262B-Lecture 24

Date created: 2021.04.22
N. of Pages: 16

$$\begin{cases} \min_x & f(x) \\ \text{s.t.} & h_i(x) = 0 \quad i = 1, \dots, m \\ & g_j(x) \le 0 \quad j = 1, \dots, r \end{cases}$$

approximate this arbitrarily precisely under smoothness

$\downarrow$ polynomial optimization

$$\left.\begin{array}{ll} \min_x & p_0(x) \\ \text{s.t} & p_i(x) \le 0 \quad i = 1, \dots, k \end{array}\right\} \Rightarrow \text{Canonical form}$$

This includes mixed-integer programming.

Ex: $\min \; x^T Q x \quad \text{s.t.} \quad x_i \in \{-1, 1\} \quad i = 1, \dots, n$

$$\downarrow$$

$$\min \; x^T Q x \quad \text{s.t.} \quad x_i^2 - 1 = 0 \quad i = 1, \dots, n$$

Ex: $x_i \in \{a_i, a_i + 1, a_i + 2, \dots, b_i\}$

$\downarrow$ base $\underline{\underline{2}}$

$x_i = y_{i0} + y_{i1} \times 2 + y_{i2} \times 2^2 + \dots \longrightarrow$ Linear in $\underline{\underline{y}}$

$$\downarrow$$

$y_{ij} \in \{0, 1\} \longrightarrow y_{ij}^2 - y_{ij} = 0 \quad , \quad a_i \le x_i \le b_i$

polynomial: $\underbrace{x_1^2} + \underbrace{5x_1x_2x_3} + \underbrace{4x_3^5}$

monomial

degree: 2 3 5

high-order monomials can be broken down to monomials of degrees 1 and 2 by introducing new variables.

Ex:
$$\min_{x_1, x_2} \; x_1^4 x_2 + x_1$$
$$\text{s.t.} \quad x_1^2 + x_2^2 = 1$$

$$x_1^4 x_2 = \underbrace{\underbrace{x_1^2}_{x_3} \times \underbrace{x_1^2}_{x_3}}_{x_4} \times x_2$$

$$\Rightarrow \quad \left.\begin{aligned} \min_{x_1, x_2, x_3, x_4} \quad & x_4 x_2 + x_1 \\ \text{s.t.} \quad & x_1^2 + x_2^2 - 1 = 0 \\ & x_1^2 - x_3 = 0 \\ & x_3^2 - x_4 = 0 \end{aligned}\right\} \text{linear \& quadratic}$$

$\Rightarrow$ every polynomial optimization can be equivalently converted to a non-convex quadratically-constrained quadratic program (QCQP):

$$\min_{x \in \mathbb{R}^n} \; x^T P_0 x + 2 q_0^T x + r_0$$

$$\text{s.t.} \quad x^T \underbrace{P_i}_{\text{matrix}} x + 2 \underbrace{q_i}_{\text{vector}}^T x + \underbrace{r_i}_{\text{scalar}} \leq 0 \qquad i = 1, \dots, m$$

- $P_0, \dots, P_m$ may be sign-indefinite $\Rightarrow$ NP-hard problem

- Compute the dual to get a lower bound.

$$L(x, \mu) = (x^T P_0 x + 2 q_0^T x + r_0) + \sum_{i=1}^{m} \mu_i (x^T P_i x + 2 q_i^T x + r_i)$$

$$= x^T \underbrace{\Big(P_0 + \sum_{i=1}^{m} \mu_i P_i \Big)}_{P(\mu)} x + 2 \underbrace{\Big(q_0 + \sum_{i=1}^{m} \mu_i q_i \Big)}_{q(\mu)}^T x + \underbrace{\Big(r_0 + \sum_{i=1}^{m} \mu_i r_i \Big)}_{r(\mu)}$$

$$\min_x L(x,\mu) = \min_x x^T P(\mu) x + 2\, q(\mu)^T x + r(\mu)$$

$$\min \; x^T A x + 2 b^T x + C$$

$$\Big\downarrow \text{ FOC}$$

$$Ax + b = 0$$

$$\Big\downarrow$$

$$x = \underbrace{-A^+ b}_{\text{pseudo-inverse of } A} + \underbrace{N(A)}_{\text{null space of } A}$$

- If $A \not\succeq 0$ $\Rightarrow$ optimal obj value $= -\infty$
- If $A \succ 0$ $\Rightarrow$ $x_* = -A^{-1} b$ (unique)
- If $A \succeq 0$ with zero eigenvalues:

Ex: $\min \; x_1^2 + x_2 \longrightarrow$ optimal obj $= -\infty$

- If $b \notin \text{range}(A)$ $\Rightarrow$ optimal obj $= -\infty$

(you can find x_0 in $N(A)$ s.t. $A x_0 = 0$, $b^T x_0 \neq 0$ $\Rightarrow$ add x_0 to solution

- If $b \in \text{range}(A) \implies$ optimal obj = finite

optimal obj value =

$$(-A^+b + N(A))^T A (-A^+b + N(A))$$

$$+ 2b^T(-A^+b + N(A)) + C$$

$$= b^TA^+b - 2b^TA^+b + C = -b^TA^+b + C$$

$$\min_x L(x,\mu) = \min_x x^T P(\mu) x + 2q(\mu)^T x + r(\mu)$$

$$= \begin{cases} -q(\mu)^T P(\mu)^+ q(\mu) + r(\mu) & \text{if } P(\mu) \succeq 0, \ q(\mu) \in \text{range of } P(\mu) \\ -\infty & \text{otherwise} \end{cases}$$

$$\implies \max_{\mu \in \mathbb{R}^m} \quad -q(\mu)^T P(\mu)^+ q(\mu) + r(\mu)$$

s.t. $\quad \mu \geq 0$

$\quad P(\mu) \succeq 0$

$\quad q(\mu) \in \text{range of } P(\mu)$

or :

$$\max_{\substack{\mu \in \mathbb{R}^m \\ \alpha \in \mathbb{R}}} \quad \alpha$$

s.t. $\quad \mu \geq 0$

Schur Complement

$$\alpha \leq -q(\mu)^T P(\mu)^+ q(\mu) + r(\mu)$$

$$P(\mu) \succeq 0$$

$$q(\mu) \in \text{range of } P(\mu)$$

$\Updownarrow$

$$\max_{\mu, \alpha} \quad \alpha$$

s.t. $\quad \mu \geq 0$

$$\begin{bmatrix} P(\mu) & q(\mu) \\ q(\mu)^T & r(\mu) - \alpha \end{bmatrix} \succeq 0$$

LMI

SDP

$P(\mu), q(\mu), r(\mu)$ are linear in μ

LMI : $\sum_{i=1}^{m} \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix} \mu_i + \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}$

$+ \begin{bmatrix} 0 & 0 \\ 0 & -\alpha \end{bmatrix} \succeq 0$

polynomial opt $\xrightarrow{\text{Dual}}$ SDP $\xrightarrow{\text{Dual}}$ SDP ?

$\mu \geq 0 \longrightarrow \lambda \geq 0$

$\begin{bmatrix} P(\mu) & q(\mu) \\ q(\mu)^T & r(\mu) - \alpha \end{bmatrix} \longrightarrow Y \succeq 0$, $Y \in \mathbb{R}^{(n+1)\times(n+1)}$

$\max \ \alpha \longrightarrow \min \ -\alpha$

$$L(\underbrace{\mu, \alpha}_{\text{primal}}, \underbrace{\lambda, Y}_{\text{dual}}) = -\alpha - \mu^T \lambda - \left\langle \begin{bmatrix} P(\mu) & q(\mu) \\ q(\mu)^T & r(\mu) - \alpha \end{bmatrix}, Y \right\rangle$$

$$= (-1 + \underbrace{Y_{n+1,n+1}}_{\text{Last entry of } Y}) \alpha - \left\langle \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}, Y \right\rangle$$

$$- \sum_{i=1}^{m} \left(\left\langle \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix}, Y \right\rangle + \lambda_i \right) \mu_i$$

Dual of Dual:

$$\max_{\substack{Y \in \mathbb{R}^{(n+1)\times(n+1)} \\ \lambda}} \; - \left\langle \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}, Y \right\rangle$$

$$\text{s.t.} \quad \left\langle \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix}, Y \right\rangle + \lambda_i = 0 \qquad i = 1, \dots, m$$

$$\lambda_i \geq 0 \qquad i = 1, \dots, m$$

$$Y \succeq 0$$

$$1 - Y_{n+1, n+1} = 0$$

$$\Updownarrow$$

$$\min_{Y} \; \left\langle \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}, Y \right\rangle$$

$$\text{s.t.} \quad \left\langle \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix}, Y \right\rangle \leq 0$$

$$Y_{n+1, n+1} = 1$$

$$Y \succeq 0$$

Dual of Dual of QCQP, provides a lower bound

This is the same as SDP relaxation.

$$\min_{x} \; x^T P_0 x + 2q_0^T x + r_0$$

$$\text{s.t. } x^T P_i x + 2q_i^T x + r_i \leq 0 \qquad i=1,\dots,m$$

$$\Updownarrow$$

$$\min_{x} \; \begin{bmatrix} x^T & 1 \end{bmatrix} \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}$$

$$\text{s.t. } \begin{bmatrix} x^T & 1 \end{bmatrix} \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0 \qquad i=1,\dots,m$$

$$\Updownarrow$$

$$\min_{x \in \mathbb{R}^n} \; \left\langle \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}, \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x^T & 1 \end{bmatrix} \right\rangle$$

$$\text{s.t. } \left\langle \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix}, \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x^T & 1 \end{bmatrix} \right\rangle \leq 0 \qquad i=1,\dots,m$$

$$\Updownarrow$$

$$\min_{Y \in \mathbb{R}^{(n+1)\times(n+1)}} \; \left\langle \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix}, Y \right\rangle$$

$$\left\langle \begin{bmatrix} P_i & q_i \\ q_i^T & r_i \end{bmatrix}, Y \right\rangle \leq 0 \qquad i=1,\dots,m$$

s.t. $\quad Y \succeq 0$

$\quad Y_{n+1,n+1} = 1$

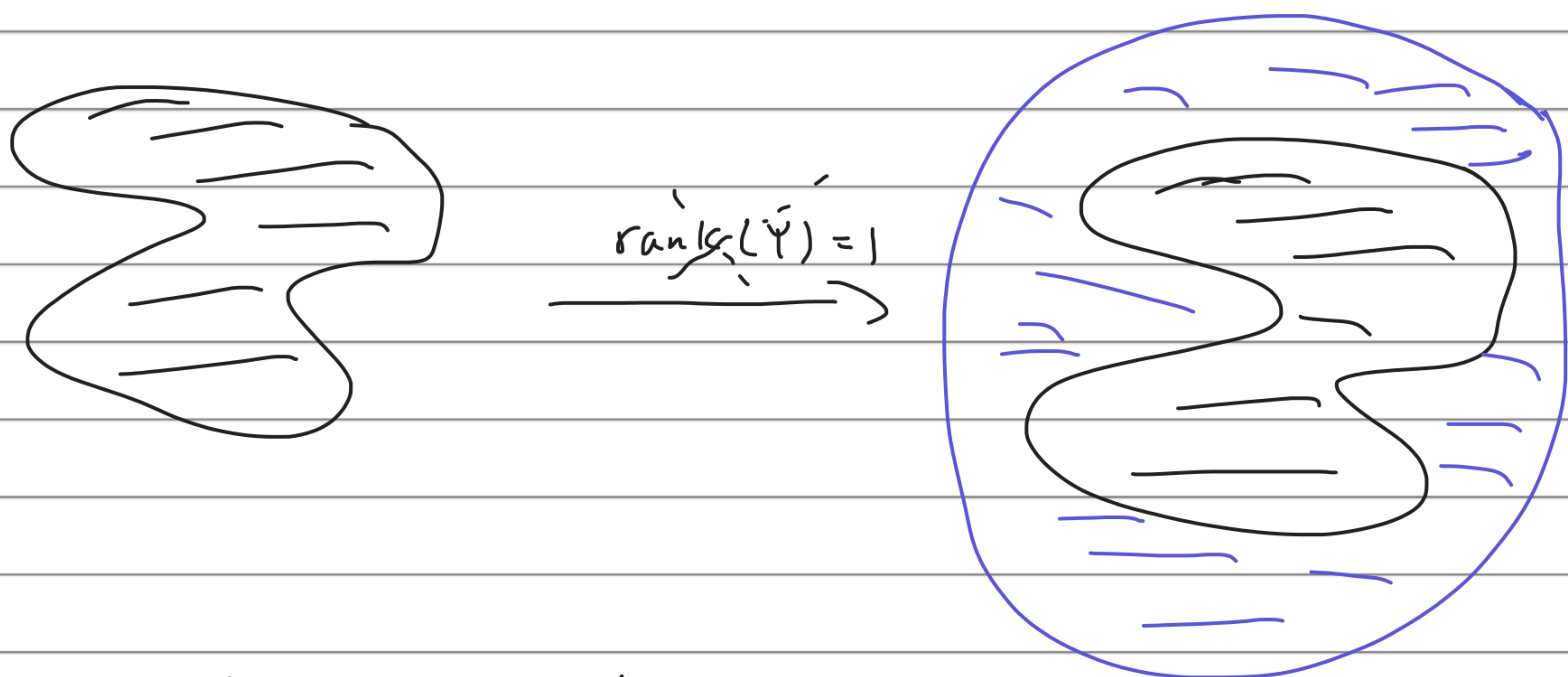
$\quad \text{rank}(Y) = 1$

$$Y = \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x^T & 1 \end{bmatrix}$$

- This is SDP + non-convex rank-1 constraint.

- All non-convexity of polynomial optimization is pushed into a rank constrained.

$\Rightarrow$ Drop it $\longrightarrow$ SDP relaxation



Dual of Dual of non-convex QCQP = SDP relaxation

S-Lemma : (nonlinear version of Farkas)

$m=1$: $\min \; x^T P_0 x + 2q_0^T x + r_0$

$\text{s.t.} \; x^T P_1 x + 2q_1^T x + r_1 \leq 0$

- Assume: $\exists \bar{x}$: $\bar{x}^T P_1 \bar{x} + 2q_1^T \bar{x} + r_1 < 0$

or $\left\langle \begin{bmatrix} P_1 & q_1 \\ q_1^T & r_1 \end{bmatrix}, \begin{bmatrix} \bar{x} \\ 1 \end{bmatrix} [\bar{x}^T \; 1] \right\rangle < 0$

$\Rightarrow$ slater holds for dual of dual.

$\Rightarrow$ can be shown that :

zero duality gap $\rightarrow$ SDP relaxation

is exact $\rightarrow$ $\exists$ rank-1 solution Y_*

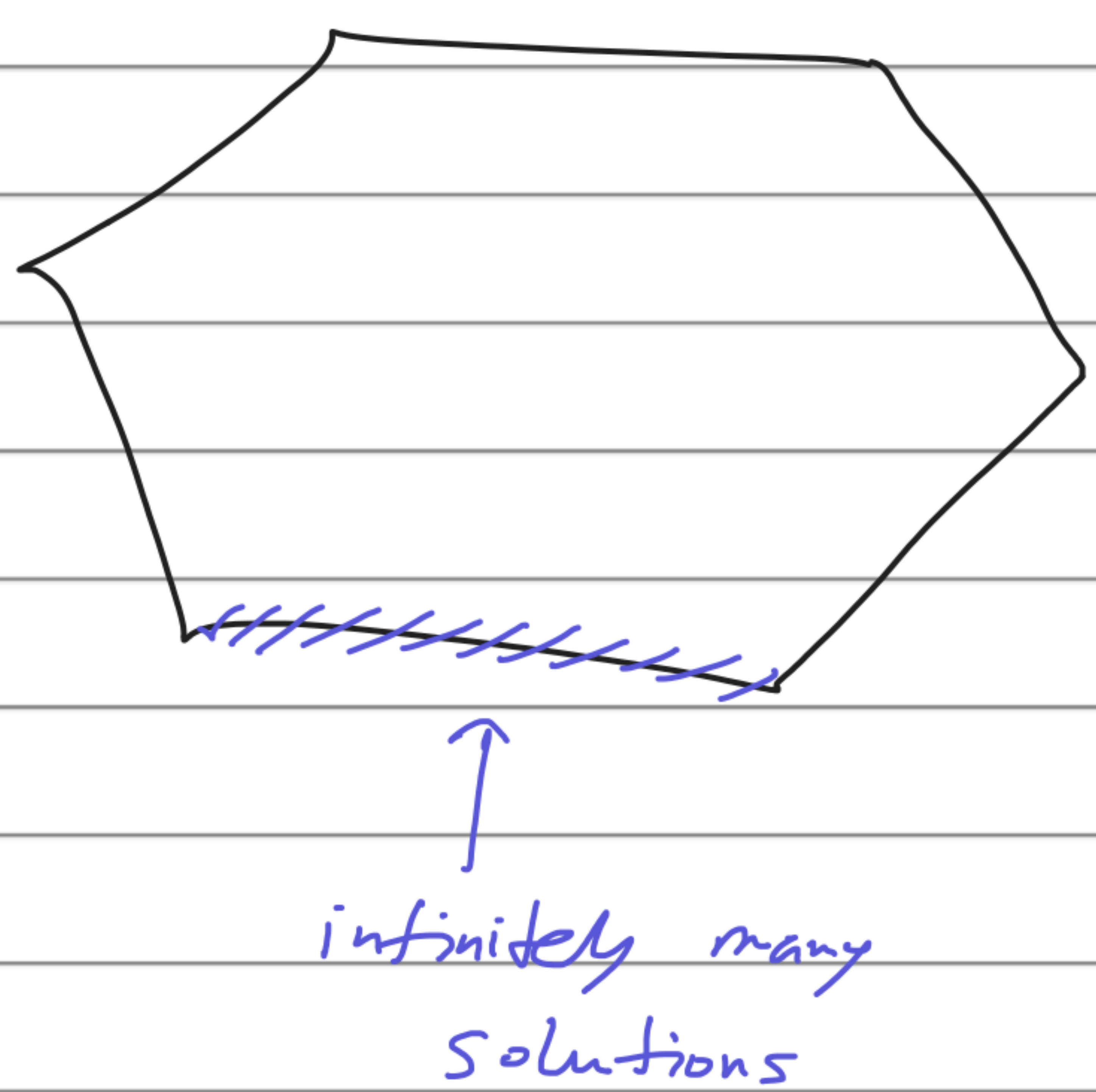
(Proof: convexity of the set used to analyse geometric multipliers)

- Rank of SDP solution $= 1$ $\longrightarrow$ find a global solution x_*
- If the rank is not 1, is it close to 1?

General SDP:

$$\min_{X} \ \langle D, X \rangle$$
$$\text{s.t.} \ \langle A_i, X \rangle = b_i \quad i = 1, \dots, m$$
$$X \succeq 0$$

Let X_* denote a solution to SDP.



general: a face of solutions, interior of face = high-rank solutions, boundary = low-rank solutions

- we may have infinitely many solutions

- Can we use X_* to design a lower-rank solution?

$$X_* = \text{rank } \underline{\underline{k}} \longrightarrow X_* = \underbrace{U}_{n \times k} \underbrace{U^T}_{k \times n}$$

Define $X(\Delta) = U(I + \underbrace{\alpha}_{\text{scalar to be designed}} \underbrace{\Delta}_{k \times k \text{ matrix to be found.}}) U^T$

Also, $X(0) = X_*$

If $\langle U^T D U, \Delta \rangle = 0$ ①

$\langle U^T A_i U, \Delta \rangle = 0 \qquad i = 1, \ldots, m$ ②

$I + \alpha \Delta \succeq 0$ ③

$I + \alpha \Delta = \text{singular}$ ④

1, 2, 3 $\Longrightarrow$ $X(\Delta)$ is another solution

since objective & constraint values won't change

4 $\Longrightarrow$ $\operatorname{rank}(X(\Delta)) < \operatorname{rank}(X_*)$

Always exists α : 3, 4 are satisfied.

If 2 is satisfied $\Longrightarrow$ 1 is satisfied (KKT)

Instead 1, 2, 3, 4 $\longrightarrow$ just need 2.

$\exists\, \Delta$ satisfying 2 if there are more $\underbrace{\text{variables}}_{\frac{k(k+1)}{2}}$ than $\underbrace{\text{equations}}_{m}$

If $k(k+1)/2 > m$ $\Longrightarrow$ we reduce the rank.

Pataki theorem: SDP has a solution of rank r s.t. $\frac{r(r+1)}{2} \leq m$

special case: $m = 1$ $\Longrightarrow$ $r = 1$ $\Longrightarrow$ rank-1 solution (S-Lemma)

$\Longrightarrow$ worst-case rank $= O(\sqrt{m})$

Polynomial opt $\longrightarrow$ QCQP

$\downarrow$

non-rank-1 solution $\longleftarrow$ SDP relaxation

$\downarrow$

rounding: round it to rank $\underline{1}$

Find a sub-optimal solution (approximation ratio)

special case: $\min \; x^T M x$

$\quad\quad\quad\quad\quad$ s.t $x_i^2 = 1 \quad i = 1, \dots, n$

$\Longleftrightarrow \max \; x^T(-M)x \quad$ s.t. $\; x_i^2 = 1 \quad i = 1, \dots, n$

$\Longleftrightarrow \max \; x^T(-M + \alpha I)x \quad$ s.t. $\; x_i^2 = 1 \quad i = 1, \dots, n$

$= \alpha$ over feasible set

Design α s.t. $\underbrace{-M + \alpha I}_{Q} \succeq 0$

$$\Longrightarrow \quad \max_{x} \; x^T Q x \quad \text{s.t.} \quad x_i^2 = 1 \quad i = 1, \dots, n$$

$\underbrace{\text{maximize}}$ a convex quadratic objective over a discrete set.

$\boxed{\begin{array}{l} X = xx^T \Longrightarrow \\ X \succeq 0, \; \cancel{\operatorname{rank}(X) = 1} \end{array}}$

SDP relaxation

$$\Longrightarrow \quad \max_{X \succeq 0} \; \langle Q, X \rangle$$

$$\text{s.t.} \quad X_{ii} = 1 \quad i = 1, \dots, n$$

SDP

Nestrov:

$$\frac{2}{\pi} f_{SDP} \le f_{\text{original}} \le f_{SDP} \Longrightarrow$$

optimal obj of SDP

0.6366

SDP: suboptimal by at most 37% $\longrightarrow$ approximation ratio

and can find a suboptimal $\hat{x}_1, \dots, \hat{x}_n \in \{-1, 1\}$