal{D}) \text{ and } \Leftrightarrow (X_*, X_*) \in \mathrm{crit}\Psi_\tau.$$

The limiting subdifferential of a proper and lower semicontinuous function $k : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, where $\mathcal{H}$ is a real finite-dimensional space

- The Fréchet (viscosity) subdifferential of h at $x \in \mathrm{dom} k$:

$$\hat{\partial} k(x) = \left\{ v \in \mathcal{H} : \liminf_{y \to x} \frac{k(y) - k(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

- The limiting (Mordukhovich) subdifferential of h at $x \in \mathrm{dom} k$:

$$\partial k(x) = \{ v \in \mathcal{H} : \exists x_n \to x, k(x_n) \to k(x) \text{ and } \exists v_n \in \hat{\partial} k(x_n), v_n \to v \text{ as } n \to +\infty \}$$

Properties of the limiting subdifferential

- if $x \in \mathcal{H}$ is a local minimizer of k, then $x \in \mathrm{crit} k := \{ z \in \mathcal{H} : 0 \in \partial k(z) \}$;
- if k is C^1 around $x \in \mathcal{H}$, then $\partial k(x) = \{ \nabla k(x) \}$;
- if k is convex, then $\partial k(x) = \{ v \in \mathcal{H} : k(y) \geq k(x) + \langle v, y - x \rangle \ \forall y \in \mathcal{H} \} \ \forall x \in \mathrm{dom} k$;
- closedness criterion: $v_n \in \partial k(x_n) \ \forall n \geq 0$, $(x_n, v_n) \to (x, v)$ and $k(x_n) \to k(x)$ as $n \to +\infty$, then $v \in \partial k(x)$;
- sum formula: if $l : \mathcal{H} \to \mathbb{R}$ is C^1, then $\partial(k + l)(x) = \partial k(x) + \nabla l(x)$ for all $x \in \mathcal{H}$.

The limiting subdifferential of a proper and lower semicontinuous function $k : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, where $\mathcal{H}$ is a real finite-dimensional space

- The Fréchet (viscosity) subdifferential of h at $x \in \mathrm{dom} k$:

$$\hat{\partial} k(x) = \left\{ v \in \mathcal{H} : \liminf_{y \to x} \frac{k(y) - k(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

- The limiting (Mordukhovich) subdifferential of h at $x \in \mathrm{dom} k$:

$$\partial k(x) = \{ v \in \mathcal{H} : \exists x_n \to x, k(x_n) \to k(x) \text{ and } \exists v_n \in \hat{\partial} k(x_n), v_n \to v \text{ as } n \to +\infty \}$$

Properties of the limiting subdifferential

- if $x \in \mathcal{H}$ is a local minimizer of k, then $x \in \mathrm{crit} k := \{ z \in \mathcal{H} : 0 \in \partial k(z) \}$;
- if k is C^1 around $x \in \mathcal{H}$, then $\partial k(x) = \{ \nabla k(x) \}$;
- if k is convex, then $\partial k(x) = \{ v \in \mathcal{H} : k(y) \geq k(x) + \langle v, y - x \rangle \ \forall y \in \mathcal{H} \} \ \forall x \in \mathrm{dom} k$;
- closedness criterion: $v_n \in \partial k(x_n) \ \forall n \geq 0$, $(x_n, v_n) \to (x, v)$ and $k(x_n) \to k(x)$ as $n \to +\infty$, then $v \in \partial k(x)$;
- sum formula: if $l : \mathcal{H} \to \mathbb{R}$ is C^1, then $\partial(k + l)(x) = \partial k(x) + \nabla l(x)$ for all $x \in \mathcal{H}$.

Cluster points are critical points

1. The sequence $\{\Psi_\tau(Z_k, X_{k-1})\}_{k\geq 2}$ is monotonically decreasing and convergent;
2. It holds $\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 < +\infty$, thus $X_{k+1} - X_k \to 0$ as $k \to +\infty$, and so $X_{k+1} - Y_k \to 0$ and $Z_{k+1} - Y_k \to 0$ as $k \to +\infty$, hence the sequences $\{X_k\}_{k\geq 0}$, $\{Y_k\}_{k\geq 1}$ and $\{Z_k\}_{k\geq 2}$ have the same cluster points.

Let $\Omega := \Omega\left(\{(Z_k, X_{k-1})\}_{k\geq 2}\right)$ be the set of cluster points of the sequence $\{(Z_k, X_{k-1})\}_{k\geq 2}$ The following statements are true:

- $\Omega \subseteq \operatorname{crit}\Psi_\tau = \{(X_*, X_*) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} : X_* \in \operatorname{crit}\Psi\}$;
- it holds $\lim_{k\to+\infty} \operatorname{dist}[(Z_k, X_{k-1}), \Omega] = 0$;
- the set Ω is nonempty, connected and compact;
- the function Ψ_τ takes on Ω the value $\Psi_* := \lim_{k\to+\infty} \Psi_\tau(Z_k, X_{k-1})$.

Subsequence convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. Then every cluster point of $\{X_k\}_{k\geq 0}$ is a critical point of $\mathcal{E} + \delta_{\mathcal{D}}$.

Cluster points are critical points

1. The sequence $\{\Psi_\tau(Z_k, X_{k-1})\}_{k\geq 2}$ is monotonically decreasing and convergent;
2. It holds $\sum_{k\geq 0} \|X_{k+1} - X_k\|_\mathsf{F}^2 < +\infty$, thus $X_{k+1} - X_k \to 0$ as $k \to +\infty$, and so $X_{k+1} - Y_k \to 0$ and $Z_{k+1} - Y_k \to 0$ as $k \to +\infty$, hence the sequences $\{X_k\}_{k\geq 0}$, $\{Y_k\}_{k\geq 1}$ and $\{Z_k\}_{k\geq 2}$ have the same cluster points.

Let $\Omega := \Omega\left(\{(Z_k, X_{k-1})\}_{k\geq 2}\right)$ be the set of cluster points of the sequence $\{(Z_k, X_{k-1})\}_{k\geq 2}$ The following statements are true:

- $\Omega \subseteq \operatorname{crit}\Psi_\tau = \{(X_*, X_*) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} : X_* \in \operatorname{crit}\Psi\}$;
- it holds $\lim_{k\to+\infty} \operatorname{dist}[(Z_k, X_{k-1}), \Omega] = 0$;
- the set Ω is nonempty, connected and compact;
- the function Ψ_τ takes on Ω the value $\Psi_* := \lim_{k\to+\infty} \Psi_\tau(Z_k, X_{k-1})$.

Subsequence convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. Then every cluster point of $\{X_k\}_{k\geq 0}$ is a critical point of $\mathcal{E} + \delta_\mathcal{D}$.

Cluster points are critical points

1. The sequence $\{\Psi_\tau(Z_k, X_{k-1})\}_{k\geq 2}$ is monotonically decreasing and convergent;
2. It holds $\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 < +\infty$, thus $X_{k+1} - X_k \to 0$ as $k \to +\infty$, and so $X_{k+1} - Y_k \to 0$ and $Z_{k+1} - Y_k \to 0$ as $k \to +\infty$, hence the sequences $\{X_k\}_{k\geq 0}$, $\{Y_k\}_{k\geq 1}$ and $\{Z_k\}_{k\geq 2}$ have the same cluster points.

Let $\Omega := \Omega\left(\{(Z_k, X_{k-1})\}_{k\geq 2}\right)$ be the set of cluster points of the sequence $\{(Z_k, X_{k-1})\}_{k\geq 2}$ The following statements are true:

- $\Omega \subseteq \operatorname{crit}\Psi_\tau = \{(X_*, X_*) \in \mathbb{R}^{n\times r} \times \mathbb{R}^{n\times r} : X_* \in \operatorname{crit}\Psi\}$;
- it holds $\lim_{k\to+\infty} \operatorname{dist}\left[(Z_k, X_{k-1}), \Omega\right] = 0$;
- the set Ω is nonempty, connected and compact;
- the function Ψ_τ takes on Ω the value $\Psi_* := \lim_{k\to+\infty} \Psi_\tau(Z_k, X_{k-1})$.

Subsequence convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. Then every cluster point of $\{X_k\}_{k\geq 0}$ is a critical point of $\mathcal{E} + \delta_{\mathcal{D}}$.

The Kurdyka-Łojasiewicz property

Let $k : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous. The function k is said to have the Kurdyka-Łojasiewicz (KL) property at $x \in \operatorname{dom}\partial k = \{z \in \mathcal{H} : \partial k(z) \neq \emptyset\}$ if there exist

- $\eta \in (0, +\infty]$;
- a neighborhood U of x;
- a concave and continuous function $\varphi : [0, \eta) \to [0, +\infty)$ such that $\varphi(0) = 0$, φ is C^1 on $(0, \eta)$ and $\varphi'(s) > 0$ for every $s \in (0, \eta)$

such that

$$\varphi'(k(y) - k(x))\operatorname{dist}(0, \partial k(y)) = \varphi'(k(y) - k(x)) \inf\{\|v\| : v \in \partial k(y)\} \geq 1 \qquad \text{(KL)}$$

for every

$$y \in U \cap \{z \in \mathcal{H} : k(x) < k(z) < k(x) + \eta\}.$$

If k has the KL property at every point in $\operatorname{dom}\partial k$, then k is called KL function.

- The KL property is satisfied at every noncritical point $x \in \operatorname{dom}\partial k$ of k.

The Kurdyka-Łojasiewicz property

Let $k : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be proper and lower semicontinuous. The function k is said to have the Kurdyka-Łojasiewicz (KL) property at $x \in \mathrm{dom}\partial k = \{z \in \mathcal{H} : \partial k(z) \neq \emptyset\}$ if there exist

- $\eta \in (0, +\infty]$;
- a neighborhood U of x;
- a concave and continuous function $\varphi : [0, \eta) \to [0, +\infty)$ such that $\varphi(0) = 0$, φ is C^1 on $(0, \eta)$ and $\varphi'(s) > 0$ for every $s \in (0, \eta)$

such that

$$\varphi'(k(y) - k(x))\mathrm{dist}(0, \partial k(y)) = \varphi'(k(y) - k(x)) \inf\{\|v\| : v \in \partial k(y)\} \geq 1 \qquad \text{(KL)}$$

for every

$$y \in U \cap \{z \in \mathcal{H} : k(x) < k(z) < k(x) + \eta\}.$$

If k has the KL property at every point in $\mathrm{dom}\partial k$, then k is called KL function.

- The KL property is satisfied at every noncritical point $x \in \mathrm{dom}\partial k$ of k.

If k is C^1 around x, then (KL) becomes

$$\varphi'(k(y) - k(x))\|\nabla k(y)\| = \|\nabla(\varphi \circ (k - k(x)))(y)\| \geq 1 \qquad \text{(smoothKL)}$$

for every

$$y \in U \cap \{z \in \mathcal{H} : k(x) < k(z) < k(x) + \eta\}.$$

Łojasiewicz (1963)

If $k : \mathcal{H} \to \mathbb{R}$ is a real-analytic function and $x \in \mathcal{H}$ a critical point, then there exist $\theta \in [1/2, 1)$ and $C, \varepsilon > 0$ such that (Łojasiewicz property)

$$|k(y) - k(x)|^\theta \leq C\|\nabla k(y)\| \text{ for every } y \in \mathcal{H} \text{ with } \|y - x\| < \varepsilon.$$

Thus, (smoothKL) is fulfilled for $\varphi(s) = \frac{1}{1-\theta}Cs^{1-\theta}$ and every

$$y \in B(x, \varepsilon) \cap \{z \in \mathcal{H} : k(x) < k(z) < +\infty\}.$$

- the Kurdyka-Łojasiewicz property: Kurdyka (Ann. I. Fourier, 1998); Bolte, Daniilidis, Lewis (SIOPT, 2006); Bolte, Daniilidis, Lewis, Shiota (SIOPT, 2007); Bolte, Daniilidis, Ley, Mazet (TAMS, 2010)

If k is C^1 around x, then (KL) becomes

$$\varphi'(k(y) - k(x))\|\nabla k(y)\| = \|\nabla(\varphi \circ (k - k(x)))(y)\| \geq 1 \qquad \text{(smoothKL)}$$

for every

$$y \in U \cap \{z \in \mathcal{H} : k(x) < k(z) < k(x) + \eta\}.$$

Łojasiewicz (1963)

If $k : \mathcal{H} \to \mathbb{R}$ is a real-analytic function and $x \in \mathcal{H}$ a critical point, then there exist $\theta \in [1/2, 1)$ and $C, \varepsilon > 0$ such that (Łojasiewicz property)

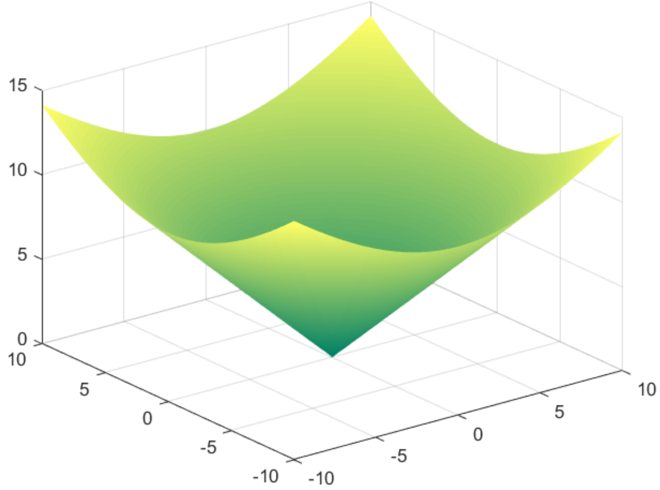
$$|k(y) - k(x)|^\theta \leq C\|\nabla k(y)\| \text{ for every } y \in \mathcal{H} \text{ with } \|y - x\| < \varepsilon.$$

Thus, (smoothKL) is fulfilled for $\varphi(s) = \frac{1}{1-\theta}Cs^{1-\theta}$ and every

$$y \in B(x, \varepsilon) \cap \{z \in \mathcal{H} : k(x) < k(z) < +\infty\}.$$

- the Kurdyka-Łojasiewicz property: Kurdyka (Ann. I. Fourier, 1998); Bolte, Daniilidis, Lewis (SIOPT, 2006); Bolte, Daniilidis, Lewis, Shiota (SIOPT, 2007); Bolte, Daniilidis, Ley, Mazet (TAMS, 2010)

Kurdyka-Łojasiewicz (KL) property



Examples of KL functions

- semi-algebraic functions, i.e., functions having as graph semi-algebraic sets, namely, sets of the form

$$\bigcup_{j=1}^{p} \bigcap_{i=1}^{q} \{u \in \mathbb{R}^m : g_{ij}(u) = 0 \text{ and } h_{ij}(u) < 0\},$$

  where $g_{ij}, h_{ij} : \mathbb{R}^m \to \mathbb{R}$ are polynomial functions;
- real polynomial functions;
- indicator functions of semi-algebraic sets;
- finite sums and product of semi-algebraic functions;
- compositions of semi-algebraic functions;
- $\|\cdot\|_p$ for $p \in \mathbb{Q}$ (including the case $p = 0$);
- convex functions fulfilling a certain growth condition;
- uniformly convex functions.

Convergence of the iterates

Global convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. The sequence $\{X_k\}_{k\geq 0}$ converges to a critical point of $\mathcal{E} + \delta_{\mathcal{D}}$.

- Since Ψ_τ is semi-algebraic, it fulfills the Kurdyka - Łojasiewicz property. This can be used to show that

$$\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 < +\infty.$$

implies

$$\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}} < +\infty.$$

In other words, $\{X_k\}_{k\geq 0}$ is a Cauchy sequence, hence it is convergent.

Convergence of the iterates

Global convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. The sequence $\{X_k\}_{k\geq 0}$ converges to a critical point of $\mathcal{E} + \delta_{\mathcal{D}}$.

▸ Since Ψ_τ is semi-algebraic, it fulfills the Kurdyka - Łojasiewicz property. This can be used to show that

$$\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}}^2 < +\infty.$$

implies

$$\sum_{k\geq 0} \|X_{k+1} - X_k\|_{\mathsf{F}} < +\infty.$$

In other words, $\{X_k\}_{k\geq 0}$ is a Cauchy sequence, hence it is convergent.

Rates of convergence

Let $\{X_k\}_{k\geq 0}$ be the sequence generated by RIPG. Let $X_* \in \operatorname{int} \mathcal{D}$ be the critical point of $\mathcal{E} + \delta_{\mathcal{D}}$ to which the sequence $\{X_k\}_{k\geq 0}$ converges as $k \to +\infty$. Then there exists $k_1 \geq 2$ such that the following statements are true:

- if $\theta = 0$, then $\{\mathcal{E}(Z_k) - \Psi_*\}_{k\geq 2}$ and $\{X_k\}_{k\geq 0}$ converge in finitely many steps;
- if $\theta \in (0, 1/2]$, then there exist $C_1', C_2' > 0$ and $Q_1, Q_2 \in [0, 1)$ such that

$$0 \leq \mathcal{E}(Z_k) - \Psi_* \leq C_1' Q_1^k \text{ and } \|X_k - X_*\|_{\mathsf{F}} \leq C_2' Q_2^k;$$

- if $\theta \in (1/2, 1)$, then there exist $C_3', C_4' > 0$ such that

$$0 \leq \mathcal{E}(Z_k) - \Psi_* \leq C_3' (k-1)^{-\frac{1}{2\theta-1}} \text{ and } \|X_k - X_*\|_{\mathsf{F}} \leq C_4' (k-1)^{-\frac{1-\theta}{2\theta-1}}.$$

Some particular cases of RIPG

Relaxed projected gradient algorithm

Choosing $\alpha_k = 0$ for all $k \geq 1$, RIPG reduces to the relaxed projected gradient algorithm

$$(\forall k \geq 1) \begin{cases} Z_{k+1} & := \mathsf{Pr}_{\mathcal{D}} \left(X_k - \dfrac{1}{L_{\mathsf{F}}(0)} \nabla \mathcal{E}\left(X_k\right) \right), \\ X_{k+1} & := (1 - \rho)\, X_k + \rho Z_{k+1}. \end{cases}$$

In this case, $\alpha_+ = 0$ and condition (RelaxInertial) becomes

$$\frac{\sqrt{L_{\mathsf{F}}\left(0\right) + 2\left\|A\right\|_2}}{\sqrt{L_{\mathsf{F}}\left(0\right) + 2\left\|A\right\|_2} + \sqrt{L_{\mathsf{F}}\left(0\right)}} < \rho \leq 1 < \frac{\sqrt{L_{\mathsf{F}}\left(0\right) + 2\left\|A\right\|_2}}{\sqrt{L_{\mathsf{F}}\left(0\right) + 2\left\|A\right\|_2} - \sqrt{L_{\mathsf{F}}\left(0\right)}}.$$

Notice that the choice $\rho = 1$ is allowed, which leads to the classical projected gradient algorithm (PG).

Inertial projected gradient algorithm

For $\rho = 1$, RIPG reduces to the inertial projected gradient algorithm (IPG)

$$(\forall k \geq 1) \begin{cases} Y_k & := X_k + \alpha_k \left(X_k - X_{k-1} \right), \\ X_{k+1} & := \mathsf{Pr}_{\mathcal{D}} \left(Y_k - \dfrac{1}{L_{\mathsf{F}}(\alpha_+)} \nabla \mathcal{E} \left(Y_k \right) \right). \end{cases}$$

In this setting, condition (RelaxInertial) is equivalent to

$$0 \leq \alpha_+ < \sqrt{\frac{L_{\mathsf{F}}\left(\alpha_+\right)}{L_{\mathsf{F}}\left(\alpha_+\right) + 2\left\|A\right\|_2}}. \qquad \text{(Inertial)}$$

- Condition (Inertial) is nothing else than

$$\alpha_+^2 \left(\|A\|_2 + \left(3 + 8\alpha_+ + 6\alpha_+^2\right) \operatorname{trace}(A) - \lambda_{\min}(A) \right) \leq \left(3 + 8\alpha_+ + 6\alpha_+^2\right) \operatorname{trace}(A) - \lambda_{\min}(A)$$

and it is fulfilled for every $0 < \alpha_+ \leq 0.967$.

- In our numerical experiments we used 0.0967 as the starting point for a bisection procedure aimed to find larger α_+ which fulfill (Inertial).

Inertial projected gradient algorithm

For $\rho = 1$, RIPG reduces to the inertial projected gradient algorithm (IPG)

$$(\forall k \geq 1) \begin{cases} Y_k & := X_k + \alpha_k \left(X_k - X_{k-1} \right), \\ X_{k+1} & := \mathsf{Pr}_{\mathcal{D}} \left(Y_k - \dfrac{1}{L_{\mathsf{F}}(\alpha_+)} \nabla \mathcal{E} \left(Y_k \right) \right). \end{cases}$$

In this setting, condition (RelaxInertial) is equivalent to

$$0 \leq \alpha_+ < \sqrt{\frac{L_{\mathsf{F}}\left(\alpha_+\right)}{L_{\mathsf{F}}\left(\alpha_+\right) + 2\left\|A\right\|_2}}. \tag{Inertial}$$

- Condition (Inertial) is nothing else than

$$\alpha_+^2 \left(\left\|A\right\|_2 + \left(3 + 8\alpha_+ + 6\alpha_+^2\right) \operatorname{trace}\left(A\right) - \lambda_{\min}\left(A\right) \right) \leq \left(3 + 8\alpha_+ + 6\alpha_+^2\right) \operatorname{trace}\left(A\right) - \lambda_{\min}\left(A\right)$$

and it is fulfilled for every $0 < \alpha_+ \leq 0.967$.

- In our numerical experiments we used 0.0967 as the starting point for a bisection procedure aimed to find larger α_+ which fulfill (Inertial).

Variable inertial parameters for IPG

- $$\alpha_k := \kappa \cdot \frac{t_k - 1}{t_{k+1}}, \qquad \text{where } \begin{cases} t_1 & := 1 \\ t_{k+1} & := \dfrac{1 + \sqrt{1 + 4t_k^2}}{2} \end{cases} \quad \forall k \geq 1. \qquad (\kappa\text{Nes})$$

- (László (MathProg, 2020))

$$\alpha_k := \frac{\kappa k}{k+3} \quad \forall k \geq 1, \qquad \text{where } \kappa \in (0,1). \qquad (\kappa\text{ModNes})$$

- In both cases $\alpha_+ = \sup_{k \geq 1} \alpha_k = \kappa$, thus, according to (Inertial), κ must be chosen such that

$$0 \leq \kappa < \sqrt{\frac{L_{\mathsf{F}}(\kappa)}{L_{\mathsf{F}}(\kappa) + 2\|A\|_2}}.$$

Choosing α_+ even closer to 1

As far as α_+ satisfies (Inertial), we can choose $\rho = 1$. For α_+ close to 1 such that (Inertial) is not satisfied, in other words, if

$$\sqrt{\frac{L_{\mathsf{F}}(\alpha_+)}{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}} \leq \alpha_+,$$

then we have to choose

$$0 < \frac{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2} + \sqrt{L_{\mathsf{F}}(\alpha_+)}} < \rho < \frac{\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}{(1+\alpha_+)\sqrt{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2} - \sqrt{L_{\mathsf{F}}(\alpha_+)}} < 1. \tag{Relax}$$

For $\alpha_k = 1$ for every $k \geq 1$, and thus $\alpha_+ = 1$, RIPG becomes

$$(\forall k \geq 1) \begin{cases} Z_{k+1} & := \mathsf{Pr}_{\mathcal{D}}\left(2X_k - X_{k-1} - \dfrac{1}{L_{\mathsf{F}}(1)}\nabla\mathcal{E}(2X_k - X_{k-1})\right), \\ X_{k+1} & := (1-\rho)X_k + \rho Z_{k+1}. \end{cases}$$

- The strategy of choosing α_+ close to 1 and ρ according to (Relax) yields the best numerical performances of the algorithm.

Numerical experiments

- **Number of runs and starting points:** For $A \in \mathbb{R}^{n \times n}$ with $n < 100$, we run:
  - RIPG 100 times for randomly chosen initial matrices in $\mathcal{D}$;
  - ModMAP and SpFeasDC also 100 times for randomly chosen initial matrices in $\mathbb{O}_r$ (computed via singular value decomposition) and for matrices B computed via Cholesky decomposition.

If $n \geq 100$, then we do this for each of the algorithms 10 times.

- **Parameter choice:** We choose the constant α_+:
  - by running a simple bisection routine with upodate rule $\alpha_+ := (3\alpha_+ + 1)/4$ which starts at 0.967 in order to find greater values for α_+ that satisfy

$$0 \leq \alpha_+ < \sqrt{\frac{L_{\mathsf{F}}(\alpha_+)}{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}.$$

    Then we choose $\alpha_+ := \widehat{\alpha}_+$, which is the last value at which this inequality holds, and $\rho := 1$.
  - by taking $\widehat{\alpha}_1 := (3\widehat{\alpha}_+ + 1)/4$, $\widehat{\alpha}_2 := (\widehat{\alpha}_+ + 1)/2$, and $\widehat{\alpha}_3 := (\widehat{\alpha}_+ + 3)/4$, which, when $\widehat{\alpha}_+$ is obtained as above, all violate the above inequality. The corresponding relaxation parameters will be denoted by $\rho(\widehat{\alpha}_1)$, $\rho(\widehat{\alpha}_2)$ and $\rho(\widehat{\alpha}_3)$, respectively, and chosen to satisfy (Relax).
  - by taking $\alpha_+ := 1$ and the relaxation parameter $\rho(1)$ to satisfy (Relax).

Numerical experiments

- Number of runs and starting points: For $A \in \mathbb{R}^{n \times n}$ with $n < 100$, we run:
  - RIPG 100 times for randomly chosen initial matrices in $\mathcal{D}$;
  - ModMAP and SpFeasDC also 100 times for randomly chosen initial matrices in $\mathbb{O}_r$ (computed via singular value decomposition) and for matrices B computed via Cholesky decomposition.

If $n \geq 100$, then we do this for each of the algorithms 10 times.

- Parameter choice: We choose the constant α_+:
  - by running a simple bisection routine with upodate rule $\alpha_+ := (3\alpha_+ + 1)/4$ which starts at 0.967 in order to find greater values for α_+ that satisfy

$$0 \leq \alpha_+ < \sqrt{\frac{L_{\mathsf{F}}(\alpha_+)}{L_{\mathsf{F}}(\alpha_+) + 2\|A\|_2}}.$$

    Then we choose $\alpha_+ := \widehat{\alpha}_+$, which is the last value at which this inequality holds, and $\rho := 1$.
  - by taking $\widehat{\alpha}_1 := (3\widehat{\alpha}_+ + 1)/4$, $\widehat{\alpha}_2 := (\widehat{\alpha}_+ + 1)/2$, and $\widehat{\alpha}_3 := (\widehat{\alpha}_+ + 3)/4$, which, when $\widehat{\alpha}_+$ is obtained as above, all violate the above inequality. The corresponding relaxation parameters will be denoted by $\rho(\widehat{\alpha}_1)$, $\rho(\widehat{\alpha}_2)$ and $\rho(\widehat{\alpha}_3)$, respectively, and chosen to satisfy (Relax).
  - by taking $\alpha_+ := 1$ and the relaxation parameter $\rho(1)$ to satisfy (Relax).

- **Stopping criteria:** For $A \in \mathbb{R}^{n\times n}$, we run each of the algorithms at most 10000 iterations, if $n < 100$, and at most 50000 iterations, otherwise.
  - **Stopping criterion for ModMAP and SpFeasDC:** $\min\{(BQ_k)_{i,j}\} \geq -\mathtt{Tol}_{\mathtt{fea}}$, with $\mathtt{Tol}_{\mathtt{fea}} := 10^{-16}$, if the matrix A belongs to $\operatorname{int}(\mathcal{CP}_n)$, and $\mathtt{Tol}_{\mathtt{fea}} := 10^{-7}$, otherwise.
  - **Stopping criterion for RIPG:** $\dfrac{\|A - X_k X_k^T\|_{\mathsf{F}}^2}{\|A\|_{\mathsf{F}}^2} < \mathtt{Tol}_{\mathtt{val}}$, with $\mathtt{Tol}_{\mathtt{val}} := 10^{-16}$, if A belongs to $\operatorname{int}(\mathcal{CP}_n)$, and $\mathtt{Tol}_{\mathtt{val}} := 10^{-7}$, otherwise.

- Algorithms:
  - ModMAP: the Modified Method of Alternating Projections (Groetzner, Dür (LAA, 2020));
  - SpFeasDC: the algorithm in (Chen, Pong, Tan, Zeng (JOGO, 2020)) enhanced with a nonmonotone linesearch procedure;
  - PG: the classical projected gradient algorithm ($\rho = 1$ and $\alpha_+ = 0$);
  - IPG-Nes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy Nesterov's rule;
  - IPG-const: $\rho = 1$ with constant inertial parameters and α_+ chosen to satisfy (Inertial);
  - IPG-κNes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy (κNes);
  - IPG-κModNes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy (κModNes);
  - RIPG-const, RIPG-κNes and RIPG-κModNes: elaxed versions of IPG-const, IPG-κNes and IPG-κModNes, respectively, for different values of α_+ that violate (Inertial) and corresponding relaxation parameters ρ satisfying (Relax).

- Stopping criteria: For $A \in \mathbb{R}^{n\times n}$, we run each of the algorithms at most 10000 iterations, if $n < 100$, and at most 50000 iterations, otherwise.
  - Stopping criterion for ModMAP and SpFeasDC: $\min\{(BQ_k)_{i,j}\} \geq -\mathtt{Tol}_{\mathtt{fea}}$, with $\mathtt{Tol}_{\mathtt{fea}} := 10^{-16}$, if the matrix A belongs to $\operatorname{int}(\mathcal{CP}_n)$, and $\mathtt{Tol}_{\mathtt{fea}} := 10^{-7}$, otherwise.
  - Stopping criterion for RIPG: $\dfrac{\|A - X_k X_k^T\|_{\mathsf{F}}^2}{\|A\|_{\mathsf{F}}^2} < \mathtt{Tol}_{\mathtt{val}}$, with $\mathtt{Tol}_{\mathtt{val}} := 10^{-16}$, if A belongs to $\operatorname{int}(\mathcal{CP}_n)$, and $\mathtt{Tol}_{\mathtt{val}} := 10^{-7}$, otherwise.

- Algorithms:
  - ◇ ModMAP: the Modified Method of Alternating Projections (Groetzner, Dür (LAA, 2020));
  - ◇ SpFeasDC: the algorithm in (Chen, Pong, Tan, Zeng (JOGO, 2020)) enhanced with a nonmonotone linesearch procedure;
  - ◇ PG: the classical projected gradient algorithm ($\rho = 1$ and $\alpha_+ = 0$);
  - ◇ IPG-Nes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy Nesterov's rule;
  - ◇ IPG-const: $\rho = 1$ with constant inertial parameters and α_+ chosen to satisfy (Inertial);
  - ◇ IPG-κNes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy (κNes);
  - ◇ IPG-κModNes: $\rho = 1$ and $(\alpha_k)_{k\geq 1}$ chosen to satisfy (κModNes);
  - ◇ RIPG-const, RIPG-κNes and RIPG-κModNes: elaxed versions of IPG-const, IPG-κNes and IPG-κModNes, respectively, for different values of α_+ that violate (Inertial) and corresponding relaxation parameters ρ satisfying (Relax).

Numerical experiment 1

- In each test we generate a random $n \times 2n$ matrix B_0 and set

$$A := |B_0||B_0|^T .$$

- We test the algorithms on 50 randomly generated 40×40 matrices and 10 randomly generated 500×500 matrices.
- We use in each test $r := 1.5n + 1$ and $r := 3n + 1$.

Findings

- SpFeasDC outperforms the other methods with respect to the number of iterations, possibly due to the fact that it uses a linesearch routine to improve the step size, while the others have quite conservative step size rules.
- Some of the instances of RIPG can compete with SpFeasDC in terms of computational time, in particular, the more the dimension grows.

Numerical experiment 1

- In each test we generate a random $n \times 2n$ matrix B_0 and set

$$A := |B_0||B_0|^T .$$

- We test the algorithms on 50 randomly generated 40×40 matrices and 10 randomly generated 500×500 matrices.
- We use in each test $r := 1.5n + 1$ and $r := 3n + 1$.

Findings

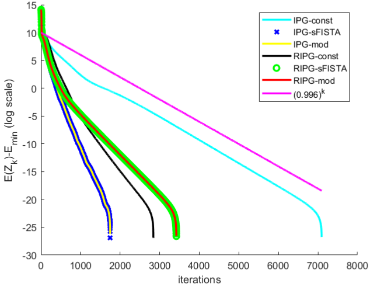
- SpFeasDC outperforms the other methods with respect to the number of iterations, possibly due to the fact that it uses a linesearch routine to improve the step size, while the others have quite conservative step size rules.
- Some of the instances of RIPG can compete with SpFeasDC in terms of computational time, in particular, the more the dimension grows.

$n = 40$ and $r = 61$

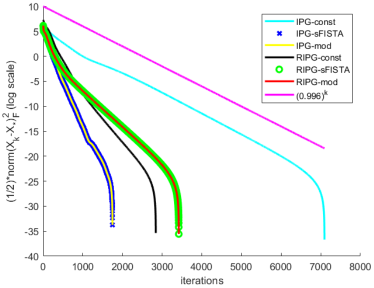
Method	Rate	Time (s)	Time (f)	Iter.
ModMAP	0.80	2.5137×10^{0}	7.0416×10^{0}	3467.08
SpFeasDC	1.00	$\mathbf{4.1259 \times 10^{-2}}$	–//–	**38.51**
PG	0.00	–//–	4.5239×10^{-1}	–//–
IPG-const	1.00	1.3017×10^{-1}	–//–	2554.45
IPG-κNes	1.00	1.2994×10^{-1}	–//–	2561.51
IPG-κModNes	1.00	1.3122×10^{-1}	–//–	2562.88
RIPG-const	1.00	2.8331×10^{-1}	–//–	5490.14
RIPG-κNes	1.00	$\mathbf{8.8411 \times 10^{-2}}$	–//–	**1752.14**
RIPG-κModNes	1.00	$\mathbf{8.9617 \times 10^{-2}}$	–//–	**1751.66**

The nonnegative factorization of random completely positive matrices for $n = 40$ and $r = 61$.

$n = 40$ and $r = 61$



$$\mathcal{E}(Z_k) - \mathcal{E}_{\min}$$



$$\frac{1}{2}\|X_k - X_{\mathrm{sol}}\|_{\mathsf{F}}^2$$

$n = 500$ and $r = 751, 1501$

Method	Rate	Time (s)	Time (f)	Iter.
SpFeasDC	1.00	1.6557×10^2	–//–	**929.38**
RIPG-κNes	1.00	$\mathbf{1.4526 \times 10^2}$	–//–	7919.40
RIPG-κModNes	1.00	1.4861×10^2	–//–	7921.64

The nonnegative factorization of random completely positive matrices for $n = 500$ and $r = 751$.

Method	Rate	Time (s)	Time (f)	Iter.
SpFeasDC	1.00	1.3813×10^3	–//–	**914.15**
RIPG-κNes	1.00	$\mathbf{2.2975 \times 10^2}$	–//–	7776.30
RIPG-κModNes	1.00	2.3037×10^2	–//–	7779.60

The nonnegative factorization of random completely positive matrices for $n = 500$ and $r = 1501$.

Numerical experiment 2

We examine the efficiency of the factorization algorithms when the number of columns r varies:

- for matrices of the form

$$A_n := \begin{pmatrix} 0 & \mathbf{j}_{n-1}^T \\ \mathbf{j}_{n-1} & \mathbb{I}_{n-1} \end{pmatrix}^T \begin{pmatrix} 0 & \mathbf{j}_{n-1}^T \\ \mathbf{j}_{n-1} & \mathbb{I}_{n-1} \end{pmatrix} \in \mathbb{R}^{n \times n},$$

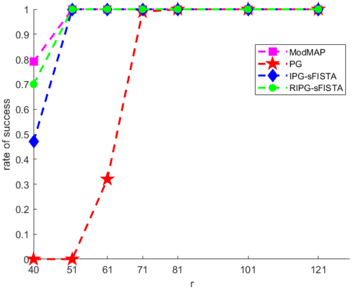
  where $\mathbb{I}_n$ and $\mathbf{j}_n$ denote the $n \times n$ identity matrix and the all-ones-vector in $\mathbb{R}^n$, respectively. We set $n := 40$, choose $r \in \{40, 51, 61, 71, 81, 101, 121\}$ and consider 100 random initial points.

- for a completely positive matrix $A := |B_0||B_0|^T$ constructed from a randomly generated 100×200 matrix B_0, for $r \in \{151, 176, 201, 251, 301\}$, and random initial points.

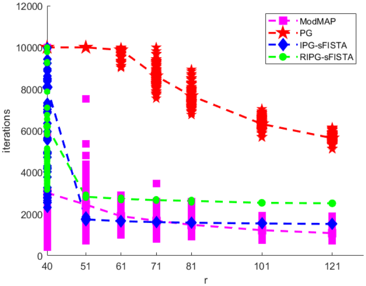
Findings

- The rate of success for different variants RIPG increases with higher values for r.
- RIPG requires less iterations than ModMAP to provide a nonnegative factorization.

A_{40} for $r \in \{40, 51, 61, 71, 81, 101, 121\}$



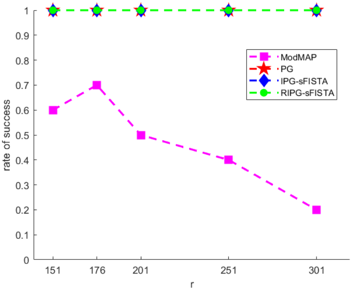
The rate of success



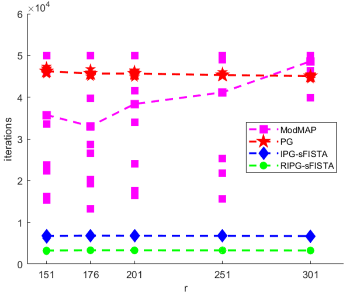
The number of iterations required

The rate of success and number of iterations required for the factorization of A_{40} for different values of r and random initial points. The dash-lines show the average value.

$A \in \mathcal{CP}_{100}$ for $r \in \{151, 176, 201, 251, 301\}$



The rate of success



The number of iterations required

The rate of success and number of iterations required for the factorization of a randomly generated matrix $A \in \mathcal{CP}_{100}$ for different values of r and random initial points. The dash-lines show the average value.

Numerical experiment 3

We consider the perturbed matrix A_ω defined by

$$A_\omega := \omega A + (1-\omega)\,P, \quad \text{for } \omega \in [0,1]\,,$$

where

$$A := \begin{pmatrix} 8 & 5 & 1 & 1 & 5 \\ 5 & 8 & 5 & 1 & 1 \\ 1 & 5 & 8 & 5 & 1 \\ 1 & 1 & 5 & 8 & 5 \\ 5 & 1 & 1 & 5 & 8 \end{pmatrix} \quad \text{and} \quad P := \begin{pmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{pmatrix}.$$

- Both A and A_ω, $\omega \in [0,1]$, belong to $\mathcal{CP}_5$.
- It is more difficult to factorize A then A_ω, for $\omega < 1$. The reason is that $A \in \mathcal{CP}_5 \setminus \mathrm{int}\,(\mathcal{CP}_5)$.
- All known factorization algorithms can successfully factorize A_ω for various values of $\omega < 1$, but fail to do so for $\omega = 1$ (when $A_\omega = A$).

Findings

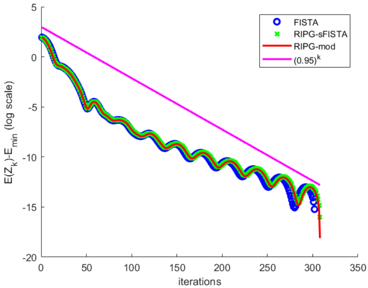
- The inertial methods `IPG-const`, IPG-κNes and IPG-κModNes also face some difficulties when factorizing A.
- The methods RIPG-κNes and RIPG-κModNes, which combine relaxation and inertial parameters, always return nonnegative factorizations.

$A_{0.99}$

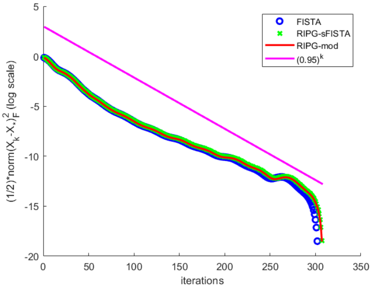
Method	Rate	Time (s)	Time (f)	Iter.
ModMAP	0.00	–//–	4.7649×10^{-1}	–//–
SpFeasDC	0.02	7.0223×10^{-1}	7.5259×10^{-1}	9220.50
PG	0.27	1.8571×10^{-2}	2.7675×10^{-2}	7069.00
IPG-Nes	1.00	$\mathbf{2.1624 \times 10^{-3}}$	–//–	**728.32**
IPG-const	1.00	7.2203×10^{-3}	–//–	2385.20
IPG-κNes	1.00	7.9190×10^{-3}	–//–	2474.65
IPG-κModNes	1.00	7.7214×10^{-3}	–//–	2473.84
RIPG-const	0.94	1.3217×10^{-2}	3.2318×10^{-2}	4446.59
RIPG-κNes	1.00	$\mathbf{2.5225 \times 10^{-3}}$	–//–	**742.12**
RIPG-κModNes	1.00	$\mathbf{2.4953 \times 10^{-3}}$	–//–	**744.37**

The nonnegative factortization of $A_{0.99}$ for $r = 12$.

$A_{0.99}$



$$\mathcal{E}(Z_k) - \mathcal{E}_{\min}$$



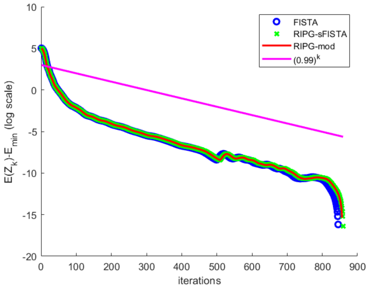
$$\frac{1}{2}\|X_k - X_{\text{sol}}\|_{\mathsf{F}}^2$$

$A_1 = A$

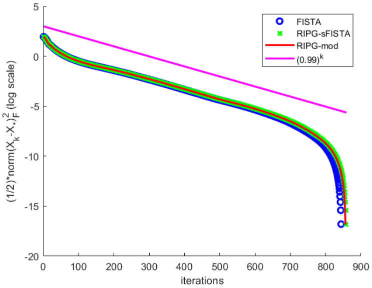
Method	Rate	Time (s)	Time (f)	Iter.
ModMAP	0.00	−//−	5.0659×10^{-1}	−//−
SpFeasDC	0.00	−//−	9.1030×10^{-1}	−//−
`PG`	0.01	1.7454×10^{-2}	2.7524×10^{-2}	7531.00
`IPG-Nes`	1.00	$\mathbf{3.1237 \times 10^{-3}}$	−//−	**1067.09**
`IPG-const`	0.99	1.1232×10^{-2}	2.9201×10^{-2}	3785.31
`IPG-`κ`Nes`	0.95	1.2694×10^{-2}	3.3234×10^{-2}	4052.98
`IPG-`κ`ModNes`	0.95	1.2337×10^{-2}	3.0064×10^{-2}	4041.04
`RIPG-const`	0.76	1.7549×10^{-2}	2.9381×10^{-2}	5908.16
`RIPG-`κ`Nes`	1.00	$\mathbf{3.6109 \times 10^{-3}}$	−//−	**1083.75**
`RIPG-`κ`ModNes`	1.00	$\mathbf{3.6073 \times 10^{-3}}$	−//−	**1084.20**

The nonnegative factortization of $A_1 = A$ for $r = 11$.

$A_1 = A$



$$\mathcal{E}(Z_k) - \mathcal{E}_{\min}$$



$$\frac{1}{2}\|X_k - X_{\mathrm{sol}}\|_{\mathsf{F}}^2$$

Numerical experiment 4

Consider

$$A_{2n} := \begin{pmatrix} n\mathbb{I}_n & \mathbf{J}_n \\ \mathbf{J}_n & n\mathbb{I}_n \end{pmatrix} \in \mathcal{CP}_{2n} \setminus \operatorname{int}(\mathcal{CP}_{2n}),$$

where $\mathbb{I}_n$ and $\mathbf{J}_n$ denote the identity matrix and the all-ones-matrix in $\mathbb{R}^{n\times n}$, respectively.

Findings

- The methods RIPG-κNes and RIPG-κModNes, which combine relaxation and inertial parameters, provide nonnegative factorizations in reasonable time.
- IPG-Nes outperforms all the other methods.

Method	Rate	Time (s)	Time (f)	Iter.
ModMAP	0.00	−//−	3.4746×10^2	−//−
SpFeasDC	0.00	−//−	5.8390×10^2	−//−
`IPG-Nes`	1.00	$\mathbf{9.9557 \times 10^{-1}}$	−//−	**6959.95**
IPG-κNes	0.00	−//−	1.5584×10^0	−//−
IPG-κModNes	0.00	−//−	1.5747×10^0	−//−
RIPG-κNes	1.00	1.4564×10^0	−//−	7037.52
RIPG-κModNes	1.00	1.4641×10^0	−//−	7036.06

The nonnegative factorization of A_{30} for $r = 30$.

Method	Rate	Time (s)	Time (f)	Iter.
`IPG-Nes`	1.00	$\mathbf{1.9818 \times 10^2}$	−//−	**22246.50**
RIPG-κNes	1.00	2.3330×10^2	−//−	22467.40
RIPG-κModNes	1.00	2.3290×10^2	−//−	22463.90

The nonnegative factorization of A_{100} for $r = 100$.

Further perspectives

- Numerical evidence suggests that the convergence rates are linear, which at its turn suggests that the Łojasiewicz exponent of the energy function is at most $1/2$.
- Use in RIPG variable step sizes.
- Extend the convergence analysis beyond the current setting, in order to cover the parameter choice of the `IPG-Nes` method.
- Replace the closed ball with radius $\sqrt{\text{trace}\,(A)}$ with the sphere of the same radius.

References

H. Attouch, J. Bolte. *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*. Mathematical Programming 116(1), 5–16 (2009)

H.H. Bauschke, M.N. Bui, X. Wang. *Projecting onto the intersection of a cone and a sphere*. SIAM Journal on Optimization 28(3), 2158–2188 (2018)

I.M. Bomze, P.J.C. Dickinson, G. Still. *The structure of completely positive matrices according to their cp-rank and cp-plus-rank*. Linear Algebra and its Applications 482, 191–206 (2015)

R. I. Boţ, D.-K. Nguyen. *Factorization of completely positive matrices using iterative projected gradient steps*. Numerical Linear Algebra with Applications, DOI: 10.1002/nla.2391 (2021)

C. Chen, T.K. Pong, L. Tan, L. Zeng *A difference-of-convex approach for split feasibility with applications to matrix factorizations and outlier detection*. Journal of Global Optimization 78, 107–136 (2020)

P.J.C. Dickinson. *An improved characterisation of the interior of the completely positive cone*. Electronic Journal of Linear Algebra 20, 723–729 (2010)

P. Groetzner, M. Dür. *A factorization method for completely positive matrices*. Linear Algebra and its Applications 591, 1–24 (2020)

K. Kurdyka. *On gradients of functions definable in o-minimal structures*. Annales de l'Institut Fourier 48, 769–783 (1998)

S. Łojasiewicz. *Une propriété topologique des sous-ensembles analytiques réels*. in: Colloques internationaux du C.N.R.S.: Les équations aux dérivées partielles, Paris (1962), Éditions du Centre National de la Recherche Scientifique, Paris, 87–89 (1963)

Thank you for your attention!

https://www.mat.univie.ac.at/~rabot/