总成绩 (100%) = 考勤 (10%) + 习题 (30%) + 测验 (60%)，习题每个点至少完成一半

课程概况：

- 绪论 2h
- 凸集和凸函数 3h
- 凸优化问题 6h
- 拉格朗日乘子 5h
- 凸优化应用 6h
- 无约束凸优化问题求解 5h
- 有约束凸优化问题求解 4h
- 课程测试 2h

1 绪论

Remark 1.1. f 是凸函数，$\nabla f_x = 0 \Leftrightarrow$ 点 x 是最小值点

2 凸集和凸函数

Remark 2.1. 凸集：$\forall x, y \in C, \theta \in [0,1] \Longrightarrow \theta x + (1-\theta)y \in C$。空集、点、线段都是凸集。

Remark 2.2. 凸集的例子：

- 超平面：$\{x : a^t x = b\}$，$a \in \mathbb{R}^n - \{0\}$ 是法向量
- 半平面：$\{x : a^t x \le b\}$，$a \in \mathbb{R}^n - \{0\}$
- 欧几里得球：$B(x_c, r) = \{x : \|x - x_c\| < r\} = \{x_c + ru : \|u\| < 1\}$
- 椭球：$\{x : (x-x_c)^t P^{-1}(x-x_c) < 1\}, P \in S^n_{++}$ 或 $\{x_c + Au : \|u\| < 1\}, A \in S^n_{++}$
- 多面体：$\{x : Ax \preceq b, Cx = d\}, A \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{p\times n}, b \in \mathbb{R}^m, d \in \mathbb{R}^p$，多面体是半空间和超平面的交集，任意个数的凸集的交集是凸集。

Remark 2.3. 凸集 S：是 S 中所有点的凸组合的最小集合

- 如果 C 是一个凸集，则 $aC + b = \{ax + b : x \in C\}, a \in \mathbb{R}, b \in \mathbb{R}^n$ 也是一个凸集
- 对于仿射函数 $f(x) = Ax + b, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$

  - 凸集在 f 下的像是凸集

$$S \subset \mathbb{R}^n \text{ convex } \Longrightarrow f(S) = \{f(x) : x \in S\} \subset \mathbb{R}^m \text{ convex}$$

  - 凸集在 f 下的逆像是凸集

$$C \subset \mathbb{R}^m \text{ convex } \Longrightarrow f^{-1}(C) = \{x : f(x) \in C\} \subset \mathbb{R}^n \text{ convex}$$

- 两个凸集可以用一个超平面分离（证明困难）
- 支撑超平面：$\{x : a^t x = a^t x_0\}, a \in \mathbb{R}^n - \{0\}, a^t x \leq a^t x_0, \forall x \in C$，其中 C 是一个凸集，x_0 是凸集上一边界点
- 支撑超平面定理：如果 C 是凸的，则在 C 的每个边界点上都存在一个支撑超平面

Remark 2.4. 凸函数：$f : \mathbb{R}^n \to \mathbb{R}$ 的定义域 $\mathrm{dom}(f)$ 是一个凸集，满足 $\forall x, y \in \mathrm{dom}(f), \theta \in [0, 1]$

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

f 是一个凸函数，则 f 在定义域的任何内点都是连续的，并且 f 是局部有界的：$\exists B(x, r) \subset \mathrm{dom}(f)$

$$z = \theta x + (1-\theta)y \Longrightarrow \frac{f(z) - f(x)}{1-\theta} \leq \frac{f(y) - f(z)}{\theta} \Longrightarrow \frac{f(z) - f(x)}{\|z - x\|} \leq \frac{f(y) - f(z)}{\|y - z\|}$$

Remark 2.5. 一些凸函数的例子：

- 仿射函数：$a^t x + b, \forall a \in \mathbb{R}^n, b \in \mathbb{R}$
- 幂函数 x^α on $\mathbb{R}_{++} = (0, \infty), \alpha \geq 1$ or $\alpha \leq 0$
- l^p 范数：$\|x\|_p = \sum_{i=1}^n (|x_i|^p)^{\frac{1}{p}}$ on $\mathbb{R}^n$ for $p \geq 1$

Remark 2.6. 凸函数的性质

- f is convex $\Longrightarrow \alpha f$ for $\alpha > 0$ is convex.
- $f_1, \ldots, f_m$ are convex $\Longrightarrow f_1 + \cdots + f_m$ is convex.
- Composition with affine function: f is convex $\Longrightarrow f(Ax + b)$ is convex.
- $f_1, \ldots, f_m$ are convex $\Longrightarrow f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex.
  - 分段线性函数 (piecewise-linear function): $f(x) = \max_{1 \leq i \leq m}\{a_i^t x + b_i\}$

Remark 2.7. 严格凸函数 strictly convex function，$f : \mathbb{R}^n \to \mathbb{R}$

- dom(f) is convex set
- $\forall x \neq y \in \text{dom}(f), \theta \in (0,1)$

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

l^p-norm and l^∞-norm are not strictly convex.

Remark 2.8. 强凸函数 strongly convex function，$f : \mathbb{R}^n \to \mathbb{R}$

- dom(f) is convex set
- $\exists m > 0$, satisfy $f(x) - \frac{m}{2}\|x\|^2$ is convex.
- 判定方法：$\nabla^2 f \succ 0$
- 谱分解：$A = P^t Q P$，其中 Q 的对角线是 A 的特征值
- $\nabla^2 f - mI \succeq 0 \Longrightarrow u^t(\nabla^2 f - mI)u \geq 0 \Longrightarrow u^t \nabla^2 f u \geq m\|u\|^2 \Longrightarrow u^t P^t Q P u \geq m\|u\|^2 \Longrightarrow \forall \lambda \geq m, \lambda$ 是 $\nabla^2 f(x)$ 的特征值
- $f(x)$ is convex $\Longrightarrow f(x) + \frac{m}{2}\|x\|^2$ is strictly convex.

strongly convex $\Longrightarrow$ strictly convex $\Longrightarrow$ convex

Remark 2.9. 凸函数判定：

- First-order condition: **differentiable** f with **convex domain** is convex iff $\forall x, y \in$ dom(f)

$$f(y) \geq f(x) + \nabla f(x)^t (y - x)$$

  - *Proof.*
    * f is convex $\Longleftrightarrow \frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(z)}{y-z} \Longrightarrow$ with $z \to x$, $\frac{f(y)-f(x)}{y-x} \geq f'(x) \Longrightarrow f(y) \geq f(x) + f'(x)(y-x)$
    * $f(y) \geq f(x) + f'(x)(y-x) \Longrightarrow \frac{f(y)-f(z)}{y-z} \geq f'(z) \geq \frac{f(x)-f(z)}{x-z} \Longrightarrow$ by $z = \theta x + (1-\theta)y, f$ is convex

  $f(x^*) = 0 \Longleftrightarrow x^*$ is global minimum of f.

  f is strictly convex $\Longleftrightarrow \forall x \neq y \in \text{dom}(f), f(y) > f(x) + \nabla f(x)^t(y-x)$

- Second-order conditions: for **twice differentiable** f with **convex domain** is convex iff $x \in$ dom(f)

$$\nabla^2 f(x) \succeq 0$$

  - if $\nabla^2 f(x) \succ 0$ for $\forall x \in \text{dom}(f) \Longrightarrow f$ is strictly convex.

  - $\exists m > 0$, satisfy $\nabla^2 f(x) \succeq mI$ for $\forall x \in \mathrm{dom}(f) \Longleftrightarrow f$ is strongly convex.

- Restriction of a convex function to a line:
  - $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff the function $g : \mathbb{R} \to \mathbb{R}$

$$g(t) = f(x + tv), \quad \mathrm{dom}(g) = \{t : x + tv \in \mathrm{dom}(f)\}$$

    is convex for any $x \in \mathrm{dom}(f)$ and $v \in \mathbb{R}^n$
  - *Proof.* $g(t) = f(x + t(y - x))$ is convex $\Longleftrightarrow g(\theta) \le \theta g(0) + (1 - \theta)g(1) \Longleftrightarrow f(\theta x + (1 - \theta y)) \le \theta f(x) + (1 - \theta)f(y) \Longleftrightarrow f$ is convex. □

Remark 2.10. $X \in \mathbb{S}^n_{++} \Longrightarrow X = P^t Q P = P^t \,\mathrm{diag}(q_1, \ldots, q_n)P \Longrightarrow X^\alpha \triangleq P^t \,\mathrm{diag}(q_1^\alpha, \ldots, q_n^\alpha)P$, satisfy $X^\alpha X^\beta = X^{\alpha+\beta}$, $X^0 = I$.

Remark 2.11.

- sublevel set(下水平集) of $f : \mathbb{R}^n \to \mathbb{R}$

$$C_\alpha = \{x \; in \, \mathrm{dom}(f) : f(x) \le \alpha\}$$

  sublevel set of convex functions are convex.

- Epigraph(上境图) set of $f : \mathbb{R}^n \to \mathbb{R}$

$$\mathrm{epi}(f) = \big\{(x, t) \in \mathbb{R}^{n+1} : x \in \mathrm{dom}(f), t \ge f(x)\big\}$$

  f is convex iff epi is convex set.

3 凸优化问题

Remark 3.1. Optimization problem:

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & f_i(x) \le 0, \quad 1 \le i \le m \\ & h_i(x) = 0, \quad 1 \le i \le p \end{cases}$$

- feasible set $X \subset \mathcal{D}$

$$\mathcal{D} = \mathrm{dom}(f) \cap \left(\cap_{i=1}^m \mathrm{dom}\,(f_i)\right) \cap \left(\cap_{i=1}^p \mathrm{dom}\,(h_i)\right)$$

- optimal value: $p^* = \inf\{f(x) : x \text{ is feasible}\}$

- a feasible x is an optimal solution(minimizer) if $f(x) = p^*$

Remark 3.2. Convex optimization problem(COP):

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & f_i(x) \leq 0, \quad 1 \leq i \leq m \\ & Ax = b \end{cases}$$

- objective function f is convex
- inequality constraints $f_1.,\ldots,f_m$ are convex
- equality constraints are affine: $A \in \mathbb{R}^{p\times n}, b \in \mathbb{R}^p$
- the feasible set X of COP is convex
  - $X = \text{dom}(f) \cap (\cap_{i=1}^m X_i) \cap \text{hyperplanes}$
    - $X_i = \{x \in \text{dom}(f_i) : f_i(x) \leq 0\}$
  - so a COP is actually an unconstrained COP defined on a convex set.
- any local minimum of a COP is globally optimal.
  - x^* is a local minimum: a solution of the COP in $B(x^*, r) \cap X$
  - $\forall y \in X$, take $\theta \to 0$, satisfy $z = \theta x^* + (1-\theta)y \in B(x^*, r)$, by convexity, $\theta f(x^*) + (1-\theta)f(y) \geq f(z) \geq f(x^*)$, thus $f(y) \geq f(x^*)$.
- the set of optimal solutions is convex.
- For **differentiable** f, $x \in X$ is optimal iff

$$\nabla f(x)^t (y - x) \geq 0, \quad \forall y \in X$$

  - if x is optimal, let $g(\theta) = f(x + \theta(y - x))$

$$0 \leq \lim_{\theta \downarrow 0} \frac{f(x + \theta(y-x)) - f(x)}{\theta} = g'(0) = \nabla f(x)^t (y - x)$$

  - conversely, by convexity(First order condition)

$$\begin{cases} f(y) \geq f(x) + \nabla f(x)^t (y - x) \\ \nabla f(x)^t (y - x) \geq 0 \end{cases} \Longrightarrow f(y) \geq f(x) \Longrightarrow x \text{ is optimal}$$

Important, for COP:

$$x \text{ optimal} \Longleftrightarrow \nabla f(x)^t (y - x) \geq 0, \forall y \in X$$

Remark 3.3. Important examples:

- Linear program(LP):

$$\begin{cases} \text{minimize} & c^t x \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{cases}$$

  - convex problem with affine object over a polyhedron
  - standard from

$$\begin{cases} \text{minimize} & c^t x \\ \text{subject to} & x \succeq 0 \\ & Ax = b \end{cases}$$

- Quadratic program(QP):

$$\begin{cases} \text{minimize} & \frac{1}{2}x^t Px + q^t x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{cases}$$

  $P \in \mathbb{S}^n_+$, convex problem with quadratic object over a polyhedron.

- Quadratically constrained quadratic program(QCQP)

$$\begin{cases} \text{minimize} & \frac{1}{2}x^t Px + q^t x + r \\ \text{subject to} & \frac{1}{2}x^t P_i x + q_i^t x + r \preceq 0, \quad 1 \leq i \leq m \\ & Ax = b \end{cases}$$

  - $P, P_i \in \mathbb{S}^n_+$, objective and constraints are convex quadratic.
  - if $P_1, \ldots, P_m \in \mathbb{S}^n_{++}$, feasible region is intersection of m ellipsoids andan affine set.

Remark 3.4. Unconstrained COP, with differentiable f

$$\text{minimize}\, f(x)$$

- $x \in \operatorname{dom}(f)$ (open set!) is optimal iff $\nabla f(x) = 0$
  - if x is optimal, $\nabla f(x)^t(y-x) \geq 0$ for any feasible y, take $y = x - \lambda\nabla f(x)$ for sufficient small $\lambda > 0$, thus $\nabla f(x) = 0$
  - conversely, $\nabla f(x)^t(y-x) = 0$

- Intuitive interpretation: x is optimal, then $\langle \nabla f(x), y-x\rangle \geq 0$. if $\nabla f(x) \neq 0$, $\exists y$ satisfy $\langle \nabla f(x), y-x\rangle < 0$, so $\nabla f(x) = 0$.

Remark 3.5. Equality constrained COP, with differentiable f, $A \in \mathbb{R}^{p\times n}, b \in \mathbb{R}^p$

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{cases}$$

- $x \in \mathrm{dom}\,(f)$ is optimal iff:

$$Ax = b, \text{ and there exists } v \in \mathbb{R}^p, \text{ s.t. } \nabla f(x) = A^t v, \nabla f(x) \in \mathcal{R}(A^t)$$

Proof. x optimal $\Longleftrightarrow Ax = b, \nabla f(x) = A^t v$.

- "$\Longleftarrow$": $\forall y \in X$, $Ay = b \Longrightarrow (\nabla f(x))^t(y-x) = v^t A(y-x) = v^t(b-b) = 0 \Longrightarrow x$ is optimal.
- $\mathcal{N}(A) = \mathcal{R}(A^t)^\perp, \mathcal{N}(A)^\perp = \mathcal{R}(A^\perp)$
- "$\Longrightarrow$": $\forall u \in \mathcal{N}(A)$, $x + \theta u \in \mathrm{dom}\,(f)$ for $\theta \to 0$. make $y = x + \theta u$, $\nabla f(x)(y-x) \geq 0 \Longrightarrow \theta\langle \nabla f(x), u\rangle \geq 0$ satisfy for all $\theta \to 0$, so $\langle \nabla f(x), u\rangle = 0$. As a result, $\nabla f(x)^t \in \mathcal{N}(A)^\perp$, then $\exists v \in \mathbb{R}^p$ s.t. $\nabla f(x) = A^t v$.

□

Remark 3.6. Equality constrained QP: $P \in \mathbb{S}^n_+, q \in \mathbb{R}^n, r \in \mathbb{R}, A \in \mathbb{R}^{p\times n}$ with $\mathrm{rank}(A) = p$, $b \in \mathbb{R}^p$

$$\begin{cases} \text{minimize} & \frac{1}{2}x^t P x + q^t x + r \\ \text{subject to} & Ax = b \end{cases}$$

- x^* is optimal $\Longleftrightarrow \exists v^* \in \mathbb{R}^p$ s.t.

$$\begin{bmatrix} P & A^t \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

- coefficient matrix is called KKT matrix.
- KKT matrix is nonsingular $\Longleftrightarrow$ "$Ax = 0, x \neq 0 \Longrightarrow x^t P x > 0$"

Remark 3.7.

$$\begin{cases} \text{minimize}_x & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad 1 \leq i \leq m \end{cases} \Longleftrightarrow \begin{cases} \text{minimize}_{(x,y)} & f(x) \\ \text{subject to} & g_i(x) + y_i^2 = 0, \quad 1 \leq i \leq m \end{cases}$$

- $L(x,y,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) + y_i^2), \quad \partial_x L = \partial_y L = 0$
- $\begin{cases} \nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0 \\ \lambda_i y_i = 0 \end{cases} \Longrightarrow \begin{cases} \lambda_i = 0 \\ y_i = 0 \end{cases} \Longrightarrow \lambda_i g_i(x) = 0$

Remark 3.8. Inequality constrained COP, with differentiable $f, f_1, \ldots, f_m$

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & f_i(x) \le 0, \quad 1 \le i \le m \end{cases}$$

- sufficient condition: for a feasible x, if exists $\lambda_i \ge 0$ for $i \in [1,m]$ and $\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0$, $\lambda_1 f_1(x) = \ldots = \lambda_m f_m(x) = 0$, x is optimal.

  Proof.

  - $f_i(x) \neq 0 \Longrightarrow \lambda_i = 0$
  - $f_i(x) = 0 \Longrightarrow \nabla f_i(x)^t (y - x) \le 0$ for any feasible y
    - if $f(x_i) = 0$, then $\forall y \in X \Longrightarrow f_i(y) \le f_i(x) \Longrightarrow \nabla f_i^t(x)(y-x) \le 0$
    - no more proof...
  - for any feasible y, $\nabla f(x)^t (y-x) = -\sum_{i=1}^{m} \lambda_i \nabla f_i(x)^t (y-x) \ge 0$
  - x is optimal.

  ∎

- the converse is false.

Remark 3.9. COP over nonnegative orthant, with differentiable f

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & x \succeq 0 \end{cases}$$

- $x \in \mathrm{dom}(f)$ is optimal iff

$$x \succeq 0, \qquad \begin{cases} \nabla f(x)_i \ge 0 & \text{if } x_i = 0 \\ \nabla f(x)_i = 0 & \text{if } x_i > 0 \end{cases}$$

- *Proof. by observe.*
  - x is optimal $\Longrightarrow \nabla f(x)^t (y-x) \ge 0$ holds **for all feasible** y
  - for $x_i > 0 \Longrightarrow y_i - x_i$ can be positive or negative, so $\nabla f(x)_i = 0$

- for $x_i = 0 \Longrightarrow y_i - x_i \geq 0, \nabla f(x)_i \geq 0$

$\square$

Remark 3.10.

$$\lim_{\varepsilon \to 0} \frac{f(x + a\varepsilon) - f(x)}{\varepsilon} = \langle \nabla f(x), a \rangle, \qquad \varepsilon \to x + a\varepsilon \to f(x + a\varepsilon)$$

Remark 3.11. COP over a simplex, with differentiable f

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & x \succeq 0, \sum_{i=1}^n x_i = 1 \end{cases}$$

- $x \in \mathrm{dom}(f)$ is optimal iff

$$\partial_j f(x) \geq \partial_i f(x) \text{ for all } 1 \leq j \leq n \text{ when } x_i > 0$$

- *Proof.* by observe
  - if x is optimal, $\nabla f(x)^t (y - x) \geq 0$ holds for all feasible y
  - for $x_i > 0 \Longrightarrow y_i - x_i$ can be positive or negative, so $\partial_j f(x) \geq \partial_i f(x) = 0$ for any j
  - $\partial_i f(x)$ is a constant C for all $x_i > 0$, and $\partial_j f(x) \geq C$ for all $x_j = 0$.
  - the converse is obvious

$$\nabla f(x)^t (y - x) = \sum_{x_i > 0} \partial_i f(x)(y_i - x_i) + \sum_{x_i = 0} \partial_i f(x)(y_i - x_i) \geq C \sum_{i=1}^n (y_i - x_i) = 0$$

$\square$

4 拉格朗日乘子

Remark 4.1. The Lagrange multiplier only tells the properties satisfied by the solution. The solution can not always be obtained by Lagrange multiplier.

Remark 4.2. Standard form optimization problem (not necessarily convex)

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & f_i(x) \leq 0, \quad 1 \leq i \leq m \\ & h_i(x) = 0, \quad 1 \leq i \leq p \end{cases}$$

$x \in \mathcal{D} = \mathrm{dom}(f) \cap (\cap_{i=1}^m \mathrm{dom}\,(f_i)) \cap (\cap_{i=1}^p \mathrm{dom}\,(h_i))$, optimal value denoted p^*. $x \in \mathcal{D}$ **do not need to satisfy constraints.**

- Lagrange function, $L : \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R}$

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

- Lagrange dual function, $g : \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R}$

$$g(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu)$$

  - g is concave, can be $-\infty$ for some (λ, μ)
  - lower bound property: if $\lambda \succeq 0$, then $g(\lambda, \mu) \leq p^*$
    * $L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) \leq f(x)$
    * $g(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu) \leq L(x^*, \lambda, \mu) \leq f(x^*) = p^*$.
    * $\max g(\lambda, \mu) = d^*$ with $\lambda \succeq 0$, get $d^* \leq p^*$.

Remark 4.3. Equality constrained norm minimization, with any norm $\|\cdot\|$ of $\mathbb{R}^n$

$$\begin{cases} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{cases}$$

- Lagrangian: $L(x, \mu) = \|x\| + \mu^t(Ax - b)$
- dual function:

$$\begin{aligned} g(\mu) = \inf_{x \in \mathcal{D}} \left(\|x\| + \mu^t Ax - b^t \mu\right) &= \inf_{x \in \mathcal{D}} \left(\|x\|(\mathbf{1} + \mu^t A \frac{x}{\|x\|}) - b^t \mu\right) \\ &= \begin{cases} -b^t \mu, & \text{if } \|A^t \mu\|_* \leq 1 \\ -\infty, & \text{otherwise} \end{cases} \end{aligned}$$

  - $\frac{x}{\|x\|}$ is vector of norm 1.
  - $\|v\|_* = \sup_{\|u\| \leq 1} u^t v$ is the dual norm of $\|\cdot\|$

- lower bound property: $p^* \geq -b^t \mu$ if $\|A^t u\|_* \leq 1$.

Remark 4.4. Standard form LP:

$$\begin{cases} \text{minimize} & c^t x \\ \text{subject to} & x \succeq 0 \\ & Ax = b \end{cases}$$

- Lagrangian:

$$L(x, \lambda, \mu) = c^t x - \lambda^t x + \mu^t (Ax - b) = -\mu^t b + (c + A^t \mu - \lambda)^t x$$

- dual function:

$$\begin{aligned} g(\lambda,\mu) &= \inf_{x\in\mathcal{D}}\left(-\mu^t b + (c + A^t\mu - \lambda)^t x\right) \\ &= \begin{cases} -\mu^t b, & \text{if } c + A^t\mu - \lambda = 0 \\ -\infty, & \text{otherwise} \end{cases} \end{aligned}$$

- lower bound property: $p^* \geq -\mu^t b$ if $c + A^t\mu \succeq 0$.

Remark 4.5. Tow-way partitioning, for $W \in \mathbb{S}^n$

$$\begin{cases} \text{minimize} & x^t W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1,\dots,n \end{cases}$$

- a nonconvex problem, feasible set contains 2^n discrete points
- Lagrangian: $L(x,\mu) = f(x) + \sum_{i=1}^n \mu_i(x_i^2 - 1)$
- dual function:

$$\begin{aligned} g(\mu) &= \inf_{x\in\mathcal{D}}\ (x^t W x + \sum_{i=1}^n \mu_i(x_i^2 - 1)) \\ &= \inf_{x\in\mathcal{D}}\ x^t(W + \operatorname{diag}(\mu))x - \sum_{i=1}^n \mu_i \\ &= \begin{cases} -\sum_{i=1}^n \mu_i, & \text{if } W + \operatorname{diag}(\mu) \succeq 0 \\ -\infty, & \text{otherwise} \end{cases} \end{aligned}$$

- lower bound property: $p^* \geq -\sum_{i=1}^n \mu_i$ if $W + \operatorname{diag}(\mu) \succeq 0$.
  - $p^* \geq n\lambda_{\min}(W)$, where $\lambda_{min}(W)$ is the smallest eigenvalue of W.
  - *Proof.* $W + \operatorname{diag}(\mu) = P^t Q P + \theta I = P^t(Q + \theta I)P \succeq 0$, let $\operatorname{diag}(\mu) = \theta I$, so $\lambda_i + \theta \geq 0, \theta = -\lambda_{\min}(W)$, $p^* \geq n\lambda_{\min}(W)$. □

Remark 4.6. Lagrange dual problem

$$\begin{cases} \text{maximize} & g(\lambda,\mu) \\ \text{subject to} & \lambda \succeq 0 \end{cases}$$

- COP, optimal value denoted d^*
- finds best lower bound on $p*$, obtained from Lagrange dual function.
- week duality: $d^* \leq p^*$

    - always holds (for both convex and nonconvex problems)
    - can be used to find nontrivial lower bounds for difficult problems

- strong duality: $d^* = p^*$
    - does not hold in general, but usually holds for COP
    - an example that the strong duality does not hold

$$\begin{cases} \text{minimize} & e^{-x} \\ \text{subject to} & x^2/y \leq 0 \end{cases}$$

        * $\mathcal{D} = \{(x, y) : x \in \mathbb{R}, y > 0\}, g(\lambda) = 0 \Longrightarrow d^* = 0 < p^* = 1$

5 凸优化应用

6 无约束凸优化问题求解

7 有约束凸优化问题求解