# COMPUTATIONAL METHODS IN OPTIMIZATION

November 6, 2023

Today: First lecture on IPMs and conic programming (SDP)
Coming up: More conic programming + examples

Tomorrow: Lab session on QP
Coming up: 2 more labs/tutorials (Nov 14 + Nov 21)
December 5: 2 lab sessions
Week of December 11: 4 lab sessions } course project

## Recall

(LP) $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^T x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0_{\mathbb{R}^n}$

(QP) $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} x^T Q x + c^T x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0_{\mathbb{R}^n}$

→ IPMs originally developed for LPs, very easy to adapt them to quadratic programs

→ They handle the constraint $x \geq 0$ quite well.

Q) How much can we generalize that approach?

⇒ Today: SDPs, that are problems involving matrix variables

## 1 Semidefinite programs (SDPs)

### Matrix tools and notations

- $M \in \mathbb{R}^{n \times n}$ is positive semidefinite (psd) if $x^T M x \geq 0 \quad \forall x \neq 0_{\mathbb{R}^n}$

  ⇒ We write $M \succeq 0$ (0 here: $0_{\mathbb{R}^{n \times n}}$)

  $\left( M \succeq 0 \neq \underbrace{M \geq 0}_{M_{ij} \geq 0} \right)$

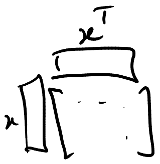
- $M \in \mathbb{R}^{n \times n}$ is positive definite if $x^T M x > 0 \quad \forall x \neq 0_{\mathbb{R}^n}$

  ⇒ We write $M \succ 0$

⇒ We will only work with symmetric psd matrices $(M^T = M)$

- $\mathcal{S}^{n \times n}$ : set of symmetric matrices in $\mathbb{R}^{n \times n}$

Ex) $\forall x \in \mathbb{R}^n$, $\underbrace{x}_{n \times 1}\underbrace{x^T}_{1 \times n} \in \mathbb{R}^{n \times n}$




- $(xx^T)^T = (x^T)^T x^T = xx^T$

$xx^T \in \mathcal{S}^{n \times n}$

- $\forall y \in \mathbb{R}^n$, $y \neq 0_{\mathbb{R}^n}$,

$y^T(xx^T)y = (y^Tx)(x^Ty) = (x^Ty)^2 \geq 0$

$\Rightarrow \quad xx^T \succeq 0$

- $\forall (A,B) \in (\mathcal{S}^{n \times n})^2$,

$$A \bullet B := \text{trace}(AB) = \text{trace}(BA) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij}$$

"Matrix inner product"

(Equivalent to $a^Tb$ for vectors)

$$\|A\|_F := (A \bullet A)^{1/2} \qquad \text{Frobenius norm}$$

<u>NB</u>: $\forall x \in \mathbb{R}^n$, $\forall A \in \mathcal{S}^{n \times n}$, $\quad A \bullet (xx^T) = x^TAx$

## Semidefinite program in standard form

linear equality constraints (w.r.t. $X$)

$$\underset{X \in \mathcal{S}^{n \times n}}{\text{minimize}} \quad C \bullet X \quad \text{s.t.} \quad A_i \bullet X = b_i \quad i = 1..m$$
$$X \succeq 0$$

Matrix optimization

$C \in \mathcal{S}^{n \times n}$

positive semidefiniteness

$A_1, \dots, A_m \in \mathcal{S}^{n \times n}$

$A_i \bullet X \Leftrightarrow a_i^T x$

$[b_i] = b \in \mathbb{R}^m$

<u>Remark</u>: The problem could be written as a vector optimization problem, but the constraints would be more complex to define and we would not take

advantage of the matrix structure

$$[X_{ij}] = X \in \mathbb{R}^{n\times n} \implies x = \begin{bmatrix} X_{11} \\ \vdots \\ X_{1n} \\ X_{21} \\ \vdots \\ X_{n1} \\ \vdots \\ X_{nn} \end{bmatrix} \in \mathbb{R}^{n^2}$$

<u>Remark</u>: LP and QP are special cases of SDP

$X = \mathrm{diag}(x)$

LP

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \ c^T x \quad \text{s.t.} \quad a_i^T x = b_i \quad i=1..m, \quad x \geq 0$$

$x_i \geq 0 \ \forall i=1..n$

$$\implies \underset{X\in\mathbb{S}^{n\times n}}{\text{minimize}} \ C\cdot X \quad \text{s.t.} \quad A_i \cdot X = b_i \quad i=1..m$$
$$X_{ij} = 0 \ \forall i\neq j$$
$$X \succeq 0$$

linear equality constraints

$$C = \mathrm{diag}(c) = \begin{bmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{bmatrix}$$

$$A_i = \mathrm{diag}(a_i)$$

$$X_{ij} = 0 \iff E_{ij}\cdot X = 0 \qquad E_{ij} = \begin{array}{c} j \\ \begin{pmatrix} 0 & \vdots & 0 \\ 0 & 1 & \cdots \\ & & 0 \end{pmatrix} \end{array} i$$

For any $X \in \mathbb{S}^{n\times n}$ that satisfies the linear equality constraints

$$X \succeq 0 \iff X_{ii} \geq 0 \quad \forall i=1..n$$

## ② Solving SDPs

$$(P) \quad \underset{X\in\mathbb{S}^{n\times n}}{\text{minimize}} \ C\cdot X \quad \text{s.t.} \quad A_i \cdot X = b_i \quad i=1..m$$
$$X \succeq 0$$

<u>pb data</u>: $A_1, \dots, A_m \in \mathbb{S}^{n\times n}$, $b\in\mathbb{R}^m$, $C \in \mathbb{S}^{n\times n}$

The dual of (P) is an SDP defined by

$$(D) \quad \underset{\substack{y \in \mathbb{R}^m \\ S \in \mathbb{S}^{n\times n}}}{\text{maximize}} \quad b^T y \quad \text{s.t.} \quad \sum_{i=1}^m y_i A_i + S = C, \quad S \succeq 0$$

(For LP: $\underset{\substack{y \in \mathbb{R}^m \\ s \in \mathbb{R}^n}}{\text{maximize}} \; b^T y$ s.t. $A^T y + s = c$, $s \geq 0$ $\qquad A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}$ $\quad A^T y = \sum_{i=1}^m y_i a_i$)

↳ Define $\mathcal{A} : \mathbb{S}^{n\times n} \longrightarrow \mathbb{R}^m$, $X \longmapsto [A_i \bullet X]_{i=1}^m$

and $\mathcal{A}^* : \mathbb{R}^m \longrightarrow \mathbb{S}^{n\times n}$, $y \longmapsto \sum_{i=1}^m y_i A_i$

Then (P) and (D) can be rewritten as

$$(P) \quad \underset{X \in \mathbb{S}^{n\times n}}{\text{minimize}} \quad C \bullet X \quad \text{s.t.} \quad \mathcal{A} X = b, \quad X \succeq 0$$

$\mathcal{A} X \equiv \mathcal{A}(X)$

linear operator

$$(D) \quad \underset{\substack{y \in \mathbb{R}^m \\ S \in \mathbb{S}^{n\times n}}}{\text{maximize}} \quad b^T y \quad \text{s.t.} \quad \mathcal{A}^* y + S = C, \quad S \succeq 0$$

$\mathcal{A}^* y \equiv \mathcal{A}^*(y)$

↳ Optimality conditions: Suppose that both (P) and (D) have interior feasible points

(i.e. $\exists X \in \mathbb{S}^{n\times n}$, $\mathcal{A} X = b$, $X \succ 0$
$\exists y \in \mathbb{R}^m$, $\exists S \in \mathbb{S}^{n\times n}$, $\mathcal{A}^* y + S = C$, $S \succ 0$)

Then $(\bar{X}, \bar{y}, \bar{S}) \in \mathbb{S}^{n\times n} \times \mathbb{R}^m \times \mathbb{S}^{n\times n}$ is a primal-dual solution of (P) if and only if

$$A\bar{X} = b,\quad \bar{X} \succeq 0 \quad , \quad A^*\bar{y} + \bar{S} = C,\quad \bar{S} \succeq 0 \quad , \quad \bar{X} \cdot \bar{S} = 0$$

$\bar{X}$ feasible for (P)

$(\bar{y}, \bar{S})$ feasible for (D)

complementarity

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \bar{X}_{ij} \bar{S}_{ij} = 0$$

↳ These conditions are very similar to that of LP and QP

but $\quad \begin{matrix} \bar{X} \cdot \bar{S} = 0 \\ \bar{X} \succeq 0 \\ \bar{S} \succeq 0 \end{matrix} \quad \cancel{(\Rightarrow)} \quad \bar{X}_{ij} \bar{S}_{ij} = 0 \quad \forall (i,j)$

in general

However, $\quad \begin{matrix} \bar{X} \cdot \bar{S} = 0 \\ \bar{X} \succeq 0 \\ \bar{S} \succeq 0 \end{matrix} \quad \Rightarrow \quad \begin{cases} \bar{X}\bar{S} = 0 \\ \bar{S}\bar{X} = 0 \end{cases}$

There are several ways to define an interior-point method for SDPs depending on the condition used to replace complementarity ($X \cdot S = 0 \Rightarrow$ ? ), unlike LP, and QP, ($x^T s = 0 \Rightarrow x_i s_i = 0 \ \forall i = 1..n$)

## A (classical) example of IPMs for SDPs

$$\begin{matrix} X \cdot S = 0 \\ X \succeq 0 \\ S \succeq 0 \end{matrix} \quad \Rightarrow \quad XS = 0$$

For any $\begin{matrix} X \succ 0 \\ S \succ 0 \end{matrix}$, $\quad XS \succ 0 \Rightarrow X \cdot S \neq 0$

IPMs want to stay in the interior of the domain, i.e. only consider positive definite matrices

## Iteration k of IPM

- Start with $(X^k, y^k, S^k)$ primal-dual strictly feasible
  $(\mathcal{A}X^k = b,\ X^k \succ 0,\ \mathcal{A}^* y^k + S^k = C^k,\ S^k \succ 0)$

- Compute a Newton step based on the nonlinear system of equations

system of $m+2n^2$ equations with $m+2n^2$ unknowns

$$\left\{\begin{array}{ll} \mathcal{A}X = b & \text{) } m \text{ equations} \\ \mathcal{A}^* y + S = C & \text{) } n^2 \text{ equations} \\ XS = 0_{\mathbb{R}^{n\times n}} & \text{) } n^2 \text{ equations} \end{array}\right.$$

Matrix-matrix product

- Define $(X^{k+1}, y^{k+1}, S^{k+1})$ based on $(X^k, y^k, S^k)$ and the Newton step so that the new point is primal-dual strictly feasible

↳ In theory, SDPs can be solved to arbitrary accuracy using IPMs

↳ In practice, the dimension of the problem quickly becomes an issue

$\Rightarrow$ Most solvers handle SDPs with $n \approx 10^3$ $(n^2 \approx 10^6)$
(vs $n \approx 10^6$ for LP)

$\Rightarrow$ Key challenge: Compute the Newton step

(Recall: For LP,

$$\underbrace{\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S^k & 0 & X^k \end{bmatrix}}_{\text{large matrix}} \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{bmatrix} = -\begin{bmatrix} 0 \\ 0 \\ X^k S^k e \end{bmatrix}$$

diagonal matrices

vectors )

For SDPs, the Newton system looks like

$$\mathcal{A}\,\Delta X^k = 0$$

$$\mathcal{A}^*\Delta y^k + \Delta S^k = 0$$

$$\overbrace{(X^k \bullet S^k)}^{\in \mathbb{R}}\underbrace{\left((X^k)^{-1} \circ (X^k)^{-1}\right)}_{\text{operator}}\underbrace{\Delta X^k}_{\in \mathbb{R}^{n\times n}} + \underbrace{\Delta S^k}_{\in \mathbb{R}^{n\times n}} = \underbrace{X^k S^k}_{\in \mathbb{R}^{n\times n}}$$

$\forall U \in \mathbb{S}^{n\times n}$, $(P \circ Q)(U) = \frac{1}{2}\left(PUQ^T + QUP^T\right)$