# DATA130026.01 Optimization Assignment 8

**Due Time:** at the beginning of the class, May 15, 2024
**姓名：**雍崔扬
**学号：**21307140051

## Problem 1

Consider the Semi-Definite Programming (SDP):

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & \begin{bmatrix} D & Ax \\ (Ax)^T & t \end{bmatrix} \succeq 0 \end{aligned}$$

Suppose that $D$ is an $n \times n$ diagonal matrix with positive entries in the diagonal positions, and $A \in \mathbb{R}^{n \times m}$ is given.
Rewrite the above SDP as an SOCP (Second-Order Cone Programming).

**Solution:**

根据 Schur 补的结论 (参考 [FDU 高等线性代数 专题D. Schur补 https://zhuanlan.zhihu.com/p/680073237](https://zhuanlan.zhihu.com/p/680073237))
由题设可知 $D \succ 0$，
在此条件下，我们有 $\begin{bmatrix} D & Ax \\ (Ax)^T & t \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad s = t - (Ax)^T D^{-1} Ax \geq 0$
其中 $s$ 是矩阵 $\begin{bmatrix} D & Ax \\ (Ax)^T & t \end{bmatrix}$ 的 $(2,2)$ 分块 $t$ 的 Schur 补
(由于在本题的假设下，它是 $1 \times 1$ 的标量，故用小写的 $s$ 表示)
因此**半定锥约束** (SDPC) $\begin{bmatrix} D & Ax \\ (Ax)^T & t \end{bmatrix} \succeq 0$ 等价于 $\|D^{-\frac{1}{2}}Ax\|_2^2 \leq t$

要使得约束成立，必然有 $t \geq 0$
令 $w = \sqrt{t}$ 即等价于**二阶锥约束** (SOCC) $\|D^{-\frac{1}{2}}Ax\|_2 \leq w$.
问题转化为 SOCP 形式：

$$\begin{aligned} \min_{x,w} \quad & w^2 \\ \text{s.t.} \quad & \|D^{-\frac{1}{2}}Ax\|_2 \leq w \end{aligned}$$

## Problem 2

**(a)** Find a dual problem to the following QCQP (Quadratically Constrained Quadratic Program):

$$(\text{P}) \begin{array}{ll} \min & x^T A_0 x + 2b_0^T x + c_0 \\ \text{s.t.} & x^T A_i x + 2b_i^T x + c_i \leq 0, \quad i = 1, \ldots, m \end{array}$$

where $A_i \succ 0$ is an $n \times n$ matrix, $b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, $i = 0, 1, \ldots, m$.

- **Lemma:**
  设 $f(x) = x^T Ax + 2b^T x + c$,
  其中 $A \in \mathbb{R}^{n \times n}$ 是对称阵，$b \in \mathbb{R}^n$ 且 $c \in \mathbb{R}$.
  则以下两个命题是等价的：
  - $f(x) = x^T Ax + 2b^T x + c \geq 0$ 对于任意 $x \in \mathbb{R}^n$ 都成立.
  - $\begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \succeq 0$

- **Solution:**
  Lagrange 函数为:
  $$L(x,\lambda) = x^TA_0x + 2b_0^Tx + c_0 - \sum_{i=1}^m (x^TA_ix + 2b_i^Tx + c_i)$$
  $$= x^T(A_0 + \sum_{i=1}^m \lambda_iA_i)x + 2(b_0 + \sum_{i=1}^m \lambda_ib_i)^Tx + \sum_{i=1}^m \lambda_ic_i$$
  Lagrange 对偶函数为:
  $$\begin{aligned} d(\lambda) &= \inf_{x\in\mathbb{R}^n} L(x,\lambda) \\ &= \inf_{x\in\mathbb{R}^n} \{x^T(A_0 + \sum_{i=1}^m \lambda_iA_i)x + 2(b_0 + \sum_{i=1}^m \lambda_ib_i)^Tx + \sum_{i=1}^m \lambda_ic_i\} \\ &= \sup_{\tau\in\mathbb{R}} \{\tau : \tau \le x^T(A_0 + \sum_{i=1}^m \lambda_iA_i)x + 2(b_0 + \sum_{i=1}^m \lambda_ib_i)^Tx + \sum_{i=1}^m \lambda_ic_i \ \ (\forall\, x\in\mathbb{R}^n)\} \\ &= \sup_{\tau\in\mathbb{R}} \{\tau : \begin{bmatrix} A_0 + \sum\limits_{i=1}^m \lambda_iA_i & b_0 + \sum\limits_{i=1}^m \lambda_ib_i \\ (b_0 + \sum\limits_{i=1}^m \lambda_ib_i)^T & \sum\limits_{i=1}^m \lambda_ic_i - \tau \end{bmatrix} \succeq 0\} \qquad (\text{应用引理}) \end{aligned}$$
  于是 Lagrange 对偶问题的形式为:
  $$\begin{aligned} \sup_{\lambda\in\mathbb{R}_+^m} d(\lambda) &= \sup_{\lambda\in\mathbb{R}_+^m,\tau\in\mathbb{R}} \{\tau : \begin{bmatrix} A_0 + \sum\limits_{i=1}^m \lambda_iA_i & b_0 + \sum\limits_{i=1}^m \lambda_ib_i \\ (b_0 + \sum\limits_{i=1}^m \lambda_ib_i)^T & \sum\limits_{i=1}^m \lambda_ic_i - \tau \end{bmatrix} \succeq 0\} \\ &= \begin{array}{ll} \max & \tau \\ \text{s.t.} & \lambda \succeq 0_m \\ & \begin{bmatrix} A_0 + \sum\limits_{i=1}^m \lambda_iA_i & b_0 + \sum\limits_{i=1}^m \lambda_ib_i \\ (b_0 + \sum\limits_{i=1}^m \lambda_ib_i)^T & \sum\limits_{i=1}^m \lambda_ic_i - \tau \end{bmatrix} \succeq 0 \end{array} \end{aligned}$$

**(b)** Write down the semidefinite relaxation (SDR) for the QCQP problem in **(a)**.
Write the Lagrangian dual problem for the SDR
and show that it is equivalent to the Lagrangian dual problem for the original QCQP.

**Solution:**

- **① 推导 QCQP 问题的半定松弛形式 (SDR):**
  对于任意 $A\in\mathbb{S}^n$ 我们都有 $x^TAx = \operatorname{tr}(x^TAx) = \operatorname{tr}(Axx^T) = A\bullet xx^T$
  其中 $\bullet$ 代表**矩阵内积**.
  于是 QCQP 问题 $(\mathrm{P})\ \begin{array}{ll} \min & x^TA_0x + 2b_0^Tx + c_0 \\ \text{s.t.} & x^TA_ix + 2b_i^Tx + c_i \le 0, \quad i = 1,\dots,m \end{array}$
  可以表示为:
  $$\begin{array}{ll} \min & A_0\bullet xx^T + 2b_0^Tx + c_0 \\ \text{s.t.} & A_i\bullet xx^T + 2b_i^Tx + c_i \le 0, \quad i = 1,\dots,m \end{array} \tag{1}$$
  我们很容易验证 $X = xx^T \quad\Leftrightarrow\quad \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0,\ \operatorname{rank}\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \le 1$
  - 右侧命题的必要性 "$\Rightarrow$" 显然成立，我们只需验证充分性 "$\Leftarrow$":
    根据 $\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0,\ \operatorname{rank}\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \le 1$
    我们知道 $\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix}$ 可以表示为某一向量 $v = \begin{bmatrix} y \\ z \end{bmatrix}$ 的外积 $vv^T$
    其中 $y\in\mathbb{R}^n,\ z\in\mathbb{R}$.
    因此 $\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} = vv^T = \begin{bmatrix} yy^T & yz \\ zy^T & z^2 \end{bmatrix}$

故我们有 $\begin{cases} X = yy^T \\ x = yz \\ 1 = z^2 \end{cases}$

解得 $\begin{cases} X = xx^T \\ y = \pm x \\ z = \pm 1 \end{cases}$，于是就得到 $X = xx^T$，从而验证了充分性 "$\Leftarrow$" 成立.

注意到问题 $(1)$ 等价于:

$$\begin{array}{ll} \min & A_0 \bullet X + 2b_0^T x + c_0 \\ \text{s.t.} & A_i \bullet X + 2b_i^T x + c_i \le 0, \quad i = 1, \dots, m \\ & X = xx^T \end{array}$$

进而等价于:

$$\begin{array}{ll} \min & A_0 \bullet X + 2b_0^T x + c_0 \\ \text{s.t.} & A_i \bullet X + 2b_i^T x + c_i \le 0, \quad i = 1, \dots, m \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0, \ \text{rank} \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \le 1 \end{array} \tag{2}$$

由于秩约束 $\text{rank} \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \le 1$ 是严格非凸的，
我们可以直接舍弃这个秩约束，就得到了问题的半定松弛 (SDR) 形式:

$$\begin{array}{ll} \min & A_0 \bullet X + 2b_0^T x + c_0 \\ \text{s.t.} & A_i \bullet X + 2b_i^T x + c_i \le 0, \quad i = 1, \dots, m \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0 \end{array} \tag{SDR}$$

- **② 推导 $(\text{SDR})$ 形式的对偶:**
  Lagrange 函数为:

$$\begin{aligned} L(X, x, \lambda, Z, z, \tau) &= A_0 \bullet X + 2b_0^T x + c_0 + \sum_{i=1}^m \lambda_i (A_i \bullet X + 2b_i^T x + c_i) - \begin{bmatrix} Z & z \\ z^T & \tau \end{bmatrix} \bullet \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \\ &= (A_0 + \sum_{i=1}^m \lambda_i A_i - Z) \bullet X + 2(b_0 + \sum_{i=1}^m \lambda_i b_i - z)^T x + (\sum_{i=1}^m \lambda_i c_i - \tau) \end{aligned}$$

  Lagrange 对偶函数为:

$$\begin{aligned} d(\lambda, Z, z, \tau) &= \inf_{X,x} L(X, x, \lambda, Z, z, \tau) \\ &= \inf_{X,x} \{ (A_0 + \sum_{i=1}^m \lambda_i A_i - Z) \bullet X + 2(b_0 + \sum_{i=1}^m \lambda_i b_i - z)^T x + (\sum_{i=1}^m \lambda_i c_i - \tau) \} \\ &= \begin{cases} \sum\limits_{i=1}^m \lambda_i c_i - \tau, & \text{if } \begin{cases} A_0 + \sum\limits_{i=1}^m \lambda_i A_i - Z = 0 \\ b_0 + \sum\limits_{i=1}^m \lambda_i b_i - z = 0 \end{cases} \\ -\infty, & \text{otherwise} \end{cases} \end{aligned}$$

  因此 Lagrange 对偶问题的形式为:

$$\begin{array}{ll} \sup & d(\lambda, Z, z, \tau) \\ \text{s.t.} & \lambda \succeq 0_m \\ & \begin{bmatrix} Z & z \\ z^T & \tau \end{bmatrix} \succeq 0 \end{array} \Leftrightarrow \begin{array}{ll} \sup & \sum\limits_{i=1}^m \lambda_i c_i - \tau \\ \text{s.t.} & \lambda \succeq 0_m \\ & \begin{bmatrix} Z & z \\ z^T & \tau \end{bmatrix} \succeq 0 \\ & A_0 + \sum\limits_{i=1}^m \lambda_i A_i - Z = 0 \\ & b_0 + \sum\limits_{i=1}^m \lambda_i b_i - z = 0 \end{array}$$

  上述问题等价于:

$$
\begin{aligned}
\max \quad & \sum_{i=1}^m \lambda_i c_i - \tau \\
\text{s.t.} \quad & \lambda \succeq 0_m \\
& \begin{bmatrix} A_0 + \sum\limits_{i=1}^m \lambda_i A_i & b_0 + \sum\limits_{i=1}^m \lambda_i b_i \\ (b_0 + \sum\limits_{i=1}^m \lambda_i b_i)^T & \tau \end{bmatrix} \succeq 0
\end{aligned}
$$

令 $w = \sum\limits_{i=1}^m \lambda_i c_i - \tau$ 即有 **(a)** 中 Lagrange 对偶问题的形式：

$$
\begin{aligned}
\max \quad & w \\
\text{s.t.} \quad & \lambda \succeq 0_m \\
& \begin{bmatrix} A_0 + \sum\limits_{i=1}^m \lambda_i A_i & b_0 + \sum\limits_{i=1}^m \lambda_i b_i \\ (b_0 + \sum\limits_{i=1}^m \lambda_i b_i)^T & \sum\limits_{i=1}^m \lambda_i c_i - w \end{bmatrix} \succeq 0
\end{aligned}
$$

## Problem 3

Show that the dual of the SOCP (Second-Order Cone Programming):

$$
(\mathrm{P}) \begin{aligned} \min \quad & f^T x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \ldots, m \end{aligned}
$$

with variables $x \in \mathbb{R}^n$, can be expressed as:

$$
(\mathrm{D}) \begin{aligned} \min \quad & \sum_{i=1}^m (b_i^T u_i - d_i v_i) \\ \text{s.t.} \quad & \sum_{i=1}^m (A_i^T u_i - c_i v_i) + f = 0_n \\ & \|u_i\|_2 \le v_i, \quad i = 1, \ldots, m \end{aligned}
$$

with variables $u_i \in \mathbb{R}^{n_i}$, $v_i \in \mathbb{R}$, $i = 1, \ldots, m$.

The problem data are $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{n_i \times n}$, $b_i \in \mathbb{R}^{n_i}$, $c_i \in \mathbb{R}^n$ and $d_i \in \mathbb{R}$, $i = 1, \ldots, m$.

Derive the dual in the following two ways:

**(a)** Introduce new variables $\begin{cases} y_i \in \mathbb{R}^{n_i} \\ t_i \in \mathbb{R} \end{cases}$ and equalities $\begin{cases} y_i = A_i x + b_i \\ t_i = c_i^T x + d_i \end{cases}$, and derive the Lagrange dual.

**Solution:**

我们引入新变量 $\begin{cases} y_i \in \mathbb{R}^{n_i} \\ t_i \in \mathbb{R} \end{cases}$ 和新的等式约束 $\begin{cases} y_i = A_i x + b_i \\ t_i = c_i^T x + d_i \end{cases}$ 得到：

$$
(\mathrm{NewP}) \begin{aligned} \min \quad & f^T x \\ \text{s.t.} \quad & \|y_i\|_2 \le t_i, \quad & i = 1, \ldots, m \\ & y_i = A_i x + b_i, \quad & i = 1, \ldots, m \\ & t_i = c_i^T x + d_i, \quad & i = 1, \ldots, m \end{aligned}
$$

为简化记号，我们记 $\begin{cases} y = (y_1, \ldots, y_m) \\ t = (t_1, \ldots, t_m) \\ \lambda = (\lambda_1, \ldots, \lambda_m) \\ \nu = (\nu_1, \ldots, \nu_m) \\ \mu = (\mu_1, \ldots, \mu_m) \end{cases}$

$(\mathrm{NewP})$ 的 Lagrange 函数为：

$$
\begin{aligned}
L(x, y, t, \lambda, \nu, \mu) &= f^T x + \sum_{i=1}^m \lambda_i (\|y_i\|_2 - t_i) + \sum_{i=1}^m \nu_i^T (y_i - A_i x - b_i) + \sum_{i=1}^m \mu_i (t_i - c_i^T x - d_i) \\
&= (f - \sum_{i=1}^m A_i^T \nu_i - \sum_{i=1}^m \mu_i c_i)^T x + \sum_{i=1}^m (\lambda_i \|y_i\|_2 + \nu_i^T y_i) + \sum_{i=1}^m (-\lambda_i + \mu_i) t_i - \sum_{i=1}^m (b_i^T \nu_i + d_i \mu_i)
\end{aligned}
$$

- $\inf\limits_{x\in\mathbb{R}^n} L(x,y,t,\lambda,\nu,\mu)$ 有下界当且仅当 $f-\sum\limits_{i=1}^m A_i^T\nu_i-\sum\limits_{i=1}^m \mu_i c_i=0_n$
- 任意给定 $i=1,\dots,m$，考虑在 $y_i\in\mathbb{R}^{n_i}$ 上最小化 $\lambda_i\|y_i\|_2+\nu_i^Ty_i$:
    - 若 $\|\nu_i\|_2\le\lambda_i$，则根据 Cauchy-Schwarz 不等式有:
      $$\begin{aligned}\lambda_i\|y_i\|_2+\nu_i^Ty_i &\ge \lambda_i\|y_i\|_2-\|\nu_i\|_2\|y_i\|_2\\ &\ge \lambda_i\|y_i\|_2-\lambda_i\|y_i\|_2\\ &=0\end{aligned}$$
      显然当 $y_i=0_{n_i}$ 时可以取等,
      故此情况下 $\lambda_i\|y_i\|_2+\nu_i^Ty_i$ 的下确界为 $0$
    - 若 $\|\nu_i\|_2>\lambda_i$，则我们可以说明 $\lambda_i\|y_i\|_2+\nu_i^Ty_i$ 无下界:
      令 $y_i=-t\nu_i\ (t>0)$ 则有:
      $$\begin{aligned}\lim_{t\to\infty}\{\lambda_i\|y_i\|_2+\nu_i^Ty_i\} &= \lim_{t\to\infty}\{\lambda_i\|\nu_i\|_2t-\|\nu_i\|_2^2t\}\\ &=\lim_{t\to\infty}\{(\lambda_i-\|\nu_i\|_2)\|\nu_i\|_2t\}\\ &=-\infty\end{aligned}$$

  于是我们有 $\inf\limits_{y_i\in\mathbb{R}^{n_i}}\{\lambda_i\|y_i\|_2+\nu_i^Ty_i\}=\begin{cases}0, & \text{if } \|\nu_i\|_2\le\lambda_i\\ -\infty, & \text{otherwise}\end{cases}$
- 任意给定 $i=1,\dots,m$
  我们很容易看出 $\inf\limits_{t_i\in\mathbb{R}}\{(-\lambda_i+\mu_i)t_i\}=\begin{cases}0, & \text{if } \mu_i=\lambda_i\\ -\infty, & \text{otherwise}\end{cases}$

综上所述，我们有:

$$\begin{aligned}d(\lambda,\nu,\mu)&=\inf_{x,y,t}L(x,y,t,\lambda,\nu,\mu)\\ &=\begin{cases}-\sum\limits_{i=1}^m(b_i^T\nu_i+d_i\mu_i), & \text{if }\begin{cases}\sum\limits_{i=1}^m A_i^T\nu_i+\sum\limits_{i=1}^m\mu_ic_i=f\\ \|\nu_i\|_2\le\lambda_i,\ \ i=1,\dots,m\\ \mu=\lambda\end{cases}\\ -\infty, & \text{otherwise}\end{cases}\end{aligned}$$

因此 Lagrange 对偶问题的形式为:

$$\begin{aligned}\max\quad & -\sum_{i=1}^m(b_i^T\nu_i+d_i\mu_i)\\ \text{s.t.}\quad & \sum_{i=1}^m(A_i^T\nu_i+c_i\mu_i)=f\\ & \|\nu_i\|_2\le\lambda_i,\ \ i=1,\dots,m\\ & \mu=\lambda\end{aligned}$$

我们消去 $\lambda$ 并令 $\begin{cases}u_i=-\nu_i\\ v_i=\mu_i=\lambda_i\end{cases}(i=1,\dots,m)$ 即得到等价形式:

$$\text{(D)}\quad\begin{aligned}\min\quad & \sum_{i=1}^m(b_i^Tu_i-d_iv_i)\\ \text{s.t.}\quad & \sum_{i=1}^m(A_i^Tu_i-c_iv_i)+f=0_n\\ & \|u_i\|_2\le v_i,\ \ i=1,\dots,m\end{aligned}$$

**(b)** Start from the conic formulation of the SOCP and use the conic dual. Use the fact that the second-order cone is self-dual.

**Solution:**

记 $K_i=\{\begin{bmatrix}x\\ t\end{bmatrix}\in\mathbb{R}^{n_i+1}:\|x\|_2\le t\}\ \ (i=1,\dots,m)$

我们有 $K_i^*=K_i\ \ (i=1,\dots,m)$ 成立 **(二阶锥是自对偶锥)**

我们将原始问题的约束转为广义不等式形式:

$$\text{(NewP)}\quad\begin{aligned}\min\quad & f^Tx\\ \text{s.t.}\quad & \begin{bmatrix}A_ix+b_i\\ c_i^Tx+d_i\end{bmatrix}\succeq_{K_i}0,\ \ i=1,\dots,m\end{aligned}$$

(NewP) 的 Lagrange 函数为：

$$
\begin{aligned}
L(x,u,v) &= f^T x - \sum_{i=1}^m \begin{bmatrix} u_i \\ v_i \end{bmatrix}^T \begin{bmatrix} A_i x + b_i \\ c_i^T x + d_i \end{bmatrix} \\
&= (f - \sum_{i=1}^m \{A_i^T u_i + v_i c_i\})^T x - \sum_{i=1}^m (u_i^T b_i + v_i d_i)
\end{aligned}
$$

Lagrange 对偶函数为：

$$
\begin{aligned}
d(u,v) &= \inf_{x\in\mathbb{R}^n} L(x,u,v) \\
&= \inf_{x\in\mathbb{R}^n} \{(f - \sum_{i=1}^m \{A_i^T u_i + v_i c_i\})^T x - \sum_{i=1}^m (u_i^T b_i + v_i d_i)\} \\
&= \begin{cases} -\sum\limits_{i=1}^m (u_i^T b_i + v_i d_i), & \text{if } \sum\limits_{i=1}^m (A_i^T u_i + v_i c_i) = f \\ -\infty, & \text{otherwise} \end{cases}
\end{aligned}
$$

因此 Lagrange 对偶问题的形式为：

$$
\begin{aligned}
\max \quad & -\sum_{i=1}^m (b_i^T u_i + d_i v_i) \\
\text{s.t.} \quad & \sum_{i=1}^m (A_i^T u_i + v_i c_i) = f \\
& \begin{bmatrix} u_i \\ v_i \end{bmatrix} \succeq_{K_i^*} 0, \;\; i = 1, \ldots, m
\end{aligned}
$$

由于二阶锥 $K_i$ 是自对偶锥，因此 $K_i^* = K_i \;\; (i = 1, \ldots, m)$

我们就得到以下形式：

$$
\begin{aligned}
\max \quad & -\sum_{i=1}^m (b_i^T u_i + d_i v_i) \\
\text{s.t.} \quad & \sum_{i=1}^m (A_i^T u_i + v_i c_i) = f \\
& \|u_i\|_2 \le v_i, \;\; i = 1, \ldots, m
\end{aligned}
$$

用 $u_i$ 替换 $-u_i$ 就得到：

$$
(\text{D}) \quad
\begin{aligned}
\min \quad & \sum_{i=1}^m (b_i^T u_i - d_i v_i) \\
\text{s.t.} \quad & \sum_{i=1}^m (A_i^T u_i - c_i v_i) + f = 0_n \\
& \|u_i\|_2 \le v_i, \;\; i = 1, \ldots, m
\end{aligned}
$$

## Problem 4

We consider the two-way partitioning problem:

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & x^T W x \\
\text{s.t.} \quad & x_i^2 = 1, \;\; i = 1, \ldots, n
\end{aligned}
\tag{1}
$$

Note that this might not be a convex problem.

**(a)** Show that the Lagrange dual of this (nonconvex) problem is equivalent to the following SDP:

$$
\begin{aligned}
\max_{\nu\in\mathbb{R}^n} \quad & -1_n^T \nu \\
\text{s.t.} \quad & W + \text{diag}(\nu) \succeq 0
\end{aligned}
\tag{2}
$$

The optimal value of this SDP gives a lower bound on the optimal value of the partitioning problem.

- **Solution:**
  问题 $(1)$ 的 Lagrange 函数为:
  $$\begin{aligned} L(x,\nu) &= x^TWx + \sum_{i=1}^{n}\nu_i(x_i^2-1) \\ &= x^T(W+\operatorname{diag}(\nu))x - 1_n^T\nu \end{aligned}$$
  Lagrange 对偶函数为:
  $$d(\nu) = \inf_{x\in\mathbb{R}^n} L(x,\nu) = \begin{cases} -1_n^T\nu, & \text{if } W+\operatorname{diag}(\nu)\succeq 0 \\ -\infty, & \text{otherwise} \end{cases}$$
  Lagrange 对偶问题为:
  $$\begin{aligned} \max_{\nu\in\mathbb{R}^n} \quad & -1_n^T\nu \\ \text{s.t.} \quad & W+\operatorname{diag}(\nu)\succeq 0 \end{aligned}$$
  根据弱对偶性可知，对偶问题的最优解一定是原始问题最优解的下界.

**(b)** Two-way partitioning problem in matrix form.
Show that the two-way partitioning problem can be cast as:

$$\begin{aligned} \min_{X\in\mathbb{S}^n} \quad & \operatorname{tr}(WX) \\ \text{s.t.} \quad & X\succeq 0 \\ & \operatorname{rank}(X)=1 \\ & X_{ii}=1,\ \ i=1,\dots,n \end{aligned} \tag{3}$$

**Hint:** Show that if $X$ is feasible, then it has the form $X=xx^T$,
where $x\in\mathbb{R}^n$ satisfies $x_i\in\{-1,1\}$ (and vice versa).

**Solution:**

- **Proof of Hint:**
  - 若 $X$ 可行，即满足 $\begin{cases} X\succeq 0 \\ \operatorname{rank}(X)=1 \\ X_{ii}=1,\ \ i=1,\dots,n \end{cases}$
    根据 $\begin{cases} X\succeq 0 \\ \operatorname{rank}(X)=1 \end{cases}$ 可知 $X$ 一定可以写成某个向量 $x\in\mathbb{R}^n$ 的外积 $xx^T$
    再根据 $X_{ii}=1$ 可知 $(xx^T)_{ii}=x_i^2=1$,
    说明 $x_i\in\{-1,1\}\ \ (i=1,\dots,n)$
  - 若 $X=xx^T$ 且 $x\in\mathbb{R}^n$ 满足 $x_i\in\{-1,1\}\ \ (i=1,\dots,n)$,
    则显然有 $\begin{cases} X\succeq 0 \\ \operatorname{rank}(X)=1 \\ X_{ii}=1,\ \ i=1,\dots,n \end{cases}$ 成立，即 $X$ 可行.
- **Utilizing the Hint:**
  根据 $x^TWx=\operatorname{tr}(x^TWx)=\operatorname{tr}(Wxx^T)$ 可知,
  原始问题 $\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & x^TWx \\ \text{s.t.} \quad & x_i^2=1,\ \ i=1,\dots,n \end{aligned}$ 可以写成:
  $$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \operatorname{tr}(Wxx^T) \\ \text{s.t.} \quad & x_i^2=1,\ \ i=1,\dots,n \end{aligned}$$
  我们引入新变量 $X\in\mathbb{S}^n$ 和新约束 $X=xx^T$，并利用 **Hint** 即有:
  $$\begin{aligned} \min_{X\in\mathbb{S}^n} \quad & \operatorname{tr}(WX) \\ \text{s.t.} \quad & X\succeq 0 \\ & \operatorname{rank}(X)=1 \\ & X_{ii}=1,\ \ i=1,\dots,n \end{aligned}$$

**(c)** SDP relaxation of the two-way partitioning problem.
Using the formulation in part **(b)**, we can form the relaxation:

$$
\begin{aligned}
\min_{X\in\mathbb{S}^n}\quad & \mathrm{tr}(WX) \\
\mathrm{s.t.}\quad & X\succeq 0 \\
& X_{ii}=1,\ \ i=1,\dots,n
\end{aligned}
\tag{4}
$$

This problem is an SDP, and therefore can be solved efficiently.
Explain why its optimal value gives a lower bound
on the optimal value of the two-way partitioning problem $(1)$.
What can you say if an optimal point $X^\star$ for this SDP has rank one?

- **Solution:**
  根据弱对偶性可知,
  **(b)** 中的对偶问题 $(3)$ 作为原始问题 $(1)$ 的 Lagrange 对偶,
  其对偶最优值一定是原始问题最优值的下界.
  而 SDP 问题 $(4)$ 作为对偶问题 $(3)$ 的松弛形式,
  其最优值一定是对偶问题 $(3)$ 的最优值的下界,
  进而一定是原始问题最优值的下界.

**(d)** We now have two SDPs that give a lower bound
on the optimal value of the two-way partitioning problem $(1)$:
the SDP relaxation $(4)$ found in part **(c)**,
and the Lagrange dual $(2)$ of the two-way partitioning problem, given in part (a).

What is the relation between the two SDPs?
What can you say about the lower bounds found by them?
**Hint:** Relate the two SDPs via duality.

- **Solution:**
  我们的目的是寻找 SDP 问题 $(2)$ 和 $(4)$ 之间的关系.

$$
\begin{aligned}
\max_{\nu\in\mathbb{R}^n}\quad & -1_n^T\nu \\
\mathrm{s.t.}\quad & W+\mathrm{diag}(\nu)\succeq 0
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
\min_{X\in\mathbb{S}^n}\quad & \mathrm{tr}(WX) \\
\mathrm{s.t.}\quad & X\succeq 0 \\
& X_{ii}=1,\ \ i=1,\dots,n
\end{aligned}
\tag{4}
$$

  我们将问题 $(2)$ 转换为一个最小化问题:

$$
\begin{aligned}
\min_{\nu\in\mathbb{R}^n}\quad & 1_n^T\nu \\
\mathrm{s.t.}\quad & W+\mathrm{diag}(\nu)\succeq 0
\end{aligned}
\tag{2}
$$

  为半定约束 $W+\mathrm{diag}(\nu)\succeq 0$ 引入一个 Lagrange 乘子 $X\in\mathbb{S}_+^n$,我们有:

$$
\begin{aligned}
L(\nu,X) &= 1_n^T\nu - X\bullet[W+\mathrm{diag}(\nu)] \\
&= 1_n^T\nu-\mathrm{tr}\{X(W+\mathrm{diag}(\nu))\} \\
&= 1_n^T\nu-\mathrm{tr}(XW)-\sum_{i=1}^n\nu_iX_{ii} \\
&= -\mathrm{tr}(XW)+\sum_{i=1}^n\nu_i(1-X_{ii})
\end{aligned}
$$

  Lagrange 对偶函数为:

$$
\begin{aligned}
d(X) &= \inf_{\nu\in\mathbb{R}^n} L(\nu,X) \\
&= \inf_{\nu\in\mathbb{R}^n}\{-\mathrm{tr}(XW)+\sum_{i=1}^n\nu_i(1-X_{ii})\} \\
&= \begin{cases} -\mathrm{tr}(XW), & \text{if } X_{ii}=1 \text{ for all } i=1,\dots,n \\ -\infty, & \text{otherwise}\end{cases}
\end{aligned}
$$

  于是我们就得到对偶问题:

$$
\begin{aligned}
\max_{X\in\mathbb{S}^n} \quad & -\operatorname{tr}(WX) \\
\text{s.t.} \quad & X \succeq 0 \\
& X_{ii} = 1,\ \ i = 1,\dots,n
\end{aligned} \tag{5}
$$

再反转一次符号，变为最小值问题：

$$
\begin{aligned}
\min_{X\in\mathbb{S}^n} \quad & \operatorname{tr}(WX) \\
\text{s.t.} \quad & X \succeq 0 \\
& X_{ii} = 1,\ \ i = 1,\dots,n
\end{aligned} \tag{4}
$$

这就说明了 $(4)$ 是 $(2)$ 的对偶问题.
根据弱对偶性可知，$(4)$ 的最优值小于等于 $(2)$ 的最优值，
因此 $(4)$ 为原始问题 $(1)$ 提供的下界没有 $(2)$ 提供的好.

**The End**