A GRADUATE COURSE

IN

OPTIMIZATION

A GRADUATE COURSE

IN

OPTIMIZATION

CIE6010 Notebook

Prof. Yin Zhang

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the CIE6010 in fall semester, 2018.

CUHK(SZ)

Notations and Conventions

X	Set
$\inf X \subseteq \mathbb{R}$	Infimum over the set X
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i,j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i,j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i,j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i,j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 15

Week14

15.1. Wednesday

15.1.1. Conic Programming

The primal conic programming is given by:

$$\begin{aligned} \min \quad & \langle \boldsymbol{C}, \boldsymbol{X} \rangle \\ & \langle \boldsymbol{a}_i, \boldsymbol{X} \rangle = \boldsymbol{b}_i, i = 1, \dots, m \\ & \boldsymbol{X} \in \mathcal{K} \end{aligned}$$

LP, SDP, SOCP.

The dual form is given by:

$$\begin{aligned} \max \quad & \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y} \\ & \textstyle\sum_{i=1}^{m} y_i \boldsymbol{a}_i + S = \boldsymbol{C} \\ & S \in \mathcal{K}^* \end{aligned}$$

Most problem setting is self-dual.

$$\begin{aligned} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{A}^* \boldsymbol{y} + S = \boldsymbol{C} \end{aligned}$$

Fermat-Weber location problem: Given a set of points $\boldsymbol{p}_i$, our goal is to

$$\min_{\boldsymbol{y} \in \mathbb{R}^2} \sum_{i=1}^{m} \|\boldsymbol{y} - \boldsymbol{p}_i\|$$

Note that it is norm 2 instead of its square. It's relatively complicated problem.

Introduce variables $\eta_1, \ldots, \eta_m$:

$$\begin{aligned} \min \quad & \eta_1 + \cdots + \eta_m \\ & \|\boldsymbol{y} - \boldsymbol{p}_i\| \leq \eta_i, \quad i = 1, \ldots, m \end{aligned}$$

Or equivalently,

$$\begin{aligned} \min_{\boldsymbol{y},\boldsymbol{\eta},\boldsymbol{z}} \quad & \mathbf{1}^{\mathrm{T}}\boldsymbol{\eta} \\ & \boldsymbol{z}_i + \boldsymbol{y} = \boldsymbol{p}_i, \quad i = 1, \ldots, m \\ & \|\boldsymbol{z}_i\| \leq \eta_i \end{aligned}$$

The dual problem is

$$\begin{aligned} \max \quad & \textstyle\sum_{i=1}^m \boldsymbol{p}_i^{\mathrm{T}} \boldsymbol{x}_i \\ & \textstyle\sum_{i=1}^m \boldsymbol{x}_i = 0 \\ & \|\boldsymbol{x}_i\| \leq 1, i = 1, \ldots, m \end{aligned}$$

the last constraint is the second order cone.

For quadratic constraint with $\boldsymbol{A} \succeq 0$:

$$(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}) - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{y} - \boldsymbol{d} \leq 0$$

which is equivalent to say

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b} \\ (\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})^{\mathrm{T}} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{y} + \boldsymbol{d} \end{bmatrix} \succeq 0$$

QCQP can be converted into SDP when $\boldsymbol{A}$ is convex.

$$\left(\boldsymbol{c}^{\mathrm{T}} y + \boldsymbol{d} - \frac{1}{4}\right)^2 + \|\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}\|^2 \leq \left(\boldsymbol{c}^{\mathrm{T}} y + \boldsymbol{d} + \frac{1}{4}\right)^2$$

Thus QCQP can be converted into SOCP as well.

15.1.2. Algorithm to solve conic programming

$$\begin{aligned}\min \quad & \langle \boldsymbol{C}, \boldsymbol{X} \rangle \\ & \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle = \boldsymbol{b}_i, i = 1, \ldots, m \\ & \boldsymbol{X} \succeq 0\end{aligned}$$

$$\begin{aligned}\max \quad & \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y} \\ & \sum y_i A_i + Z = C \\ & Z \succeq 0\end{aligned}$$

If both (P) and (D) are **strictly feasible**, then there exists $\boldsymbol{X}^*, \boldsymbol{y}^*$ (feasible) such that

$$\langle \boldsymbol{C}, \boldsymbol{X}^* \rangle = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}^*$$

which follows that

$$\langle \boldsymbol{C}, \boldsymbol{X} \rangle - \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y} = \langle \boldsymbol{X}, \boldsymbol{Z} \rangle = 0$$

It suffices to let

$$0 = \operatorname{trace}(\boldsymbol{X}\boldsymbol{Z}) = \sum_i \lambda_i (\mathbf{Z}^{1/2} \mathbf{X} \mathbf{Z}^{1/2})$$

which implies

$$\mathbf{Z}^{1/2} \mathbf{X} \mathbf{Z}^{1/2} = 0 \Longleftrightarrow \mathbf{Z}^{1/2} \mathbf{X}^{1/2} = 0$$