Projection onto spectraplex

Andersen Ang

Mathématique et recherche opérationnelle, UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft: November 18, 2020
Last update: December 24, 2020

Spectraplex

- Spectraplex is an example of spectrahedron, defined as the set

$$\mathcal{S} = \Big\{ \mathbf{X} \in \mathbb{S}^n_+ : \operatorname{Tr} \mathbf{X} = 1 \Big\},$$

  where

  - $\mathbb{S}^n_+$ is the set of n-by-n positive semidefinite matrices
  - $\operatorname{Tr} \mathbf{X} =$ trace of $\mathbf{X}$, which is the sum of diagonal of $\mathbf{X}$, which also equals to the sum of eigenvalue of $\mathbf{X}$

$$\operatorname{Tr} \mathbf{X} = \sum_{ii} \mathbf{X} = \sum_i \lambda_i(\mathbf{X}).$$

- Spectraplex = "spectra" + "simplex", meaning "eigenvalues are inside simplex". i.e., spectraplex is the "semidefinite" analog of simplex.

- Question: given a matrix $\mathbf{Z} \in \mathbb{R}^{n \times n}$, what is the projection of $\mathbf{Z}$ onto the set $\mathcal{S}$?

Formulating the projection onto spectraplex

- Given a set $\mathcal{C}$ and a point $\mathbf{x}_0$, the Euclidean projection onto $\mathcal{C}$ is

$$\operatorname*{argmin}_{\mathbf{x}\in\mathcal{C}} \frac{1}{2}\|\mathbf{x}-\mathbf{x}_0\|_2^2,$$

  i.e., find the point $\mathbf{x}\in\mathcal{C}$ that is closest to $\mathbf{x}_0$. This problem always has a unique solution, details here.

- Given a matrix $\mathbf{Z}\in\mathbb{R}^{n\times n}$, the Euclidean projection onto spectraplex is

$$\operatorname*{argmin}_{\mathbf{X}\in\mathcal{S}} \frac{1}{2}\|\mathbf{X}-\mathbf{Z}\|_F^2,\ \mathcal{S}=\Big\{\mathbf{Y}\in\mathbb{S}_+^n : \operatorname{Tr}\mathbf{Y}=1\Big\},$$

  or equivalently

$$\operatorname*{argmin}_{\mathbf{X}\succeq 0,\operatorname{Tr}\mathbf{X}=1} \frac{1}{2}\|\mathbf{X}-\mathbf{Z}\|_F^2,$$

Structure of the solution

- As positive semidefinite matrices are about eigenvalues, so we consider the eigendecomposition of $\mathbf{Z} = \sum_i \lambda_i \mathbf{z}_i \mathbf{z}_i^\top$, where $\mathbf{z}_i$ is the ith (unit) eigenvector of $\mathbf{Z}$ and λ_i is the ith (nonnegative) eigenvalue of $\mathbf{Z}$.

- The unknown variable matrix $\mathbf{X} \in \mathcal{S}$ can be expressed as the sum of two part: a part that is "perpendicular" to the span of $\mathbf{z}_i$ and a part that is "insider" the span of $\mathbf{z}_i$

$$\mathbf{X} = \mathbf{X}^\perp + \sum_{i=1}^{d} \alpha_i \mathbf{z}_i \mathbf{z}_i^\top,$$

where

$$\begin{aligned} \langle \mathbf{X}^\perp, \mathbf{z}_i \mathbf{z}_i^\top \rangle &= \Big\langle \sum_i \mathbf{x}_i^\perp {\mathbf{x}_i^\perp}^\top, \mathbf{z}_i \mathbf{z}_i^\top \Big\rangle \\ &= \sum_i \langle \mathbf{x}_i^\perp {\mathbf{x}_i^\perp}^\top, \mathbf{z}_i \mathbf{z}_i^\top \rangle \\ &= \sum_i \langle \mathbf{x}_i^\perp \underbrace{{\mathbf{x}_i^\perp}^\top \mathbf{z}_i}_{=0} \mathbf{z}_i^\top \rangle = 0. \end{aligned}$$

Deriving the solution ... (1/3)

- As $\mathbf{X} = \mathbf{X}^{\perp} + \sum_i \alpha_i \mathbf{z}_i \mathbf{z}_i^{\top}$

$$
\begin{aligned}
\|\mathbf{X} - \mathbf{Z}\|_F^2 &= \|\mathbf{X}^{\perp} + \sum_{i=1}^{d} \alpha_i \mathbf{z}_i \mathbf{z}_i^{\top} - \mathbf{Z}\|_F^2 \\
&= \|\mathbf{X}^{\perp} + \sum_{i=1}^{d} \alpha_i \mathbf{z}_i \mathbf{z}_i^{\top} - \sum_{i=1}^{d} \lambda_i \mathbf{z}_i \mathbf{z}_i^{\top}\|_F^2 \\
(\mathbf{X}^{\perp} \text{ orthogonal to } \mathbf{z}_i) &= \|\mathbf{X}^{\perp}\|_F^2 + \|\sum_{i=1}^{d} \alpha_i \mathbf{z}_i \mathbf{z}_i^{\top} - \sum_{i=1}^{n} \lambda_i \mathbf{z}_i \mathbf{z}_i^{\top}\|_F^2, \\
&= \|\mathbf{X}^{\perp}\|_F^2 + \|\sum_{i=1}^{d} (\alpha_i - \lambda_i) \mathbf{z}_i \mathbf{z}_i^{\top}\|_F^2 \\
(\mathbf{z}_i \text{ are orthonormal}) &= \|\mathbf{X}^{\perp}\|_F^2 + \sum_{i=1}^{d} (\alpha_i - \lambda_i)^2.
\end{aligned}
$$

- The projection problem $\underset{\mathbf{X} \succeq 0, \operatorname{Tr} \mathbf{X} = 1}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{X} - \mathbf{Z}\|_F^2$ now becomes

$$
\underset{\substack{\mathbf{X}^{\perp} \\ \sum_i \alpha_i = 1 \\ \alpha_i \geq 0}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{X}^{\perp}\|_F^2 + \frac{1}{2} \sum_{i=1}^{d} (\alpha_i - \lambda_i)^2.
$$

Deriving the solution ... (2/3)

$$\underset{\substack{\mathbf{X}^\perp \\ \sum_i \alpha_i = 1 \\ \alpha_i \geq 0}}{\text{argmin}} \ \frac{1}{2}\|\mathbf{X}^\perp\|_F^2 + \frac{1}{2}\sum_{i=1}^{d}(\alpha_i - \lambda_i)^2.$$

- As $\mathbf{X}^\perp$ is $\perp$ to all $\mathbf{z}_i$, the two parts are independent to each other. To minimize the part $\|\mathbf{X}^\perp\|_F^2$, we can just set $\mathbf{X}^\perp = 0$, this reduces the projection problem to

$$\underset{\substack{\sum_i \alpha_i = 1 \\ \alpha_i \geq 0}}{\text{argmin}} \ \frac{1}{2}\sum_{i=1}^{d}(\alpha_i - \lambda_i)^2.$$

- Let $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_n]$ and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_n]$.
  Now projection onto spectraplex = projection onto simplex on $\boldsymbol{\lambda}$

$$\underset{\substack{\boldsymbol{\alpha}^\top \mathbf{1} = 1 \\ \boldsymbol{\alpha} \geq 0}}{\text{argmin}} \ \frac{1}{2}\|\boldsymbol{\alpha} - \boldsymbol{\lambda}\|_2^2$$

Deriving the solution ... (3/3)

- The problem $\underset{\substack{\boldsymbol{\alpha}^\top \mathbf{1}=1 \\ \boldsymbol{\alpha}\geq 0}}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{\alpha}-\boldsymbol{\lambda}\|_2^2$ has closed-form solution

  $$\boldsymbol{\alpha}^* = [\boldsymbol{\lambda} - \mu^* \mathbf{1}_n]_+,$$

  where the Lagrangian multiplier μ^* is the root to the following piecewise-linear equation

  $$\sum_{i=1}^n \max\left(0, \lambda_i - \mu\right) = 1.$$

  See here for details on the projection onto simplex.

- As spectraplex is the semidefinite analog of the simplex, so it makes sense that the sol. is given by projection onto simplex on $\boldsymbol{\lambda}$.

Last page - summary

- Given a matrix $\mathbf{Z} \in \mathbb{R}^{n \times n}$, the projection of $\mathbf{Z}$ onto spectraplex is

$$\underset{\mathbf{X} \succeq 0, \operatorname{Tr} \mathbf{X} = 1}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2.$$

- The solution is in the form of $\mathbf{X} = \sum_i \alpha_i \mathbf{z}_i \mathbf{z}_i^\top$, where $\mathbf{z}_i$ is the eigenvector of $\mathbf{Z}$, and α_i can be computed as

$$\boldsymbol{\alpha}^*(\mathbf{X}) = [\boldsymbol{\lambda}(\mathbf{Z}) - \mu^* \mathbf{1}_n]_+,$$

where the Lagrangian multiplier μ^* is the root to the piecewise-linear equation

$$\sum_{i=1}^{n} \max\left(0, \lambda_i - \mu\right) = 1.$$

End of document