

Matrix Rank Minimization problems

Andersen Ang

ECS, Uni. Southampton, UK

andersen.ang@soton.ac.uk

Homepage angms.science

Version: July 10, 2024

First draft: March 20, 2021

Content

Rank Minimization Problem
RMP with affine constraint
Matrix Completion

Rank Minimization Problem (RMP)

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad (\text{RMP})$$

- ▶ $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the optimization variable.
- ▶ There are many formulations of the rank function.
One example is the ℓ_0 -seminorm of the singular values of \mathbf{X} : let $\boldsymbol{\sigma}$ be the vector holding the singular values of \mathbf{X} , then

$$\text{rank}(\mathbf{X}) = \|\boldsymbol{\sigma}\|_0.$$

- ▶ ℓ_0 -seminorm is combinatorial \implies RMP is NP-hard.
 - ▶ it has a trivial solution: \mathbf{X} is the zero matrix, and $\text{rank}(\mathbf{X}) = 0$.
 - ▶ NP-hard refers to the case $\mathbf{X} \neq \mathbf{0}$
- ▶ RMP is more interesting if there is constraint.

RMP with constraint

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad \text{s.t.} \quad c(\mathbf{X}) = 0. \quad (\text{RMP-c})$$

- ▶ $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the optimization variable and the function $c : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ represents a constraint.
 - ▶ RMP-c is still not interesting: it is too general, c can be anything.
We now consider c with some specific structure.
-

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad \text{s.t.} \quad \mathcal{A}(\mathbf{X}) = \mathbf{b}. \quad (\text{ARMP})$$

- ▶ $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^d$ is an affine operator.
 $\mathbf{b} \in \mathbb{R}^d$ is a vector.
- ▶ ARMP is interesting because
 - ▶ The constraint has enough structure so that we can study the properties of the problem.
 - ▶ The constraint is general enough that the problem covers many applications.
- ▶ Affine constraint is a convex set, and the problem is NP-hard because of the rank.

ARMP and related problems

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad \text{s.t.} \quad \mathcal{A}(\mathbf{X}) = \mathbf{b}. \quad (\text{ARMP})$$

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) + \lambda \text{dist}(\mathcal{A}(\mathbf{X}), \mathbf{b}). \quad (\text{P-ARMP})$$

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) + \frac{\lambda}{2} \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_F^2. \quad (\text{Fro-ARMP})$$

- ▶ ARMP is the original constrained formulation
- ▶ P-ARMP casts ARMP in a penalty form by adding the constraint into the objective with the penalty weight $\lambda \geq 0$, and thereby forming a unconstrained problem. dist is a function that measures the distance between $\mathcal{A}(\mathbf{X})$ and \mathbf{b} .
- ▶ Fro-ARMP is a special case of P-ARMP where Frobenius norm is used for dist .
- ▶ ARMP is noiseless: given \mathcal{A} and \mathbf{b} we assume there is a low-rank matrix \mathbf{X} such that $\mathcal{A}(\mathbf{X}) = \mathbf{b}$.
- ▶ If we assume additive noise model, P-ARMP can be seen as robust ARMP.
E.g., in Fro-ARMP we are now trying to minimize the size of additive white Gaussian noise

$$\mathcal{A}(\mathbf{X}) = \mathbf{b} + \varepsilon.$$

The affine constraint $\mathcal{A}(\mathbf{X}) = \mathbf{b}$

- ▶ $\mathbf{b} \in \mathbb{R}^d$ is a given vector and $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^d$ is a linear map.
- ▶ An interpretation of $(\mathcal{A}, \mathbf{b})$ is that \mathcal{A} represents certain measurement operation and \mathbf{b} is the measurement.
- ▶ The notation $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ is an abstract representation of the following system of matrix inner product

$$\mathcal{A}(\mathbf{X}) = \begin{bmatrix} \langle \mathbf{A}_1, \mathbf{X} \rangle \\ \langle \mathbf{A}_2, \mathbf{X} \rangle \\ \vdots \\ \langle \mathbf{A}_d, \mathbf{X} \rangle \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{bmatrix} = \mathbf{b}$$

where \mathbf{A}_i , $i = 1, 2, \dots, d$ are all m -by- n matrices and $\langle \cdot, \cdot \rangle$ is the matrix inner product.

- ▶ Matrices \mathbf{A}_i in the expression $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ can be think as an analogue to \mathbf{a}_i in the matrix \mathbf{A} in the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Why RMP is NP-hard

- ▶ Let σ be the singular value of the matrix \mathbf{X} . Then $\text{rank}(\mathbf{X}) = \|\sigma\|_0$ where $\|\cdot\|_0$ is the ℓ_0 -norm .
- ▶ ℓ_0 -norm minimization problems are NP-hard.
- ▶ As RMP is at least as hard as ℓ_0 -norm minimization problem, so RMP is (at least) NP-hard.
- ▶ See p.7-8 [here](#) for a simple illustration of the combinatorial nature of ℓ_0 -norm.

Matrix Completion (MC)

- ▶ A famous example of RMP is the MC.
- ▶ MC is the following problem: suppose $\mathbf{X}_0 \in \mathbb{R}^{m \times n}$ is a low-rank matrix and suppose we are given $\mathbf{b} = \mathcal{P}_\Omega(\mathbf{X}_0)$ where $\Omega \subseteq [m] \times [n]$ and \mathcal{P}_Ω is the projection that, for any matrix \mathbf{M} with entries M_{ij} ,
 - ▶ if $(i, j) \in \Omega$, $\mathcal{P}_\Omega(M_{ij}) = M_{ij}$, and
 - ▶ if $(i, j) \notin \Omega$, $\mathcal{P}_\Omega(M_{ij}) = 0$.

The goal of MC is to fill in the missing entries of \mathbf{X}_0 .

- ▶ Here $\mathcal{A} = \mathcal{P}_\Omega$ is the observation operator.
- ▶ Here $d = |\Omega|$.
- ▶ As MC is a special case of RMP, so all the other formulations of RMP (P-RMP, Fro-RMP) also apply to MC, that is, we also have robust formulation of MC for dealing with noisy observation.

Illustration of MC is a special case of RMP

► Example, let $\mathbf{X} \in \mathbb{R}^{3,2}$ as $\mathbf{X} = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$, and $\mathbf{b} \in \mathbb{R}^2 = \begin{bmatrix} a \\ d \end{bmatrix}$.

► So $\Omega = \{(1, 1), (2, 2)\}$ and $|\Omega| = 2$.

► In this case $\mathcal{A}(\mathbf{X}) = \begin{bmatrix} \langle \mathbf{A}_1, \mathbf{X} \rangle \\ \langle \mathbf{A}_2, \mathbf{X} \rangle \end{bmatrix}$ with

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which is basically the single entry matrix \mathbf{J}_{ij} where the location of the 1 is at the location of (i, j) in Ω .

The two main research problems on RMP

- ▶ When does solving RMP problem recover the ground truth solution?
- ▶ How to solve RMP problem?
 - ▶ This involves the study of designing efficient optimization algorithm to solve RMP. These involves to characterize some conditions on \mathcal{A} , \mathbf{X} and \mathbf{b} such that RMP is solvable. A typical characterization is the restricted isometry property

$$1 - \delta(k) \leq \frac{\|\mathcal{A}(\mathbf{X})\|_{\text{op}}}{\|\mathbf{X}\|_F} \leq 1 + \delta(k)$$

where $\delta(k)$ is a parameter depending on k , which is the rank of the matrix \mathbf{X} .

- ▶ There are many methods:
 - ▶ Method based on nuclear norm convex relaxation
 - ▶ Method based on matrix factorization
 - ▶ Method based on singular value decomposition (e.g. SVT, SVP)
- ... and other methods.

Summary: Overview of rank minimization problem with affine constraint.

End of document