Low-Rank Plus Sparse Decompositions of Large-Scale Matrices via Semidefinite Optimization

Rui Gong

Department of Combinatorics and Optimization
Faculty of Mathematics
University of Waterloo

MMath Thesis Presentation, May 3, 2023

Table of Contents

Table of Contents

Introduction

- Given a large n and an n-by-n symmetric matrix X, can we write it as

$$X = L + S,$$

where L is a low-rank symmetric positive semidefinite (PSD) matrix and S is a sparse symmetric matrix? If so, how to find such L and S?

Introduction

- Given a large n and an n-by-n symmetric matrix X, can we write it as

$$X = L + S,$$

where L is a low-rank symmetric positive semidefinite (PSD) matrix and S is a sparse symmetric matrix? If so, how to find such L and S?

- Why do we need to consider this, or why do we want to write X in this form?

Motivation of Low-Rank Plus Sparse Matrices Decomposition

- Let $n = 1,000,000$, X is an n-by-n symmetric matrix and $X = L^* + S^* \in \mathbb{R}^{n\times n}$.

Motivation of Low-Rank Plus Sparse Matrices Decomposition

- Let $n = 1,000,000$, X is an n-by-n symmetric matrix and $X = L^* + S^* \in \mathbb{R}^{n \times n}$.
- Since L^* is a low-rank PSD matrix, we can write it as $L = VV^\top$ where $\text{rank}(L) = r \ll n$, $V \in \mathbb{R}^{n \times r}$. Since S^* is sparse, we assume it has $m \ll n$ nonzero entries and we use another $2m$ units to store the indices of them.

Motivation of Low-Rank Plus Sparse Matrices Decomposition

- Let $n = 1,000,000$, X is an n-by-n symmetric matrix and $X = L^* + S^* \in \mathbb{R}^{n \times n}$.
- Since L^* is a low-rank PSD matrix, we can write it as $L = VV^\top$ where $\text{rank}(L) = r \ll n$, $V \in \mathbb{R}^{n \times r}$. Since S^* is sparse, we assume it has $m \ll n$ nonzero entries and we use another $2m$ units to store the indices of them.
- Then we only need to store $3m + rn$ numbers to have all information of X instead of $n^2/2$.

Motivation of Low-Rank Plus Sparse Matrices Decomposition

- Let $n = 1,000,000$, X is an n-by-n symmetric matrix and $X = L^* + S^* \in \mathbb{R}^{n \times n}$.
- Since L^* is a low-rank PSD matrix, we can write it as $L = VV^\top$ where $\text{rank}(L) = r \ll n$, $V \in \mathbb{R}^{n \times r}$. Since S^* is sparse, we assume it has $m \ll n$ nonzero entries and we use another $2m$ units to store the indices of them.
- Then we only need to store $3m + rn$ numbers to have all information of X instead of $n^2/2$.
- Also the decomposition increases the efficiency for matrix operations. For example, if we have a vector $y \in \mathbb{R}^n$ and compute Xy, the direct computation costs $2n^2 - n$ operations. But if we consider $VV^\top y + S^* y$, it takes approximately $4rn + 2m - r - n - 1 \ll 2n^2 - n$ operations.

Factor Analysis Model

Factor Analysis Model in Statistics describes the difference of observed $X \in \mathbb{R}^{n_1 \times n_2}$ and the mean values $M \in \mathbb{R}^{n_1 \times n_2}$ as $X - M = UF + \epsilon$, where $U \in \mathbb{R}^{n_1 \times r}$, $F \in \mathbb{R}^{r \times n_2}$, $\mathbb{E}[F] = 0, cov(F) = I$, ϵ is an error and F and ϵ are independent. We assume U to have a low rank and the error term ϵ has independent columns. Then

$$cov(X - M) = Ucov(F)U^\top + cov(\epsilon) = UU^\top + \text{Diag}(d)$$

for some $d \in \mathbb{R}^{n_1}$.

Low-Rank Plus Diagonal Decomposition

Given $A \in \mathbb{S}^n$, we want to solve

$$\begin{aligned} \min_{x\in\mathbb{R}^n, L\in\mathbb{S}^n} \ & \operatorname{rank}(L) \\ s.t. \ \ & A = L + \operatorname{Diag}(x) \\ & L \succeq 0 \end{aligned}$$

Low-Rank Plus Diagonal Decomposition

By changing the sign of the variable, given $A \in \mathbb{S}^n$, we want to solve

$$\min_{x \in \mathbb{R}^n} \operatorname{rank}(A + \operatorname{Diag}(x))$$
$$s.t. \quad A + \operatorname{Diag}(x) \succeq 0.$$

By [Fazel, 2002], we may relax the problem by replacing the rank$(\cdot)$ by its convex envelope $\|\cdot\|_*$ and get:

$$\min_{x \in \mathbb{R}^n} \|A + \operatorname{Diag}(x)\|_*$$
$$s.t. \quad A + \operatorname{Diag}(x) \succeq 0$$

By $A + \operatorname{Diag}(x) \succeq 0$, we have

$$\min \|A + \operatorname{Diag}(x)\|_* = \min \operatorname{tr}(A + \operatorname{Diag}(x)) = \min \operatorname{tr}(\operatorname{Diag}(x)) = \min \mathbb{1}^\top x$$

Then we get the *Minimum Trace Factor Analysis (MTFA)* problem:

$$\min \ \mathbb{1}^\top x$$
$$s.t. \quad A + \operatorname{Diag}(x) \succeq 0$$

MaxCut as MTFA [Goemans and Williamson, 1995]

Given a simple graph $G := ([n], E)$, and a weight matrix $W \in \mathbb{S}^n$, then the MaxCut problem can be represented as:

$$\max \frac{1}{4} \sum_{i \in V} \sum_{j \in V} W_{ij}(1 - u_i u_j) \text{ s.t. } u \in \{1, -1\}^n.$$

MaxCut as MTFA [Goemans and Williamson, 1995]

Given a simple graph $G := ([n], E)$, and a weight matrix $W \in \mathbb{S}^n$, then the MaxCut problem can be represented as:

$$\max \frac{1}{4} \sum_{i \in V} \sum_{j \in V} W_{ij}(1 - u_i u_j) \text{ s.t. } u \in \{1, -1\}^n.$$

Then MaxCut problem can be relaxed as

$$\begin{aligned} \max \ & -\frac{1}{4}\operatorname{tr}(WX) \ \left(+\frac{1}{4}\mathbb{1}^\top W \mathbb{1}\right) \\ s.t. \ & \operatorname{diag}(X) = \mathbb{1}, \\ & X \succeq 0 \end{aligned}$$

and its dual is defined as:

$$\begin{aligned} \min \ & \mathbb{1}^\top y \ \left(+\frac{1}{4}\mathbb{1}^\top W \mathbb{1}\right) \\ s.t. \ & \frac{1}{4}W + \operatorname{Diag}(y) \succeq 0, \end{aligned}$$

which is in the form of MTFA.

Equivalent MTFA

By changing the sign of the variable, the MTFA problem:

$$\min\ \mathbb{1}^\top x \ s.t.\ A + \mathrm{Diag}(x) \succeq 0$$

is equivalent to

$$\begin{aligned} \max\ & \langle \mathbb{1}, y \rangle \\ s.t.\ & L + \mathrm{Diag}(y) = A \\ & L \succeq 0 \\ & y \in \mathbb{R}^n \end{aligned} \tag{MTFA}$$

whose dual is defined as

$$\begin{aligned} \min\ & \langle A, X \rangle \\ s.t.\ & \mathrm{diag}(X) = \mathbb{1} \\ & X \succeq 0. \end{aligned} \tag{MTFAD}$$

MTFA Theorem

Theorem 1.1 ([Riccia and Shapiro, 1982])

When an MTFA instance has a feasible solution, it has a unique optimal solution.

Table of Contents

Three Problems [Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012]

1. Suppose $X^* \in \mathbb{S}^n$ can be written in the form of $X^* = L^* + \mathrm{Diag}(y^*)$, where L^* is symmetric positive semidefinite. What properties or conditions of (L^*, y^*) will ensure that (L^*, y^*) is the unique optimal solution of (MTFA) with the input $A = X^*$?

Three Problems [Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012]

2. For a closed convex set C, a *face* F of C is a closed convex subset of C such that

$$u, v \in C, \alpha \in (0,1) \text{ and } \alpha u + (1-\alpha)v \in F \implies u, v \in F.$$

A face is *proper* if $F \neq C$ and $F \neq \emptyset$.
Recall that every face of $\mathbb{S}^n_+$, $\mathcal{F}_\mathcal{U}$, is by a subspace $\mathcal{U}$ of $\mathbb{R}^n$, where

$$\mathcal{F}_\mathcal{U} = \{X \succeq 0 : \text{Null}(X) \supseteq \mathcal{U}\}.$$

We have that every face of $\mathcal{E}_n$, the elliptope (the set of correlation matrices:$X \in \mathbb{S}^n_+, \text{diag}(X) = \mathbb{1}$), is in the form

$$\mathcal{E}_n \cap \mathcal{F}_\mathcal{U} = \{X \succeq 0 : \text{Null}(X) \supseteq \mathcal{U}, \text{diag}(X) = \mathbb{1}\},$$

where $\mathcal{U}$ is a subspace of $\mathbb{R}^n$. However, for some subspaces $\mathcal{U}$ of $\mathbb{R}^n$, the set $\mathcal{E}_n \cap \mathcal{F}_\mathcal{U}$ is empty. For example, consider $\mathcal{U} = \text{span}\{e_1\}$, there is no $X \in \mathcal{E}_n$ such that $\text{Null}(X) \supseteq \mathcal{U}$, so $\mathcal{E}_n \cap \mathcal{F}_\mathcal{U} = \emptyset$. Thus, another problem is, which subspaces define a nonempty face of $\mathcal{E}_n$?

Three Problems [Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012]

3. A *centered (degenerate) ellipsoid* in $\mathbb{R}^n$ is a set in the form: given $A \in \mathbb{S}^n_+$:

$$\Lambda := \{x \in \mathbb{R}^n : x^\top A x \leq 1\},$$

We say a centered ellipsoid passing through $x \in \mathbb{R}^n$ if $x^\top A x = 1$.
Consider the lemma:

Lemma 2.1 ([Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012])

Suppose V is a $k \times n$ matrix with row space $\mathcal{V}$. If there exists a centered ellipsoid in $\mathbb{R}^k$ passing through each column (which is a point in $\mathbb{R}^k$) of V, then there exists a centered ellipsoid such that for every matrix W with row space $\mathcal{V}$, this ellipsoid passes through all columns of W.

For which subspaces $\mathcal{V}$ of $\mathbb{R}^n$, do there exist a positive integer k and a $k \times n$ matrix V with row space $\mathcal{V}$ and a centered ellipsoid passing through all its columns?

Recoverability, Realizability and Ellipsoid Fitting

Definition 2.2 ([Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012])

1. A subspace $\mathcal{U}$ of $\mathbb{R}^n$ is *diagonally recoverable* by MTFA if for every $y^* \in \mathbb{R}^n$ and every $L^* \in \mathbb{S}^n_+$ with column space $\mathcal{U}$, (L^*, y^*) is the unique optimal solution of (MTFA) with input $A = \text{Diag}(y^*) + L^*$.
2. A subspace $\mathcal{U}$ of $\mathbb{R}^n$ is *diagonally realizable* if there exists a correlation matrix $Q \in \mathcal{E}_n$ such that $\text{Null}(Q) \supseteq \mathcal{U}$.
3. A subspace $\mathcal{V}$ of $\mathbb{R}^n$ has the *ellipsoid fitting property* if there exists $V \in \mathbb{R}^{k\times n}$ with row space $\mathcal{V}$ such that there is a centered ellipsoid in $\mathbb{R}^k$ passing through each column of V.

Equivalence

Proposition 2.3 ([Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012])

Let $\mathcal{U}$ be a subspace of $\mathbb{R}^n$, then the followings are equivalent:

1. $\mathcal{U}$ is diagonally recoverable.
2. $\mathcal{U}$ is diagonally realizable.
3. $\mathcal{U}^\perp$ has the ellipsoid fitting property.

Equivalence

Proposition 2.3 ([Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012])

Let $\mathcal{U}$ be a subspace of $\mathbb{R}^n$, then the followings are equivalent:

1. $\mathcal{U}$ is diagonally recoverable.
2. $\mathcal{U}$ is diagonally realizable.
3. $\mathcal{U}^\perp$ has the ellipsoid fitting property.

Proposition 2.4 (Gong 2023)

A subspace $\mathcal{U}$ of $\mathbb{R}^n$ is diagonally recoverable if and only if there exists $y^* \in \mathbb{R}^n$ and $L^* \in \mathbb{S}^n$ with column space $\mathcal{U}$ such that (L^*, y^*) is the unique optimal solution of (MTFA) with input $A = \mathrm{Diag}(y^*) + L^*$.

Table of Contents

Low-Rank Plus Tridiagonal Decomposition

Given $u \in \mathbb{R}^n, v \in \mathbb{R}^{n-1}$, we define $\operatorname{TriDiag}(u, v) : \mathbb{R}^n \oplus \mathbb{R}^{n-1} \to \mathbb{S}^n$, such that

$$\operatorname{TriDiag}(u, v) = \begin{bmatrix} u_1 & v_1 & & & \\ v_1 & \ddots & \ddots & \mathbf{0} & \\ & \ddots & \ddots & \ddots & \\ & \mathbf{0} & \ddots & \ddots & v_{n-1} \\ & & & v_{n-1} & u_n \end{bmatrix}$$

Its adjoint is defined as:

$$\operatorname{tridiag}(X) := \begin{bmatrix} \operatorname{diag}(X) \\ \operatorname{bidiag}(X) \end{bmatrix},$$

where $\operatorname{bidiag}(X) := 2 \begin{bmatrix} X_{12} \\ X_{23} \\ \vdots \\ X_{(n-1)n} \end{bmatrix}$.

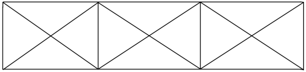
Definition 3.1 (Low-Rank Plus Tridiagonal Decomposition)

Given $A \in \mathbb{S}^n$, we want to solve

$$\begin{aligned} \min_{x \in \mathbb{R}^n, L \in \mathbb{S}^n} \quad & \operatorname{rank}(L) \\ s.t. \quad & A = L + \operatorname{TriDiag}(u, v) \\ & L \succeq 0 \end{aligned}$$

Motivations

- (Rough) Truss Design:



- Time-Dependent Models
- General chain-like Models
- Banded symmetric matrices with bandwidth one

Relaxations

By relaxing rank to $\|\cdot\|_*$ and changing the sign of variables:

$$\begin{aligned} & \min \ \mathbb{1}^\top u \\ & s.t. \ \ \mathrm{TriDiag}(u, v) \succeq -A \\ & \qquad u \in \mathbb{R}^n \\ & \qquad v \in \mathbb{R}^{n-1}. \end{aligned}$$

Relaxations

Adding regularizations on v: for $\lambda \geq 0$,

$$\begin{aligned} \min\ & \mathbb{1}^\top u + \lambda \|v\|_1 \\ s.t.\ & \mathrm{TriDiag}(u, v) \succeq -A \\ & u \in \mathbb{R}^n \\ & v \in \mathbb{R}^{n-1}. \end{aligned}$$

Tridiagonal Perturbation

$$
\begin{aligned}
\min\ & \mathbb{1}^\top u + \lambda \mathbb{1}^\top t \\
s.t.\ & \mathrm{TriDiag}(u, v) \succeq -A \\
& t - v \geq 0 \\
& t + v \geq 0 \\
& u \in \mathbb{R}^n \\
& t, v \in \mathbb{R}^{n-1},
\end{aligned}
\qquad \text{(TriRegP)}
$$

and we define the linear map on the left-hand side of the constraints as $\mathcal{A}^*(u, v, t) : \mathbb{R}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1} \to \mathbb{S}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1}$.

Tridiagonal Perturbation

$$
\begin{aligned}
\min\ & \mathbb{1}^\top u + \lambda \mathbb{1}^\top t \\
s.t.\ & \text{TriDiag}(u, v) \succeq -A \\
& t - v \geq 0 \\
& t + v \geq 0 \\
& u \in \mathbb{R}^n \\
& t, v \in \mathbb{R}^{n-1},
\end{aligned}
\qquad \text{(TriRegP)}
$$

and we define the linear map on the left-hand side of the constraints as $\mathcal{A}^*(u, v, t) : \mathbb{R}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1} \to \mathbb{S}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1}$.

$$
\begin{aligned}
\max\ & -\langle A, X\rangle \\
& \text{diag}(X) = \mathbb{1} \\
& \text{bidiag}(X) + w - \xi = 0 \\
& w + \xi = \lambda \mathbb{1} \\
& X \succeq 0, w \geq 0, \xi \geq 0
\end{aligned}
\qquad \text{(TriRegD)}
$$

where the linear map on the left-hand side of the constraints is $\mathcal{A}(X, \xi, \omega) : \mathbb{S}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1} \to \mathbb{R}^n \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1}$.

Equivalent Dual

For the dual, notice that the constraints

$$\begin{aligned}
&\operatorname{diag}(X) = \mathbb{1} \\
&\operatorname{bidiag}(X) + w - \xi = 0 \\
&w + \xi = \lambda \mathbb{1} \\
&X \succeq 0, w \geq 0, \xi \geq 0
\end{aligned}$$

are equivalent to

$$\begin{aligned}
&\operatorname{diag}(X) = \mathbb{1} \\
&-\lambda \mathbb{1} \leq \operatorname{bidiag}(X) \leq \lambda \mathbb{1} \\
&X \succeq 0.
\end{aligned}$$

Uniqueness of Optimal Solutions

Is the optimal solution to the relaxation unique?

Uniqueness of Optimal Solutions

Is the optimal solution to the relaxation unique?

If it is not unique, then our problem becomes a new problem of comparing the optimal solutions for the relaxation.

Uniqueness of Optimal Solutions

Theorem 3.2 (Gong 2023)

When $\lambda \in \mathbb{R}_+ \setminus \{2\}$, the relaxation (TriRegP) has a unique optimal solution.

Infinitely Many Optimal Solutions

Given $\lambda = 2$, $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{S}^2$. Then for every feasible solution X of (TriRegD), we have $X \succeq 0$ and $\text{tridiag}(X) = (u, 2v) = (\mathbb{1}, 2v), \xi + w = 2$, then $\langle -A, X \rangle = -2v$. Since $X \succeq 0$, we have $1 = u_1 u_2 \geq v^2$, hence $\langle -A, X \rangle = -2v \leq 2$. Then,

$$X = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, w = 2, \xi = 0$$

is an optimal solution for (TriRegD).

Infinitely Many Optimal Solutions

Given $\lambda = 2$, $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{S}^2$. Then for every feasible solution X of (TriRegD), we have $X \succeq 0$ and $\operatorname{tridiag}(X) = (u, 2v) = (\mathbb{1}, 2v), \xi + w = 2$, then $\langle -A, X \rangle = -2v$. Since $X \succeq 0$, we have $1 = u_1 u_2 \geq v^2$, hence $\langle -A, X \rangle = -2v \leq 2$. Then,

$$X = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, w = 2, \xi = 0$$

is an optimal solution for (TriRegD).
And we consider an arbitrary optimal solution (u, v, t) of (TriRegP). By Strong Duality Theorem, we know

$$X(\operatorname{TriDiag}(u, v) + A) = 0, \operatorname{Diag}(t - v) \operatorname{Diag}(\xi) = 0 = \operatorname{Diag}(t + v) \operatorname{Diag}(w)$$

which implies $t + v = 0$ and $u_1 = u_2 = 1 + v = 1 - t$, so $\mathbb{1}^\top u + \lambda t = 1 - t + 1 - t + 0 + 2t = 2$. That is, given any $1 \geq t \geq 0$, $((1 - t, 1 - t), -t, t)$ is an optimal solution of (TriRegP).

3 Problems for the Tridiagonal Perturbation

1. Given $\lambda \in \mathbb{R}^n_+ \setminus \{2\}$, suppose X^* can be written in the form of $X^* = L^* - \text{TriDiag}(u^*, v^*)$, where l^* is positive semidefinite. What properties or conditions of $(L^*, u^*, v^*, |v|^*)$ will ensure that $(L^*, u^*, v^*, |v|^*)$ is the unique optimal solution of (TriRegP) with the input $A = X^*$?

2. Given $\lambda \in \mathbb{R}^n_+ \setminus \{2\}$, we first consider a closed convex set

$$K := \{X \in \mathbb{S}^n : \mathrm{diag}(X) = \mathbb{1}, -\lambda\mathbb{1} \leq \mathrm{bidiag}(X) \leq \lambda\mathbb{1}\}.$$

Claim 3.2.1

Every proper face F of K can be written as:

$$F = \{X \in \mathbb{S}^n : \mathrm{diag}(X) = \mathbb{1}, \\ k \in [n-1] \text{ entries of } \mathrm{bidiag}(X) \text{ is fixed as } \lambda \text{ or } -\lambda\}.$$

Consider a set

$$\mathcal{E}'_n := \{X \in \mathcal{E}_n : -\lambda\mathbb{1} \leq \mathrm{bidiag}(X) \leq \lambda\mathbb{1}\} = \mathbb{S}^n_+ \cap K.$$

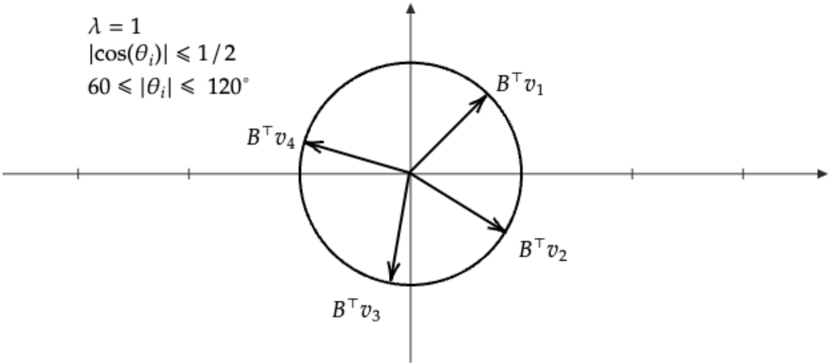
Every face of $\mathcal{E}'_n$ can be written as $F \cap \mathcal{F}_{\mathcal{U}}$, where F is a face of K and $\mathcal{F}_{\mathcal{U}}$ is a face of $\mathbb{S}^n_+$ uniquely defined by a subspace $\mathcal{U}$ of $\mathbb{R}^n$. E.g.

$$\mathcal{E}'_n \cap \mathcal{F}_{\mathcal{U}} = \{X \succeq 0 : \mathrm{Null}(X) \supseteq \mathcal{U}, \mathrm{diag}(X) = \mathbb{1}, -\lambda\mathbb{1} \leq \mathrm{bidiag}(X) \leq \lambda\mathbb{1}\},$$

is a face of $\mathcal{E}'_n$. Which subspaces $\mathcal{U}$ define a nonempty face of $\mathcal{E}'_n$?

3 Problems for the Tridiagonal Perturbation

3. Given $\lambda \in \mathbb{R}^n_+ \setminus \{2\}$, then we consider the problem: for which subspaces $\mathcal{V}$ of $\mathbb{R}^n$, do there exist a positive integer k and a $k \times n$ matrix V with row space $\mathcal{V}$ and a centered ellipsoid passing through all its columns such that when the points are projected onto the unit ball corresponding to the ellipsoid, the absolute value of the cosine value of the angle between the projected ith and $(i+1)$th columns is upper bounded by $\lambda/2$?



λ-Tridiagonal Recoverability, Realizability and Ellipsoid Fitting

Definition 3.3 (Gong 2023)

Consider a subspace $\mathcal{U}$ of $\mathbb{R}^n$ and $\lambda \in \mathbb{R}_+ \setminus \{2\}$.

- We say $\mathcal{U}$ is *λ-tridiagonally recoverable* if there exists $u \in \mathbb{R}^n, v \in \mathbb{R}^{n-1}$, $L \in \mathbb{S}^n_+$ such that $\operatorname{col}(L) = \mathcal{U}$, $(u, v, |v|)$ is the unique solution of (TriRegP) given $A = L - \operatorname{TriDiag}(u, v)$.
- We say $\mathcal{U}$ is *λ-tridiagonally realizable* if there exists $Q \in \mathbb{S}^n_+$ such that $\mathcal{U} \subseteq \operatorname{Null}(Q)$, $\operatorname{diag}(Q) = \mathbb{1}, -\lambda\mathbb{1} \leq \operatorname{bidiag}(Q) \leq \lambda\mathbb{1}$.
- We say $\mathcal{U}$ has the *λ-tridiagonal ellipsoid fitting property* if there is a $k \times n$ matrix V with row space $\mathcal{U}$ such that
    1. there is a centered ellipsoid in $\mathbb{R}^k$ passing through each column of V.
    2. Let $M \in \mathbb{S}^k_+$ represent the ellipsoid, and write $M = BB^\top$, and let $\mathcal{B} := \{B^\top v : v^\top Mv = 1\}$, which is the projected unit ball corresponding to the ellipsoid. And the angle θ_i between projections of the ith and $(i+1)$th column of V onto the ball satisfies $|\cos(\theta_i)| \leq \lambda/2$.

3 Equivalent Tridiagonal Definitions

Proposition 3.4 (Gong 2023)

Consider subspaces $\mathcal{U}$ of $\mathbb{R}^n$ with $\lambda \in \mathbb{R}_+ \setminus \{2\}$, the followings are equivalent:

1. $\mathcal{U}$ is λ-tridiagonally recoverable.
2. $\mathcal{U}$ is λ-tridiagonally realizable.
3. $\mathcal{U}^\perp$ has the λ-tridiagonal ellipsoid fitting property.

Converge to MTFA

Recall that the constraints of (TriRegD),

$$\begin{aligned} &\mathrm{diag}(X) = \mathbb{1} \\ &\mathrm{bidiag}(X) + w - \xi = 0 \\ &w + \xi = \lambda \mathbb{1} \\ &X \succeq 0, w \geq 0, \xi \geq 0 \end{aligned}$$

are equivalent to

$$\begin{aligned} &\mathrm{diag}(X) = \mathbb{1} \\ &-\lambda \mathbb{1} \leq \mathrm{bidiag}(X) \leq \lambda \mathbb{1} \\ &X \succeq 0. \end{aligned}$$

Hence, when $\lambda \to \infty$, $-\lambda \mathbb{1} \leq \mathrm{bidiag}(X) \leq \lambda \mathbb{1}$ becomes a redundant constraint. So (TriRegD) is equivalent to the dual of MTFA (MTFAD).

Converge to MTFA

When $\lambda \to \infty$, since the objective function is

$$\mathbb{1}^\top u + \lambda \mathbb{1}^\top t,$$

every optimal solution of (TriRegP) has $t = 0$ and $v = 0$. That is, (TriRegP) is equivalent to the minimum trace factor analysis problem (MTFA).

Converge to MTFA

When $\lambda \to \infty$, since the objective function is

$$\mathbb{1}^\top u + \lambda \mathbb{1}^\top t,$$

every optimal solution of (TriRegP) has $t = 0$ and $v = 0$. That is, (TriRegP) is equivalent to the minimum trace factor analysis problem (MTFA).

Corollary 3.5 (Gong 2023)

In fact, when $\lambda > 2$, both (TriRegP) and (TriRegD) are equivalent to (MTFA) and (MTFAD) respectively.

Converge to MTFA

When $\lambda \to \infty$, since the objective function is

$$\mathbb{1}^\top u + \lambda \mathbb{1}^\top t,$$

every optimal solution of (TriRegP) has $t = 0$ and $v = 0$. That is, (TriRegP) is equivalent to the minimum trace factor analysis problem (MTFA).

Corollary 3.5 (Gong 2023)

In fact, when $\lambda > 2$, both (TriRegP) and (TriRegD) are equivalent to (MTFA) and (MTFAD) respectively.

λ-tridiagonal recoverability, realizability and ellipsoid fitting property are also equivalent to the diagonal ones when $\lambda \to \infty$ (or $\lambda > 2$).

Table of Contents

Coherence

Definition 4.1 ([Candès and Recht, 2008])

Let $\mathcal{U}$ be a subspace of $\mathbb{R}^n$ of dimension r and $\mathbf{P}_\mathcal{U} \in \mathbb{S}^n$ be the orthogonal projection matrix onto $\mathcal{U}$. Then the *coherence* of $\mathcal{U}$ (with respect to the standard basis e_i) is defined to be

$$\mu(\mathcal{U}) := \max_{1 \le i \le n} \|\mathbf{P}_\mathcal{U} e_i\|^2.$$

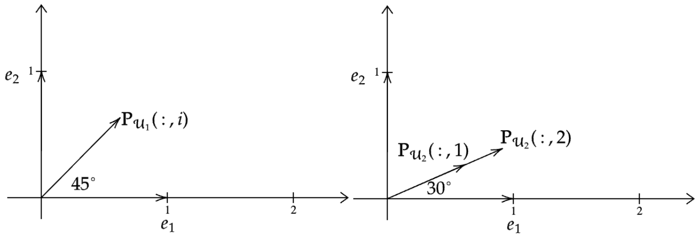
For a subspace $\mathcal{U}$ of $\mathbb{R}^n$ of dimension r, we have

$$\frac{r}{n} \le \mu(\mathcal{U}) \le 1.$$

Interpreting the Coherence

We can view the coherence of $\mathcal{U}$ as an indicator of how close it is to containing any e_i.

Consider $\mathcal{U}_1 := \text{span}\{[1/\sqrt{2}, 1/\sqrt{2}]^\top\}$, and $\mathcal{U}_2 := \text{span}\{[\sqrt{3}/2, 1/2]^\top\}$ where $\mathbf{P}_{\mathcal{U}_1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ and $\mathbf{P}_{\mathcal{U}_1} = \begin{bmatrix} 3/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/4 \end{bmatrix}$, $\mu(\mathcal{U}_1) = 1/2 < \mu(\mathcal{U}_2) = 3/4$.



Suff. Condition for Diag. Realizability

Theorem 4.2 ([Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S., 2012])

If a subspace $\mathcal{U}$ of $\mathbb{R}^n$ has $\mu(\mathcal{U}) < 1/2$ then $\mathcal{U}$ is diagonally realizable. On the other hand, for every $\alpha > 1/2$, there exists a diagonally unrealizable subspace $\mathcal{U}$ with $\mu(\mathcal{U}) = \alpha$.

Suff. Condition for TriDiag. Realizability

Proposition 4.3 (Gong 2023)

Consider subspace $\mathcal{U}$ of $\mathbb{R}^n$ with coherence $\mu(\mathcal{U}) < 1/2$ and let $p > 0$ be a constant such that $\|\mathbf{P}_{\mathcal{U}^\perp}(:, i) \circ \mathbf{P}_{\mathcal{U}^\perp}(:, i+1)\| \leq p$ for every $i \in [n-1]$. If $p\frac{\sqrt{n}}{(1-\mu)^2} < \frac{\lambda}{2}$, there exists infinitely many Q satisfying $Q \succeq 0$, $\mathrm{diag}(Q) = \mathbb{1}$, $\mathcal{U} \subseteq \mathrm{Null}(Q)$ and

$$-\lambda \mathbb{1} \leq \mathrm{bidiag}(Q) \leq \lambda \mathbb{1}.$$

Suff. Condition for TriDiag. Realizability

Proposition 4.3 (Gong 2023)

Consider subspace $\mathcal{U}$ of $\mathbb{R}^n$ with coherence $\mu(\mathcal{U}) < 1/2$ and let $p > 0$ be a constant such that $\|\mathbf{P}_{\mathcal{U}^\perp}(:, i) \circ \mathbf{P}_{\mathcal{U}^\perp}(:, i+1)\| \leq p$ for every $i \in [n-1]$. If $p\frac{\sqrt{n}}{(1-\mu)^2} < \frac{\lambda}{2}$, there exists infinitely many Q satisfying $Q \succeq 0$, $\mathrm{diag}(Q) = \mathbb{1}$, $\mathcal{U} \subseteq \mathrm{Null}(Q)$ and

$$-\lambda \mathbb{1} \leq \mathrm{bidiag}(Q) \leq \lambda \mathbb{1}.$$

Proposition 4.4 (Gong 2023)

Consider a subspace $\mathcal{U}$ with dimension r, coherence $\mu < 1/2$ and let $p > 0$ be a constant such that $\|\mathbf{P}_{\mathcal{U}^\perp}(:, i) \circ \mathbf{P}_{\mathcal{U}^\perp}(:, i+1)\| \leq p$ for every $i \in [n-1]$. If $\kappa(p, r, \mu) := p\sqrt{n + \left(\frac{1}{(1-\mu)^4} - 1\right)\frac{r}{\mu}} < \frac{\lambda}{2}$, then there exists infinitely many Q satisfying $Q \succeq 0$, $\mathrm{diag}(Q) = \mathbb{1}$, $\mathcal{U} \subseteq \mathrm{Null}(Q)$ and

$$-\lambda \mathbb{1} \leq \mathrm{bidiag}(Q) \leq \lambda \mathbb{1}.$$

Computational Example

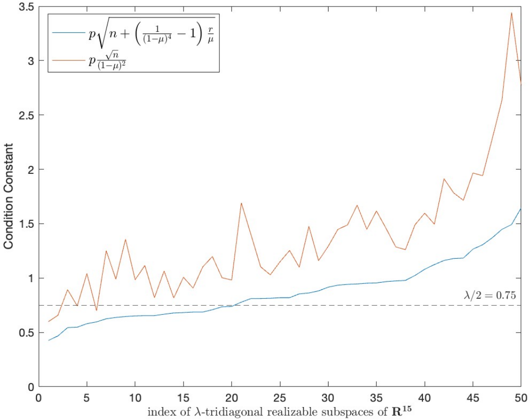


Figure: Sufficient Conditions of 1.5-tridiagonally realizable subspaces

Table of Contents

Algorithms

We can solve the low-rank plus sparse matrices decomposition in two different ways:

1. Solve one relaxation of the problem and treat the optimal solutions to the relaxation as the solutions to the original problem.

Algorithms

We can solve the low-rank plus sparse matrices decomposition in two different ways:

1. Solve one relaxation of the problem and treat the optimal solutions to the relaxation as the solutions to the original problem.
2. Solve the problem exactly.

Exact Algorithm

Given a low-rank plus tridiagonal decomposition problem whose optimal value is $\bar{r} \in O(1)$, there is an algorithm solving this instance in polynomial time, which is an extension from an algorithm in [Tunçel, Vavasis, and Xu, 2022].

Exact Algorithm

Given a low-rank plus tridiagonal decomposition problem whose optimal value is $\bar{r} \in O(1)$, there is an algorithm solving this instance in polynomial time, which is an extension from an algorithm in [Tunçel, Vavasis, and Xu, 2022].

Fancy Algorithm Solves such a Complicated Nonconvex Problem Exactly in Polynomial Time.

Exact Algorithm

Given a low-rank plus tridiagonal decomposition problem whose optimal value is $\bar{r} \in O(1)$, there is an algorithm solving this instance in polynomial time, which is an extension from an algorithm in [Tunçel, Vavasis, and Xu, 2022].

Fancy Algorithm Solves such a Complicated Nonconvex Problem Exactly in Polynomial Time.

Please see read the thesis

Table of Contents

Conclusion

- Low-Rank Plus Diagonal Decomposition, MTFA and the diagonal recoverability, realizability and ellipsoid fitting.

Conclusion

- Low-Rank Plus Diagonal Decomposition, MTFA and the diagonal recoverability, realizability and ellipsoid fitting.
- Low-Rank Plus Tridiagonal Decomposition, a relaxation of it and the λ-tridiagonal recoverability, realizability and ellipsoid fitting.

Conclusion

- Low-Rank Plus Diagonal Decomposition, MTFA and the diagonal recoverability, realizability and ellipsoid fitting.
- Low-Rank Plus Tridiagonal Decomposition, a relaxation of it and the λ-tridiagonal recoverability, realizability and ellipsoid fitting.
- Sufficient Condition for Realizability.

Conclusion

- Low-Rank Plus Diagonal Decomposition, MTFA and the diagonal recoverability, realizability and ellipsoid fitting.
- Low-Rank Plus Tridiagonal Decomposition, a relaxation of it and the λ-tridiagonal recoverability, realizability and ellipsoid fitting.
- Sufficient Condition for Realizability.
- Algorithms for solving low-rank plus tridiagonal decomposition problem exactly in polynomial time.

Future Research

- For problem (TriRegP), we used a linear objective function with a penalty parameter on the absolute values of bidiagonal entries. Can we replace this objective function with more general functions? In particular, can we replace it with other norms or other general convex functions and extend the properties like realizability, and uniqueness of optimal solutions to those cases?

Future Research

- For problem (TriRegP), we used a linear objective function with a penalty parameter on the absolute values of bidiagonal entries. Can we replace this objective function with more general functions? In particular, can we replace it with other norms or other general convex functions and extend the properties like realizability, and uniqueness of optimal solutions to those cases?
- In this thesis, we introduced the low-rank plus tridiagonal decomposition problem, and analyzed its optimality conditions and different properties. Can we apply similar analyses and expect results from more general sparsity patterns? For example, if we change tridiagonal matrices to matrices with a chordal sparsity pattern, we would expect more general results because tridiagonal matrices represent a chordal sparsity pattern, but what are we gaining by having more freedom on the sparsity pattern?

Future Research

- For problem (TriRegP), we used a linear objective function with a penalty parameter on the absolute values of bidiagonal entries. Can we replace this objective function with more general functions? In particular, can we replace it with other norms or other general convex functions and extend the properties like realizability, and uniqueness of optimal solutions to those cases?
- In this thesis, we introduced the low-rank plus tridiagonal decomposition problem, and analyzed its optimality conditions and different properties. Can we apply similar analyses and expect results from more general sparsity patterns? For example, if we change tridiagonal matrices to matrices with a chordal sparsity pattern, we would expect more general results because tridiagonal matrices represent a chordal sparsity pattern, but what are we gaining by having more freedom on the sparsity pattern?
- Since the tridiagonal matrices are banded matrices with bandwidth one, can we generalize our problem to banded matrices with larger bandwidth? If so, will similar properties like recoverability, realizability and ellipsoid fitting properties begeneralized? What about the uniqueness of optimal solutions?

Thank you!

References I

Emmanuel J. Candès and Benjamin Recht. Exact matrix completion via convex optimization. *CoRR*, abs/0805.4471, 2008. URL http://arxiv.org/abs/0805.4471.

Maryam Fazel. *Matrix Rank Minimization with Applications*. PhD thesis, Stanford University, 2002.

Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, nov 1995. ISSN 0004-5411. doi: 10.1145/227683.227684. URL https://doi.org/10.1145/227683.227684.

Giacomo D Riccia and Alexander Shapiro. Minimum rank and minimum trace of covariance matrices. *Psychometrika*, 47(4):443–448, December 1982. ISSN 1860-0980. doi: 10.1007/BF02293708. URL https://doi.org/10.1007/BF02293708.

Saunderson, J. and Chandrasekaran, V. and Parrilo, P. A. and Willsky, A. S. Diagonal and low-rank matrix decompositions, correlation matrices, and ellipsoid fitting. *SIAM J. Matrix Anal. Appl.*, 33(4):1395–1416, 2012.

References II

Levent Tunçel, Stephen A. Vavasis, and Jingye Xu. Computational complexity of decomposing a symmetric matrix as a sum of positive semidefinite and diagonal matrices, 2022. URL https://arxiv.org/abs/2209.05678. arXiv:2209.05678.