An Introduction to a Class of Matrix Cone Programming

Chao Ding

Department of Mathematics, National University of Singapore
This talk is based on a joint work with Defeng Sun and Kim-Chuan Toh at NUS

SIAM Conference on Optimization
Darmstadt, Germany
May 16, 2011

Outline

A motivating example

The matrix cone programming (MCP)

The matrix optimization problem (MOP)

The Moreau-Yosida regularization and spectral operators

A Motivating Example

The Fastest Mixing Markov Chain (FMMC)[1]:

- $(\mathcal{V}, \mathcal{E})$ is a given connected graph with vertex set $\mathcal{V} = \{1, \ldots, n\}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$.
- Let $P \in \mathbb{R}^{n \times n}$ be the transition probability matrix. This matrix must satisfy

$$P \geq 0, \quad Pe = e \quad \text{and} \quad P = P^T .$$

- Let $\pi(t) \in \mathbb{R}^n$ be the probability distribution of the state at time t. By the recursion property, we know that

$$\pi(t)^T = \pi(0)^T P^t .$$

If the chain is irreducible and aperiodic, the the distribution $\pi(t)$ converges to the uniform distribution $(1/n)e$ – the unique equilibrium distribution as t increases.

[1] S. Boyd, P. Diaconis, and L. Xiao, *Fastest mixing Markov chain on a graph*, SIAM Review, 46 (2004), pp. 667–689.

- The rate of convergence to this distribution is determined by the second largest eigenvalue modulus (SLEM):

$$\mathsf{SLEM}(P) = \max_{i=2,\ldots,n} |\lambda_i(P)| \,.$$

Find the edge transition probabilities that give the fastest mixing Markov chain, i.e.,

$$\begin{aligned} \min \quad & \mathsf{SLEM}(P) \\ \text{s.t.} \quad & P \geq 0,\ Pe = e,\ P = P^T, \\ & P_{ij} = 0,\ (i,j) \notin \mathcal{E} \,. \end{aligned} \tag{1}$$

The SDP approach

It is well-known that FMMC (1) can be written as an SDP,

$$\begin{array}{ll} \min & s \\ \text{s.t.} & -sI \preceq P - (1/n)ee^T \preceq sI, \\ & P \geq 0,\ Pe = e,\ P = P^T, \\ & P_{ij} = 0,\ (i,j) \notin \mathcal{E}\,. \end{array} \qquad (2)$$

The inequalities in (2) can be expressed as a single linear matrix inequality (LMI),

$$\operatorname{diag}\big(P - (1/n)ee^T + sI, sI - P + (1/n)ee^T, \operatorname{vec}(P)\big) \succeq 0\,.$$

As mentioned by Boyd et al. (2004), standard primal-dual interior-point algorithms for solving SDPs work well for problems with up to a thousand or so edges. But problems with 10,000 or more edges are probably beyond the capabilities of standard interior-point SDP solvers (without using the particular structures).

Large-scale FMMC

- As suggested in Boyd et al. (2009)[2], by using the graph symmetric property, one can reduce the number of variables and the size of matrices.
- Such SDP reformulations need be avoided for large-scale problems, which is also suggested in Boyd et al. (2004).

Let $p \in \mathbb{R}^m$ denote the vector of transition probabilities on the non-self-loop edges. Since $Pe = e$, we know that

$$P = \mathcal{P}(p) = I + \sum_{l=1}^{m} p_l E^{(l)} ,$$

where $E_{ij}^{(l)} = E_{ji}^{(l)} = +1$, $E_{ii}^{(l)} = E_{jj}^{(l)} = -1$ and all other entries of $E^{(l)}$ are zero.

[2] S. Boyd, P. Diaconis, P. A. Parrilo, and L. Xiao, *Fastest mixing Markov chain on graphs with Symmetries*, SIAM Journal on Optimization, 20 (2009), pp. 792–819.

Consider the following equivalent formulation of FMMC:

$$\begin{aligned} \min \quad & \mathsf{SLEM}(\mathcal{P}(p)) \\ \text{s.t.} \quad & p \geq 0,\ Bp \leq e\,, \end{aligned} \tag{3}$$

where $B \in \mathbb{R}^{n \times m}$ is the vertex-edge incidence matrix. Since P is symmetric, we know that

$$\mathsf{SLEM}(P) = \max_{i=2,\dots,n} |\lambda_i(P)| = \sigma_2(P)\,.$$

By noting that $\sigma_1(P) \equiv 1$, we know that (3) is equivalent with

$$\begin{aligned} \min \quad & \sigma_1(\mathcal{P}(p)) + \sigma_2(\mathcal{P}(p)) = \|\mathcal{P}(p)\|_{(2)} \\ \text{s.t.} \quad & p \geq 0,\ Bp \leq e\,, \end{aligned} \tag{4}$$

where $\| \cdot \|_{(k)}$ is Ky Fan k-norm of matrices, i.e., the sum of the k largest singular values of a matrix.

It is obvious that the problem (4) is equivalent with the following conic programming

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & p \geq 0,\ Bp \leq e,\ \mathcal{P}(p) = Z\,, \\ & (t, Z) \in \mathcal{K}\,, \end{aligned} \tag{5}$$

where the close convex cone $\mathcal{K} \in \mathbb{R} \times \mathbb{R}^{n\times n}$ is defined by

$$\mathcal{K} = \text{epi}\,\|\cdot\|_{(2)} = \{(t, X) \in \mathbb{R} \times \mathbb{R}^{n\times n} \mid \|X\|_{(2)} \leq t\}\,.$$

We call the problem like (5) the matrix cone programming (MCP).

The Matrix Cone Programming

The primal and dual forms:

$$\begin{array}{lll} \min\limits_{x \in \mathbb{R}^p} & c^T x \\ \text{s.t.} & \mathcal{A}x \in b + \mathcal{Q} \times \mathcal{K} \end{array} \qquad \begin{array}{ll} \max\limits_{Y \in \mathcal{Y}} & \langle b, Y \rangle \\ \text{s.t.} & \mathcal{A}^* Y = c\,, \\ & Y \in \mathcal{Q}^* \times \mathcal{K}^* \end{array}$$

- $\mathcal{Y} := \mathcal{H} \times \mathcal{X}$ with $\mathcal{H} = \mathbb{R}^q$ or $\mathcal{S}^q$ and $\mathcal{X} := \mathbb{R} \times \mathbb{R}^{m \times n}$.
- $\mathcal{Q} \in \mathcal{H}$ is the cross product of the origin $\{0\}$ and a symmetric cone ($\mathbb{R}^q_+$, the SOC cone in $\mathbb{R}^q$ and the SDP cone $\mathcal{S}^q_+$). $\mathcal{Q}^*$ is the dual cone of $\mathcal{Q}$.
- $\mathcal{K} := \text{epi}\, f \in \mathcal{X}$ is the epigraph cone of some matrix norm $f \equiv \|\cdot\|$. $\mathcal{K}^*$ is the dual cone of $\mathcal{K}$.
- $\mathcal{A} : \mathbb{R}^p \to \mathcal{Y}$ is a linear operator. $\mathcal{A}^* : \mathcal{Y} \to \mathbb{R}^p$ is the adjoint of $\mathcal{A}$.

Other Applications

- Matrix Norm Approximation

$$\min\left\{\|B_0 + \sum_{k=1}^{p} y_k B_k\|_2 \,|\, y \in \mathbb{R}^p\right\}.$$

- Matrix Completion

$$\min\left\{\|P_\Omega(X) - P_\Omega(M)\|_F^2 + \lambda\|X\|_* \,|\, X \in \mathbb{R}^{m\times n}\right\}.$$

- Robust Matrix Completion/Robust PCA

$$\min\left\{\|X\|_* + \lambda\|Y\|_1 \,|\, P_\Omega(X) + P_\Omega(Y) = P_\Omega(M)\right\}$$

and

$$\min\left\{\|P_\Omega(X) + P_\Omega(Y) - P_\Omega(M)\|_F^2 + \rho\Big(\|X\|_* + \lambda\|Y\|_1\Big) \,|\, X, Y \in \mathbb{R}^{m\times n}\right\}.$$

- Structured Low Rank Matrix Approximation

$$\min\{\|X - M\|_F + \rho \sum_{k=r+1}^{\min\{m,n\}} \sigma_k(X) \,|\, \mathcal{A}(X) \in b + \mathcal{Q}\}.$$

Matrix Cones

$$\mathcal{K} = \operatorname{epi} f = \left\{(t, X) \in \mathbb{R} \times \mathbb{R}^{m\times n} \mid f(X) \le t\right\} .$$

(1) $f = \|\cdot\|_F$, the Frobenius norm;

(2) $f = \|\cdot\|_\infty$, the l_∞ norm, i.e, $\|X\|_\infty = \max\{|x_{ij}|\}$, $X \in \mathbb{R}^{m\times n}$;

(3) $f = \|\cdot\|_1$, the l_1 norm, $\|X\|_1 = \sum_i \sum_j |x_{ij}|$, $X \in \mathbb{R}^{m\times n}$;

(4) $f = \|\cdot\|_2$, the spectral or operator norm, i.e., the largest singular value of $X \in \mathbb{R}^{m\times n}$;

(5) $f = \|\cdot\|_*$, the nuclear norm, i.e., the sum of the singular values of $X \in \mathbb{R}^{m\times n}$;

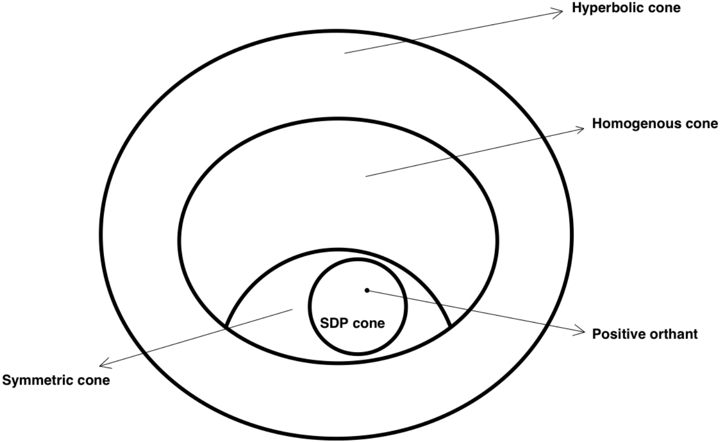
(6) $f = \|\cdot\|_{(k)}$, the Ky Fan k-norm;

(7) $f = \|\cdot\|_{(k)^*}$, the dual norm of Ky Fan k-norm $\|\cdot\|_{(k)}$, i.e., $\|X\|_{(k)^*} = \max\{\|X\|_2, \frac{1}{k}\|X\|_*\}$.

$$\mathcal{K}_2 \overset{\text{dual}}{\longleftrightarrow} \mathcal{K}_3 \quad \mathcal{K}_4 \overset{\text{dual}}{\longleftrightarrow} \mathcal{K}_5 \quad \mathcal{K}_6 \overset{\text{dual}}{\longleftrightarrow} \mathcal{K}_7$$

Except for the case when $f = \|\cdot\|_F$, the cone $\mathcal{K}$ is not self-dual. It is easy to prove that $\mathcal{K}_4$ is a proper hyperbolic cone.

The Cones Relationship



The Matrix Optimization Problem (MOP)

The primal and dual forms:

$$\begin{array}{ll} \min & c^T x + f(Z) \\ \text{s.t.} & \mathcal{A}x - B = Z \end{array} \quad \text{and} \quad \begin{array}{ll} \max & \langle B, Y\rangle - f^*(Y) \\ \text{s.t.} & \mathcal{A}^* Y = c\,, \end{array}$$

where $f : \mathcal{Y} \to (-\infty, +\infty]$ be a closed proper convex function, and f^* is the Fenchel conjugate of f, i.e.,

$$f^*(Y) := \sup_{Z \in \mathcal{Y}} \{\langle Z, Y\rangle - f(Z)\}, \quad Y \in \mathcal{Y}\,.$$

The MOP is more general (including both MCPs and "original" problems as (4) in FMMC)

- For MCPs, $f = \delta_{\mathcal{Q}\times\mathcal{K}}(\cdot)$ and $f^* = \delta_{\mathcal{Q}^\circ\times\mathcal{K}^\circ}(\cdot)$.
- $f = \|\cdot\|_\#$ be any norm function defined on $\mathcal{Y}$ and $\|\cdot\|_*$ be the dual norm of $\|\cdot\|_\#$. Since f is a positively homogeneous close convex function, $f^* = \delta_{\partial f(0)}(\cdot)$, where

$$\partial f(0) = B_*^1 := \{Y \in \mathcal{Y} \mid \|Y\|_* \le 1\}.$$

For (4) in FMMC:

$$\begin{array}{rl} \min & f(Z) = \|Z\|_{(2)} \\ \text{s.t.} & p \geq 0,\ Bp - e \leq 0\,, \\ & \mathcal{P}(p) = Z \end{array} \quad \text{and} \quad \begin{array}{rl} \max\limits_{Y \in \mathcal{Y}} & \langle e, v \rangle - \delta_{\partial f(0)}(Y) \\ \text{s.t.} & -u + B^T v + \mathcal{P}^*(Y) = 0\,, \\ & u \geq 0,\ v \geq 0\,, \end{array}$$

where $\partial f(0) = \{Y \in \mathbb{R}^{m \times n} \mid \|Y\|_{(2)^*} \leq 1\}$.

The PPA and the Moreau-Yosida regularization

The dual problem of the MOP:

$$\begin{array}{ll} \min & \langle B, Y\rangle + f^*(Y) \\ \text{s.t.} & \mathcal{A}^* Y = c\,. \end{array} \tag{6}$$

Given a sequence of parameters λ_k such that

$$0 < \lambda_k \uparrow \lambda_\infty \leq +\infty$$

and an initial point $Y_0 \in \mathcal{Y}$, the (primal) Proximal Point Algorithm (PPA) for solving (6) generates a sequence Y^k by the following scheme:

$$Y^{k+1} \approx \arg\min_{Y \in \mathcal{Y}} \left\{ \phi(Y) + \frac{1}{2\lambda_k} \|Y - Y^k\|^2 \right\} ,$$

where $\phi(Y) := \sup_{y \in \mathbb{R}^p} \{\langle B, Y\rangle + f^*(Y) + \langle y, c - \mathcal{A}^* Y\rangle\}$.

Therefore, we have

$$Y^{k+1} \approx \arg\max_{y\in\mathbb{R}^p} \left\{ \langle y, c\rangle + \frac{1}{2\lambda_k}(\|Y^k\|^2 - \|Z\|^2) + \min_{Y\in\mathcal{Y}}\{f^*(Y) + \frac{1}{2}\|Y - Z\|^*\} \right\},$$

where $Z := \lambda_k(\mathcal{A}y - B) + Y^k$.
The key: the Moreau-Yosida regularization of f^*

$$\psi_{f^*}(Z) := \min_{Y\in\mathcal{Y}} \left\{ f^*(Y) + \frac{1}{2}\|Y - Z\|^* \right\}.$$

The Moreau-Yosida regularization

In order to make MOPs tractable, the function f and f^* should be "simple" and "computable".

- The Moreau-Yosida regularization of f

$$\psi_f(Z) := \min_{Y \in \mathcal{Y}} \left\{ f(Y) + \frac{1}{2}\|Y - Z\|^2 \right\}$$

  has a closed form solution, denoted by $P_f(Z)$ (non-expansive operator), or at least admits an effective algorithm.
- We can easily compute the directional derivative of

$$\nabla \psi_f(Z) = Z - P_f(Z).$$

- The function $\nabla \psi_f$ is (strongly) semismooth.

Moreau's decomposition

$$P_f(Z) + P_{f^*}(Z) = Z \quad \forall\, Z \in \mathcal{Y}\,.$$

- $f = \delta_{\mathcal{K}}(\cdot)$:

$$P_f(Z) = \Pi_{\mathcal{K}}(Z) \quad \text{and} \quad P_{f^*}(Z) = \Pi_{\mathcal{K}^\circ}(Z)\,,$$

  where $\Pi_D(\cdot)$ is the metric projection over a closed convex set D and $\mathcal{K}^\circ$ is the polar of $\mathcal{K}$.

- $f = \|\cdot\|_{\#}$:

$$P_{f^*}(Z) = \Pi_{\partial f(0)}(Z) = \Pi_{B_*^1}(Z) \quad \text{and} \quad P_f(Z) = Z - P_{f^*}(Z)\,.$$

In both cases, f is absolutely symmetric with respect $\mathbb{R}^{m\times n}$.
One useful inequality – von Neumann's trace inequality:

$$\|\sigma(Y) - \sigma(Z)\| \le \|Y - Z\| \quad \forall\, Y, Z \in \mathbb{R}^{m\times n}\,.$$

More specifically,

- $f = \|\cdot\|_*$. $P_f(Z)$ is just the soft thresholding operator and $P_{f^*}(Z)$ is the metric projection over the unit $\|\cdot\|_2$-ball, i.e.,

$$P_f(Z) = U[(\Sigma - I)_+ \;\; 0]V^T \quad \text{and} \quad P_{f^*}(Z) = U[\min(\Sigma, I) \;\; 0]V^T \,,$$

  where $Z \in \mathbb{R}^{m\times n}$ admits the singular value decomposition (SVD) $Z = U[\Sigma \;\; 0]V^T$ with $\Sigma = \mathrm{diag}(\sigma(Z))$.

- $f = \|\cdot\|_2$.

$$P_{f^*}(Z) = U[g(\sigma(Z)) \;\; 0]V^T \quad \text{and} \quad P_f(Z) = Z - P_{f^*}(Z) \,,$$

  where $g(z) = \mathrm{argmin}_{x\in\mathbb{R}^m}\{\frac{1}{2}\|x - z\|^T \mid \|x\|_1 \le 1\}$. The breakpoint searching (BPS) algorithms[3] can be used, and the computational cost for computing $g(\sigma(Z))$ can be achieved within $\mathcal{O}(m)$ arithmetic operations.

---

[3] M. Held, P. Wolfe, and H. P. Crowder, *Validation of subgradient optimization*, Mathematical Programming, 6 (1974), pp. 62–88.

- $f = \|\cdot\|_{(k)}$.

$$P_{f^*}(Z) = U[g(\sigma(Z)) \;\; 0]V^T \quad \text{and} \quad P_f(Z) = Z - P_{f^*}(Z)\,,$$

where $g(z) = \text{argmin}_{x\in\mathbb{R}^m}\{\frac{1}{2}\|x - z\|^T \mid \|x\|_\infty \le 1,\ \|x\|_1 \le k\}$. By using the BPS algorithms, we can compute $g(\sigma(Z))$ within $\mathcal{O}(m)$ arithmetic operations.

- $f = \|\cdot\|_{(k)^*}$.

$$P_{f^*}(Z) = U[g(\sigma(Z)) \;\; 0]V^T \quad \text{and} \quad P_f(Z) = Z - P_{f^*}(Z)\,,$$

where $g(z) = \text{argmin}_{x\in\mathbb{R}^m}\{\frac{1}{2}\|x - z\|^T \mid \|x\|_{(k)} \le 1\}$. $g(z)$ can be computed at a cost of $\mathcal{O}(k(m-k+1))$ operations[4].

[4] B. WU, C. DING, D. F. SUN, AND K. C. TOH, *On the Moreau-Yosida regularization of the vector k-norm related functions*, Preprint available at http://www.optimization-online.org/DB_FILE/2011/03/2978.pdf.

- $f = \delta_{\mathcal{K}_4}(\cdot)$, where $\mathcal{K}_4 = \mathrm{epi}\,\|\cdot\|_2 = \{(t,X) \in \mathbb{R}\times\mathbb{R}^{m\times n} \mid \|X\|_2 \le t\}$.

$$P_f(t,X) = \Pi_{\mathcal{K}_4}(t,X) \quad \text{and} \quad P_{f^*}(t,X) = (t,X) - P_f(t,X) = \Pi_{-\mathcal{K}_5}(t,X)\,,$$

where $\mathcal{K}_5 = \mathrm{epi}\,\|\cdot\|_* = \{(t,X) \in \mathbb{R}\times\mathbb{R}^{m\times n} \mid \|X\|_* \le t\}$.

$$\Pi_{\mathcal{K}_4}(t,X) = \left(\bar{t}, U[\mathrm{diag}(\bar{y}) \;\; 0]V^T\right)\,,$$

where $(\bar{t},\bar{y}) = g(t,\sigma(X))$ is the unique solution of the following simple quadratic convex optimization problem

$$\begin{array}{ll} \min & \frac{1}{2}\left((\tau - t)^2 + \|y - \sigma(X)\|^2\right) \\ \text{s.t.} & \|y\|_\infty \le \tau\,. \end{array} \tag{7}$$

We can solve (8) at a cost of $\mathcal{O}(m)$ operations.

- $f = \delta_{\mathcal{K}_6}(\cdot)$, where $\mathcal{K}_6 = \mathrm{epi}\,\|\cdot\|_{(k)} = \{(t,X) \in \mathbb{R} \times \mathbb{R}^{m\times n} \mid \|X\|_{(k)} \leq t\}$.

  $P_f(t,X) = \Pi_{\mathcal{K}_6}(t,X) \quad \text{and} \quad P_{f^*}(t,X) = (t,X) - P_f(t,X) = \Pi_{-\mathcal{K}_7}(t,X)\,,$

  where $\mathcal{K}_7 = \mathrm{epi}\,\|\cdot\|_{(k)^*} = \{(t,X) \in \mathbb{R} \times \mathbb{R}^{m\times n} \mid \|X\|_{(k)^*} \leq t\}$.

$$\Pi_{\mathcal{K}_6}(t,X) = \left(\bar{t}, U[\mathrm{diag}(\bar{y}) \;\; 0]V^T\right)\,,$$

  where $(\bar{t},\bar{y}) = g(t,\sigma(X))$ is the unique solution of the following simple quadratic convex optimization problem

$$\begin{array}{ll} \min & \frac{1}{2}\left((\tau - t)^2 + \|y - \sigma(X)\|^2\right) \\ \text{s.t.} & \|y\|_{(k)} \leq \tau \end{array} \quad (8)$$

  We can solve (8) at a cost of $\mathcal{O}(k(m-k+1))$ operations. For more details, see Wu et al. (2011).

The Properties of g — (I)

g is the solution of the corresponding "vector version" M-Y regularization problem.

- g is globally Lipschitz continuous with module 1.
- g is directional differentiable everywhere.
- The sufficient and necessary conditions for the Fréchet differentiability of g are available.
- g is strongly (1-order) G-semismooth everywhere.
  Note that a locally Lipschitz function $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ is said to be ρ-order G-semismooth ($\rho > 0$) at x if for any $h \to 0$,

$$\Phi(x+h) - \Phi(x) - \partial\Phi(x+h)h = O(\|h\|^{1+\rho}) \, .$$

The Properties of g — (II)

A matrix $Q \in \mathbb{R}^{m\times m}$ is said to be a signed permutation matrix if each element of Q has exactly one nonzero entry in each row and each column, that entry being ± 1. Let $|\mathbb{P}|^m$ be the set of all signed permutation matrices in $\mathbb{R}^{m\times m}$.

$$g(z) = Q^T g(Qz) \quad \forall\, z \in \mathbb{R}^m \text{ and } \forall\, Q \in |\mathbb{P}|^m \,.$$

or

$$g(t,z) = (g_1(t,z), g_2(t,z)) = \Big(g_1(t,Qz), Q^T g_2(t,Qz)\Big) \quad \forall\, z \in \mathbb{R}^m \text{ and } \forall\, Q \in |\mathbb{P}|^m \,.$$

Then, we say the function $g : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \times \mathbb{R}^m$ is (mixed) symmetric with respect to $\mathcal{S}^1 \times \mathbb{R}^{m\times n}$. The function $G : \mathcal{S}^1 \times \mathbb{R}^{m\times n} \to \mathcal{S}^1 \times \mathbb{R}^{m\times n}$ defined by

$$G(t,X) := \Big(g_1(t,\sigma(X)), U[\operatorname{diag}(g_2(t,\sigma(X)))\ \ 0]V^T\Big)$$

is said to be the corresponding spectral operator.

The Spectral Operator

Suppose that the function $g : \mathbb{R}^{m_0} \times \mathbb{R}^m \to \mathbb{R}^{m_0} \times \mathbb{R}^m$ is (mixed) symmetric with respect to $\mathcal{S}^{m_0} \times \mathbb{R}^{m \times n}$, i.e.,

$$\mathbf{g}(\mathbf{z}) = (g_1(\mathbf{z}), g_2(\mathbf{z})) = \left(Q_1^T g_1(\mathbf{Q}\mathbf{z}), Q_2^T g_2(\mathbf{Q}\mathbf{z})\right) \quad \forall\, \mathbf{z} \in \mathbb{R}^{m_0} \times \mathbb{R}^m,\ \forall\, \mathbf{Q} \in \mathbb{P}^{m_0} \times |\mathbb{P}|^m,$$

where $\mathbf{Q} = (Q_1, Q_2) \in \mathbb{P}^{m_0} \times |\mathbb{P}|^m$, $\mathbf{Q}\mathbf{z} = (Q_1 z_1, Q_2 z_2)$, and $\mathbb{P}^{m_0}$ is the set of all permutation matrices in $\mathbb{R}^{m_0 \times m_0}$. The function $G : \mathcal{S}^{m_0} \times \mathbb{R}^{m \times n} \to \mathcal{S}^{m_0} \times \mathbb{R}^{m \times n}$ defined by

$$G(Y, Z) := \left(P \mathrm{diag}(g_1(\lambda(Y), \sigma(Z)))P^T, U[\mathrm{diag}(g_2(\lambda(Y), \sigma(Z)))\ \ 0]V^T\right)$$

is said to be the corresponding spectral operator[5].

[5] C. Ding, D. F. Sun, J. Sun and K. C. Toh, *Spectral operator of matrices*, manuscript in preparation, 2011.

The Properties of the Spectral Operator

- $\mathbf{G}$ is Hadamard directionally differentiable at (Y, Z) if and only if $\mathbf{g}$ is Hadamard directionally differentiable at $(\lambda(Y), \sigma(Z))$.
- $\mathbf{G}$ is (continuously) differentiable at (Y, Z) if and only if $\mathbf{g}$ is (continuously) differentiable at $(\lambda(Y), \sigma(Z))$.
- $\mathbf{G}$ is locally Lipschitz continuous near (Y, Z) if and only if $\mathbf{g}$ is locally Lipschitz continuous near $(\lambda(Y), \sigma(Z))$.
- Let $0 < \rho \leq 1$ be given. $\mathbf{G}$ is locally Lipschitz continuous near (Y, Z) and ρ-order G-semismooth at (Y, Z) if and only if $\mathbf{g}$ is locally Lipschitz continuous near $(\lambda(Y), \sigma(Z))$ and ρ-order G-semismooth at $(\lambda(Y), \sigma(Z))$.
- The characterization of B-subdifferenital $\partial_B G$ and Clarke's generalized Jacobian ∂G.

Future work

- The numerical implementation.
- The second order variational analysis for MCPs / MOPs.
- Other approaches for MOPs.

Thank You.