



最优化方法

东南大学

计算机&人工智能学院

宋沫飞

songmf@seu.edu.cn



对偶



- **Lagrange**对偶问题
- 弱对偶和强对偶
- 几何解释
- 最优性条件
- 扰动和灵敏度分析
- 例子
- 广义不等式



Lagrangian





Lagrangian



- 标准形式问题（不需要为凸优化）



Lagrangian



□ 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$



Lagrangian



□ 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

- **Lagrangian**: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

- **Lagrangian:** $L : \mathbf{R}^n \times \mathbf{R}_m^m \times \mathbf{R}^p \rightarrow \mathbf{R}$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

- **Lagrangian**: $L : \mathbf{R}^n \times \mathbf{R}_m^m \times \mathbf{R}^p \rightarrow \mathbf{R}_p$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

○ 其定义域为 $\mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

- **Lagrangian**: $L : \mathbf{R}^n \times \mathbf{R}_m^m \times \mathbf{R}^p \rightarrow \mathbf{R}_p$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

- 其定义域为 $\mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$
- 目标函数和约束函数的加权和



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

- **Lagrangian**: $L : \mathbf{R}^n \times \mathbf{R}_m^m \times \mathbf{R}^p \rightarrow \mathbf{R}_p$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

- 其定义域为 $\mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$
- 目标函数和约束函数的加权和
- λ_i 为第 i 个不等式约束的 **Lagrange** 乘子



Lagrangian



- 标准形式问题（不需要为凸优化）

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

○ 其中，优化变量 $x \in \mathbf{R}^n$ ，定义域为 \mathcal{D} ，最优值为 p^*

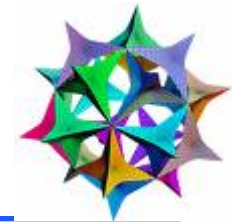
- **Lagrangian**: $L : \mathbf{R}^n \times \mathbf{R}_m^m \times \mathbf{R}_p^p \rightarrow \mathbf{R}$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

- 其定义域为 $\mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$
- 目标函数和约束函数的加权和
- λ_i 为第 i 个不等式约束的 **Lagrange** 乘子
- ν_i 为第 i 个等式约束的 **Lagrange** 乘子



Lagrange对偶函数





Lagrange对偶函数



□ Lagrange对偶函数:



Lagrange对偶函数



□ Lagrange对偶函数: $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$



Lagrange对偶函数



□ Lagrange对偶函数: $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$

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

$$= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right)$$



Lagrange对偶函数



□ Lagrange对偶函数: $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right) \end{aligned}$$

□ 函数 g 为凹函数



Lagrange对偶函数



□ Lagrange对偶函数: $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right) \end{aligned}$$

□ 函数 g 为凹函数

○ $\sup L(x, \lambda, \nu)$ 为凸函数



Lagrange对偶函数



□ Lagrange对偶函数: $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right) \end{aligned}$$

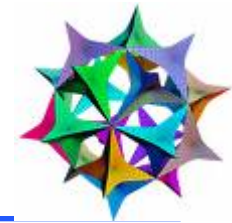
□ 函数 g 为凹函数

○ $\sup L(x, \lambda, \nu)$ 为凸函数

○ 对某些 λ, ν 可为 $-\infty$



Lagrange对偶函数





Lagrange对偶函数



□ 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$
- 因此，若 $\lambda \geq 0$ ，有



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$
- 因此，若 $\lambda \geq 0$ ，有
- $$\sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*) \leq 0$$



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$
- 因此，若 $\lambda \geq 0$ ，有
- $$\sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*) \leq 0$$
- $$L(x^*, \lambda, v) = f_0(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*)$$



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$
- 因此，若 $\lambda \geq 0$ ，有
- $$\sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*) \leq 0$$
- $$L(x^*, \lambda, v) = f_0(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*)$$
- $$\leq p^*$$



Lagrange对偶函数



- 最优值下界：若 $\lambda \geq 0$ ，则 $g(\lambda, v) \leq p^*$
- 证明：若 x^* 为原问题的最优解，则必可行，则有
- $f_i(x^*) \leq 0, h_i(x^*) = 0$
- 因此，若 $\lambda \geq 0$ ，有
- $$\sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*) \leq 0$$
- $$L(x^*, \lambda, v) = f_0(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^p v_i h_i(x^*)$$
- $$\leq p^*$$
- $g(\lambda, v) \leq p^*$



线性方程组的最小二乘解





线性方程组的最小二乘解



$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian** 是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian** 是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

- **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$
- 对偶函数 $g(\nu) = \inf_x L(x, \nu)$
- 为最小化关于 x 的 L ，令梯度为**0**:



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$

□ 为最小化关于 x 的 L ，令梯度为**0**:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \implies \quad x = -(1/2)A^T \nu$$



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$

□ 为最小化关于 x 的 L ，令梯度为**0**:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \implies \quad x = -(1/2)A^T \nu$$

□ 代入 L 得 g :



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$

□ 为最小化关于 x 的 L ，令梯度为**0**:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \implies \quad x = -(1/2)A^T \nu$$

□ 代入 L 得 g :

$$g(\nu) = L((-1/2)A^T \nu, \nu) = -\frac{1}{4}\nu^T AA^T \nu - b^T \nu$$



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$

□ 为最小化关于 x 的 L ，令梯度为**0**:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \implies \quad x = -(1/2)A^T \nu$$

□ 代入 L 得 g :

$$g(\nu) = L((-1/2)A^T \nu, \nu) = -\frac{1}{4}\nu^T AA^T \nu - b^T \nu$$

□ 为关于 ν 的凹函数



线性方程组的最小二乘解



$$\begin{aligned} & \text{minimize} && x^T x \\ & \text{subject to} && Ax = b \end{aligned}$$

□ **Lagrangian**是 $L(x, \nu) = x^T x + \nu^T (Ax - b)$

□ 对偶函数 $g(\nu) = \inf_x L(x, \nu)$

□ 为最小化关于 x 的 L ，令梯度为**0**:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \implies \quad x = -(1/2)A^T \nu$$

□ 代入 L 得 g :

$$g(\nu) = L((-1/2)A^T \nu, \nu) = -\frac{1}{4}\nu^T AA^T \nu - b^T \nu$$

□ 为关于 ν 的凹函数

□ 下界性质: $p^* \geq -(1/4)\nu^T AA^T \nu - b^T \nu$ for all ν

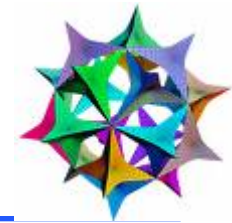


标准形式的线性规划





标准形式的线性规划



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$



标准形式的线性规划



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$

□ **Lagrangian**为



标准形式的线性规划



$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b, \quad x \succeq 0 \end{aligned}$$

□ **Lagrangian** 为 $L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x$

$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$



标准形式的线性规划



$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b, \quad x \succeq 0 \end{aligned}$$

□ **Lagrangian** 为 $L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x$

$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$

□ 函数 L 是仿射函数，因此



标准形式的线性规划



$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b, \quad x \succeq 0 \end{aligned}$$

□ **Lagrangian** 为 $L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x$

$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$

□ 函数 L 是仿射函数，因此

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$



标准形式的线性规划



$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b, \quad x \succeq 0 \end{aligned}$$

□ **Lagrangian**为 $L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x$

$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$

□ 函数 L 是仿射函数，因此

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 函数 g 是线性的，为仿射定义域 $\{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\}$ 上的凹函数



标准形式的线性规划



$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b, \quad x \succeq 0 \end{aligned}$$

□ **Lagrangian** 为 $L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x$

$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$

□ 函数 L 是仿射函数，因此

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 函数 g 是线性的，为仿射定义域 $\{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\}$ 上的凹函数

□ 下界的性质： $p^* \geq -b^T \nu$ if $A^T \nu + c \succeq 0$



双向划分问题





双向划分问题



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



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数

$$\begin{aligned} g(\nu) &= \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) = \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数

$$\begin{aligned} g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

- 下界的性质:



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数

$$\begin{aligned} g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

- 下界的性质:

$$p^* \geq -\mathbf{1}^T \nu \text{ if } W + \mathbf{diag}(\nu) \succeq 0$$



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数

$$\begin{aligned} g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

- 下界的性质：

$$p^* \geq -\mathbf{1}^T \nu \text{ if } W + \mathbf{diag}(\nu) \succeq 0$$

- 例如：



双向划分问题



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

- 非凸优化问题，可行集包含 2^n 个离散点
- 对偶函数

$$\begin{aligned} g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

- 下界的性质:

$$p^* \geq -\mathbf{1}^T \nu \text{ if } W + \mathbf{diag}(\nu) \succeq 0$$

- 例如:

- $\nu = -\lambda_{\min}(W)\mathbf{1} \quad p^* \geq n\lambda_{\min}(W)$



Lagrange对偶和共轭函数





Lagrange对偶和共轭函数



□ 共轭函数:



Lagrange对偶和共轭函数



□ 共轭函数: $f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$



Lagrange对偶和共轭函数



□ 共轭函数: $f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$

minimize $f(x)$

subject to $x = 0$



Lagrange对偶和共轭函数



□ 共轭函数: $f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$

minimize $f(x)$
subject to $x = 0$

$$g(\nu) = \inf_x (f(x) + \nu^T x)$$



Lagrange对偶和共轭函数



□ 共轭函数: $f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$

minimize $f(x)$
subject to $x = 0$

$$\begin{aligned} g(\nu) &= \inf_x (f(x) + \nu^T x) \\ &= -\sup_x ((-\nu)^T x - f(x)) \end{aligned}$$



Lagrange对偶和共轭函数



□ 共轭函数: $f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$

minimize $f(x)$
subject to $x = 0$

$$\begin{aligned} g(\nu) &= \inf_x (f(x) + \nu^T x) \\ &= -\sup_x ((-\nu)^T x - f(x)) \\ &= -f^*(-\nu) \end{aligned}$$



Lagrange对偶和共轭函数





Lagrange对偶和共轭函数



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b, \quad Cx = d \end{array}$$



Lagrange对偶和共轭函数



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b, \quad Cx = d \end{array}$$

□ 对偶函数



Lagrange对偶和共轭函数



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && Ax \preceq b, \quad Cx = d \end{aligned}$$

□ 对偶函数

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \text{dom } f_0} (f_0(x) + (A^T \lambda + C^T \nu)^T x - b^T \lambda - d^T \nu) \\ &= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu \end{aligned}$$



Lagrange对偶和共轭函数



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && Ax \preceq b, \quad Cx = d \end{aligned}$$

□ 对偶函数

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \text{dom } f_0} (f_0(x) + (A^T \lambda + C^T \nu)^T x - b^T \lambda - d^T \nu) \\ &= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu \end{aligned}$$

□ 若已知共轭函数，可简化对偶函数的表达



Lagrange对偶和共轭函数



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && Ax \preceq b, \quad Cx = d \end{aligned}$$

□ 对偶函数

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \text{dom } f_0} (f_0(x) + (A^T \lambda + C^T \nu)^T x - b^T \lambda - d^T \nu) \\ &= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu \end{aligned}$$

- 若已知共轭函数，可简化对偶函数的表达
- 如：熵的最大化



Lagrange对偶和共轭函数



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && Ax \preceq b, \quad Cx = d \end{aligned}$$

□ 对偶函数

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \text{dom } f_0} (f_0(x) + (A^T \lambda + C^T \nu)^T x - b^T \lambda - d^T \nu) \\ &= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu \end{aligned}$$

□ 若已知共轭函数，可简化对偶函数的表达

□ 如：熵的最大化

$$f_0(x) = \sum_{i=1}^n x_i \log x_i, \quad f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$



对偶问题





对偶问题



□ Lagrange对偶问题



对偶问题



□ **Lagrange**对偶问题

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$



对偶问题



- **Lagrange**对偶问题
$$\begin{aligned} & \text{maximize} && g(\lambda, \nu) \\ & \text{subject to} && \lambda \succeq 0 \end{aligned}$$
- 通过**Lagrange**对偶函数获取 p^* 的最佳下界



对偶问题



- **Lagrange**对偶问题
$$\begin{aligned} & \text{maximize} && g(\lambda, \nu) \\ & \text{subject to} && \lambda \succeq 0 \end{aligned}$$
- 通过**Lagrange**对偶函数获取 p^* 的最佳下界
- 凸优化问题，最优值表示为 d^*



对偶问题



- **Lagrange**对偶问题 $\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$
- 通过**Lagrange**对偶函数获取 p^* 的最佳下界
- 凸优化问题，最优值表示为 d^*
 $d^* \leq p^*$



对偶问题



- **Lagrange**对偶问题
$$\begin{aligned} & \text{maximize} && g(\lambda, \nu) \\ & \text{subject to} && \lambda \succeq 0 \end{aligned}$$
- 通过**Lagrange**对偶函数获取 p^* 的最佳下界
- 凸优化问题，最优值表示为 d^*
$$d^* \leq p^*$$
- 对偶可行： $\lambda \succeq 0$ 和 $g(\lambda, \nu) > -\infty$



对偶问题



- **Lagrange**对偶问题
$$\begin{aligned} & \text{maximize} && g(\lambda, \nu) \\ & \text{subject to} && \lambda \succeq 0 \end{aligned}$$
- 通过**Lagrange**对偶函数获取 p^* 的最佳下界
- 凸优化问题，最优值表示为 d^*
$$d^* \leq p^*$$
- 对偶可行: $\lambda \succeq 0$ 和 $g(\lambda, \nu) > -\infty$
- 最优**Lagrange**乘子: (λ^*, ν^*)



对偶问题





对偶问题



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



对偶问题



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

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$



对偶问题



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

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\begin{aligned} & \text{maximize} && -b^T \nu \\ & \text{subject to} && A^T \nu - \lambda + c = 0 \\ & && \lambda \succeq 0. \end{aligned}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu - \lambda + c = 0 \\ & \lambda \succeq 0. \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \preceq 0, \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \square L(x, \lambda) = c^T x + \lambda^T (Ax - b)$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \begin{array}{l} \square L(x, \lambda) = c^T x + \lambda^T (Ax - b) \\ \square \quad \quad \quad = (c + A^T \lambda)^T x - b^T \lambda \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \begin{array}{l} \square L(x, \lambda) = c^T x + \lambda^T (Ax - b) \\ \square \quad \quad \quad = (c + A^T \lambda)^T x - b^T \lambda \\ \square g(\lambda) = \inf_x L(x, \lambda) \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \begin{array}{l} \square L(x, \lambda) = c^T x + \lambda^T (Ax - b) \\ \square \quad \quad \quad = (c + A^T \lambda)^T x - b^T \lambda \\ \square g(\lambda) = \inf_x L(x, \lambda) \\ \square \quad \quad \quad = \begin{cases} -b^T \lambda & \text{if } A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \end{array}$$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0, \end{array}$$

□ $= \begin{cases} -b^T \lambda & \text{if } A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$



对偶问题



$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array} \quad \begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0, \end{array}$$



弱对偶和强对偶





弱对偶和强对偶



□ 弱对偶：



弱对偶和强对偶



□ 弱对偶：

○ 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$



弱对偶和强对偶



□ 弱对偶：

- 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$
- 可以用来寻找困难问题的非平凡下界



弱对偶和强对偶



□ 弱对偶：

- 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$
- 可以用来寻找困难问题的非平凡下界

□ 例：



弱对偶和强对偶



- 弱对偶：
 - 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$
 - 可以用来寻找困难问题的非平凡下界
- 例：
 - 求解半定规划问题



弱对偶和强对偶



□ 弱对偶：

- 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$
- 可以用来寻找困难问题的非平凡下界

□ 例：

- 求解半定规划问题

$$\begin{aligned} & \text{maximize} && -\mathbf{1}^T \boldsymbol{\nu} \\ & \text{subject to} && W + \mathbf{diag}(\boldsymbol{\nu}) \succeq 0 \end{aligned}$$



弱对偶和强对偶



□ 弱对偶：

- 对于凸优化和非凸优化问题总是满足 $d^* \leq p^*$
- 可以用来寻找困难问题的非平凡下界

□ 例：

- 求解半定规划问题

$$\begin{aligned} & \text{maximize} && -\mathbf{1}^T \boldsymbol{\nu} \\ & \text{subject to} && W + \mathbf{diag}(\boldsymbol{\nu}) \succeq 0 \end{aligned}$$

- 为双向划分问题给出了下界



弱对偶和强对偶





弱对偶和强对偶



□ 强对偶 $d^* = p^*$



弱对偶和强对偶



- 强对偶 $d^* = p^*$
- 并非总是满足



弱对偶和强对偶



- 强对偶 $d^* = p^*$
 - 并非总是满足
 - 通常在凸优化问题中满足



弱对偶和强对偶



- 强对偶 $d^* = p^*$
 - 并非总是满足
 - 通常在凸优化问题中满足
- 对偶间隙: $p^* - d^*$



弱对偶和强对偶



- 强对偶 $d^* = p^*$
 - 并非总是满足
 - 通常在凸优化问题中满足
- 对偶间隙: $p^* - d^*$
- 在凸优化问题中保证强对偶的条件称为**约束准则**



弱对偶和强对偶



- 强对偶 $d^* = p^*$
 - 并非总是满足
 - 通常在凸优化问题中满足
- 对偶间隙: $p^* - d^*$
- 在凸优化问题中保证强对偶的条件称为**约束准则**
- D 的相对内部:



弱对偶和强对偶



- 强对偶 $d^* = p^*$
 - 并非总是满足
 - 通常在凸优化问题中满足
- 对偶间隙: $p^* - d^*$
- 在凸优化问题中保证强对偶的条件称为**约束准则**
- D 的相对内部:
 - $\text{Rel int } D = \{x \in D \mid B(x, r) \cap \text{aff } D \subset D, \exists r > 0\}$



Slater约束准则





Slater约束准则



□ 若某凸优化问题



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

○ 当该问题是严格可行的，即



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

○ 当该问题是严格可行的，即

$$\exists x \in \text{relint } \mathcal{D} \quad f_i(x) < 0, \quad i = 1, \dots, m, \quad Ax = b$$



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

○ 当该问题是严格可行的，即

$$\exists x \in \text{relint } \mathcal{D} \quad f_i(x) < 0, \quad i = 1, \dots, m, \quad Ax = b$$

○ 则为强对偶，即满足 $d^* = p^*$



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

○ 当该问题是严格可行的，即

$$\exists x \in \text{relint } \mathcal{D} \quad f_i(x) < 0, \quad i = 1, \dots, m, \quad Ax = b$$

○ 则为强对偶，即满足 $d^* = p^*$

□ 弱化：



Slater约束准则



□ 若某凸优化问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

○ 当该问题是严格可行的，即

$$\exists x \in \text{relint } \mathcal{D} \quad f_i(x) < 0, \quad i = 1, \dots, m, \quad Ax = b$$

○ 则为强对偶，即满足 $d^* = p^*$

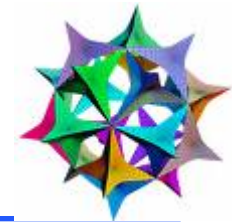
□ 弱化：

○ 若不等式约束为仿射时，只要可行域非空，必有

$$d^* = p^*$$

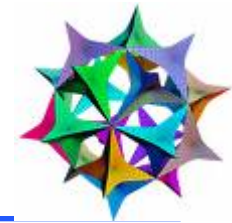


线性规划的不等式形式





线性规划的不等式形式



□ 原问题



线性规划的不等式形式



□ 原问题

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax \preceq b \end{aligned}$$



线性规划的不等式形式



- 原问题 minimize $c^T x$
subject to $Ax \preceq b$
- 对偶函数



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 对偶问题



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 对偶问题 maximize $-b^T \lambda$
subject to $A^T \lambda + c = 0, \quad \lambda \succeq 0$



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 对偶问题 maximize $-b^T \lambda$
subject to $A^T \lambda + c = 0, \quad \lambda \succeq 0$

□ 根据**Slater**条件：若对某些 \tilde{x} 有 $A\tilde{x} \prec b$ ，则 $p^* = d^*$



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 对偶问题 maximize $-b^T \lambda$
subject to $A^T \lambda + c = 0, \quad \lambda \succeq 0$

□ 根据**Slater**条件：若对某些 \tilde{x} 有 $A\tilde{x} \prec b$ ，则 $p^* = d^*$

□ 根据弱化条件，只要原问题和对偶问题可行



线性规划的不等式形式



□ 原问题 minimize $c^T x$
subject to $Ax \preceq b$

□ 对偶函数

$$g(\lambda) = \inf_x ((c + A^T \lambda)^T x - b^T \lambda) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

□ 对偶问题 maximize $-b^T \lambda$
subject to $A^T \lambda + c = 0, \quad \lambda \succeq 0$

□ 根据**Slater**条件：若对某些 \tilde{x} 有 $A\tilde{x} \prec b$ ，则 $p^* = d^*$

□ 根据弱化条件，只要原问题和对偶问题可行

○ 必有 $p^* = d^*$



线性方程组的最小二乘解





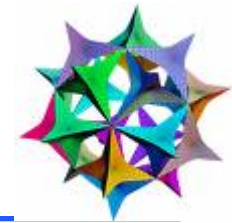
线性方程组的最小二乘解



$$\begin{aligned} &\text{minimize} && x^T x \\ &\text{subject to} && Ax = b \end{aligned}$$



线性方程组的最小二乘解



$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

□ 对偶问题:



线性方程组的最小二乘解



$$\begin{aligned} &\text{minimize} && x^T x \\ &\text{subject to} && Ax = b \end{aligned}$$

□ 对偶问题:

$$\text{maximize} \quad -(1/4)\nu^T AA^T \nu - b^T \nu$$



线性方程组的最小二乘解



$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

□ 对偶问题:

$$\text{maximize} \quad -(1/4)\nu^T AA^T \nu - b^T \nu$$

□ 强对偶成立



QCQP二次约束二次规划





QCQP二次约束二次规划



$$\begin{aligned} &\text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ &\text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

□ **Lagrangian:**



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

□ Lagrangian:

$$\bullet L(x, \lambda) = \frac{1}{2}x^T P_0 x + q_0^T x + r_0 + \sum_{i=1}^m \left(\frac{1}{2}\lambda_i x^T P_i x + \lambda_i q_i^T x + \lambda_i r_i \right)$$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

□ Lagrangian:

$$\bullet L(x, \lambda) = \frac{1}{2}x^T P_0 x + q_0^T x + r_0 + \sum_{i=1}^m \left(\frac{1}{2}\lambda_i x^T P_i x + \lambda_i q_i^T x + \lambda_i r_i \right)$$

$$\bullet = \frac{1}{2}x^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + \left(r_0 + \sum_{i=1}^m \lambda_i r_i \right)$$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

□ Lagrangian:

$$\bullet L(x, \lambda) = \frac{1}{2}x^T P_0 x + q_0^T x + r_0 + \sum_{i=1}^m \left(\frac{1}{2}\lambda_i x^T P_i x + \lambda_i q_i^T x + \lambda_i r_i \right)$$

$$\bullet = \frac{1}{2}x^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + \left(r_0 + \sum_{i=1}^m \lambda_i r_i \right)$$

$P(\lambda)$ $q(\lambda)$ $r(\lambda)$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

$$\square L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda)$$

$$\circ L(x, \lambda) = \frac{1}{2}x^T P_0 x + q_0^T x + r_0 + \sum_{i=1}^m \left(\frac{1}{2}\lambda_i x^T P_i x + \lambda_i q_i^T x + \lambda_i r_i \right)$$

$$\circ = \frac{1}{2}x^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + \left(r_0 + \sum_{i=1}^m \lambda_i r_i \right)$$

$P(\lambda)$

$q(\lambda)$

$r(\lambda)$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \\ & && L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda) \end{aligned}$$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \\ & && L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda) \end{aligned}$$

□ 对偶函数：



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

$$L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda)$$

□ 对偶函数:

$$g(\lambda) = \inf_x L(x, \lambda) = -(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda)$$



QCQP二次约束二次规划



minimize $(1/2)x^T P_0 x + q_0^T x + r_0$
subject to $(1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m,$

$$L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda)$$

□ 对偶函数:

$$g(\lambda) = \inf_x L(x, \lambda) = -(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda)$$

□ 对偶问题:



QCQP二次约束二次规划



minimize $(1/2)x^T P_0 x + q_0^T x + r_0$
subject to $(1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m,$

$$L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda)$$

□ 对偶函数:

$$g(\lambda) = \inf_x L(x, \lambda) = -(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda)$$

□ 对偶问题:

maximize $-(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda)$
subject to $\lambda \succeq 0.$



QCQP二次约束二次规划



$$\begin{aligned} & \text{minimize} && (1/2)x^T P_0 x + q_0^T x + r_0 \\ & \text{subject to} && (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \\ & && L(x, \lambda) = (1/2)x^T P(\lambda)x + q(\lambda)^T x + r(\lambda) \end{aligned}$$

□ 对偶函数:

$$g(\lambda) = \inf_x L(x, \lambda) = -(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda)$$

□ 对偶问题:

$$\begin{aligned} & \text{maximize} && -(1/2)q(\lambda)^T P(\lambda)^{-1} q(\lambda) + r(\lambda) \\ & \text{subject to} && \lambda \succeq 0. \end{aligned}$$

□ 存在 x , 满足

$$(1/2)x^T P_i x + q_i^T x + r_i < 0, \quad i = 1, \dots, m$$



几何解释





几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题



几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$



几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$



几何解释



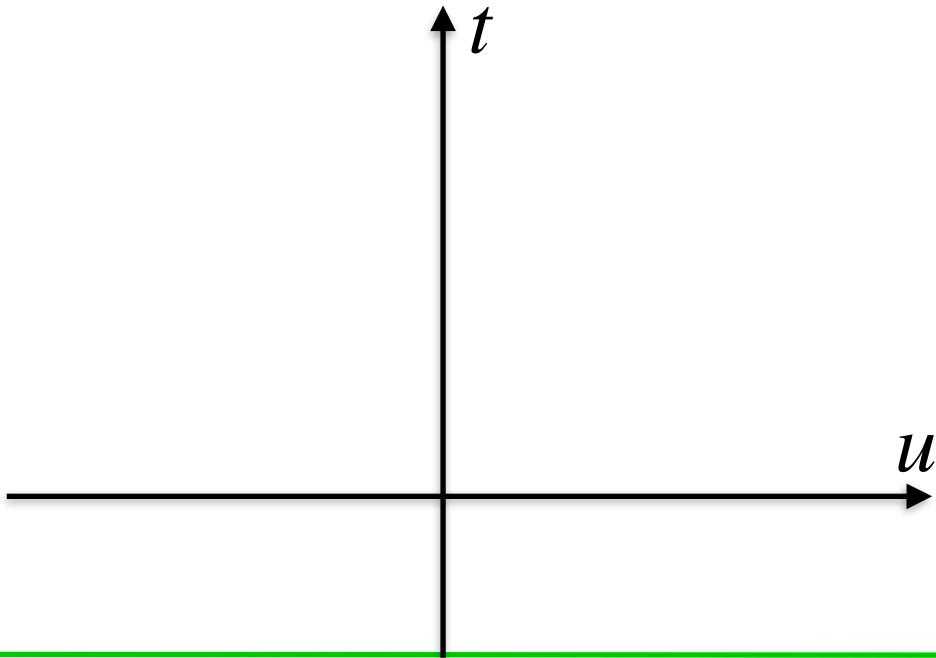
- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$



几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

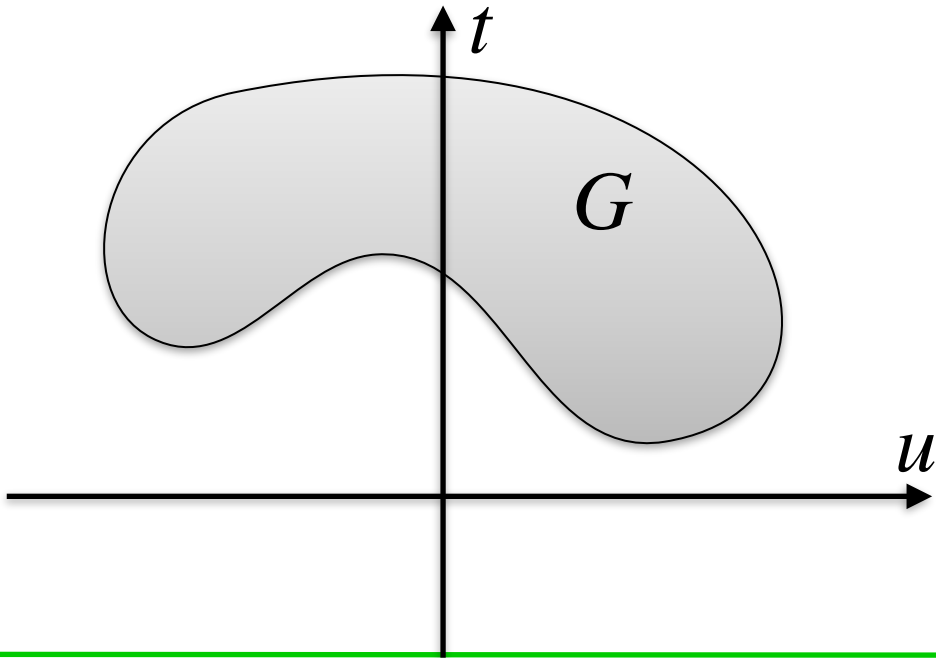




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

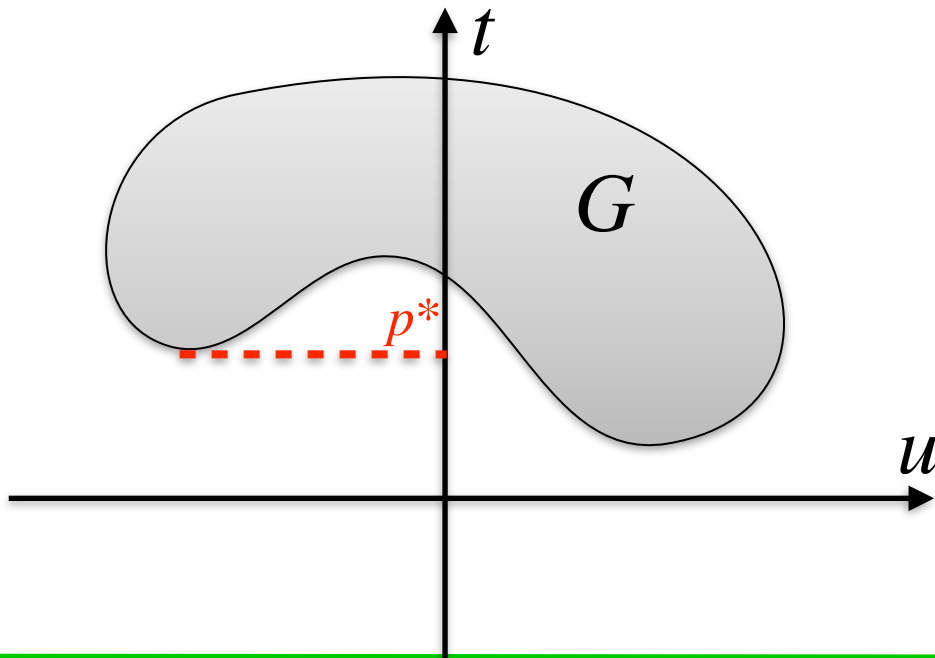




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

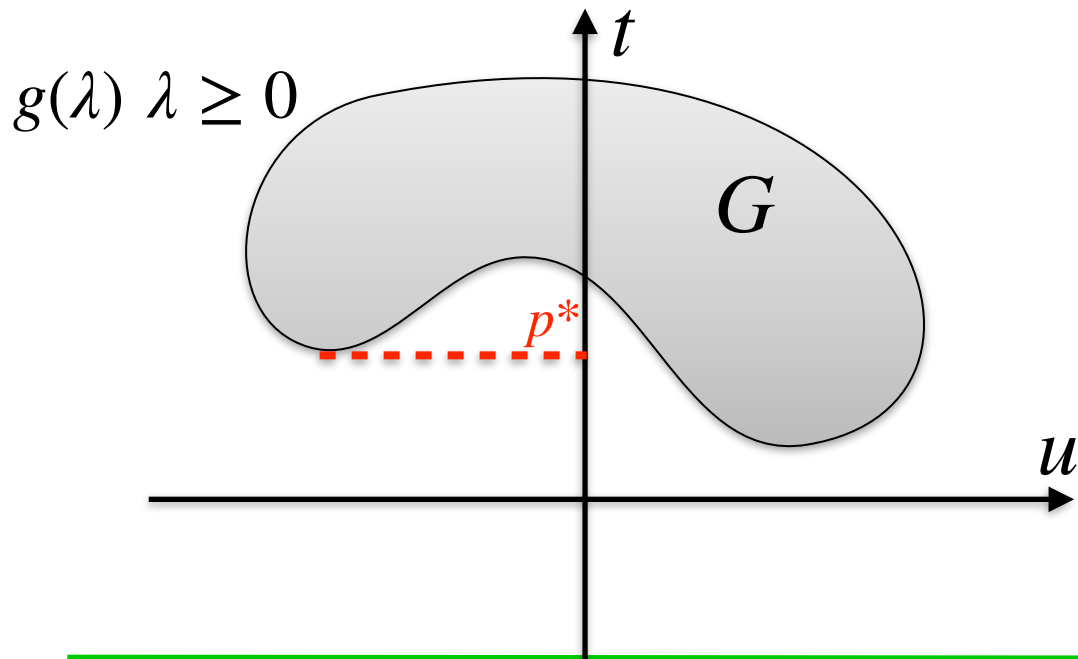




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

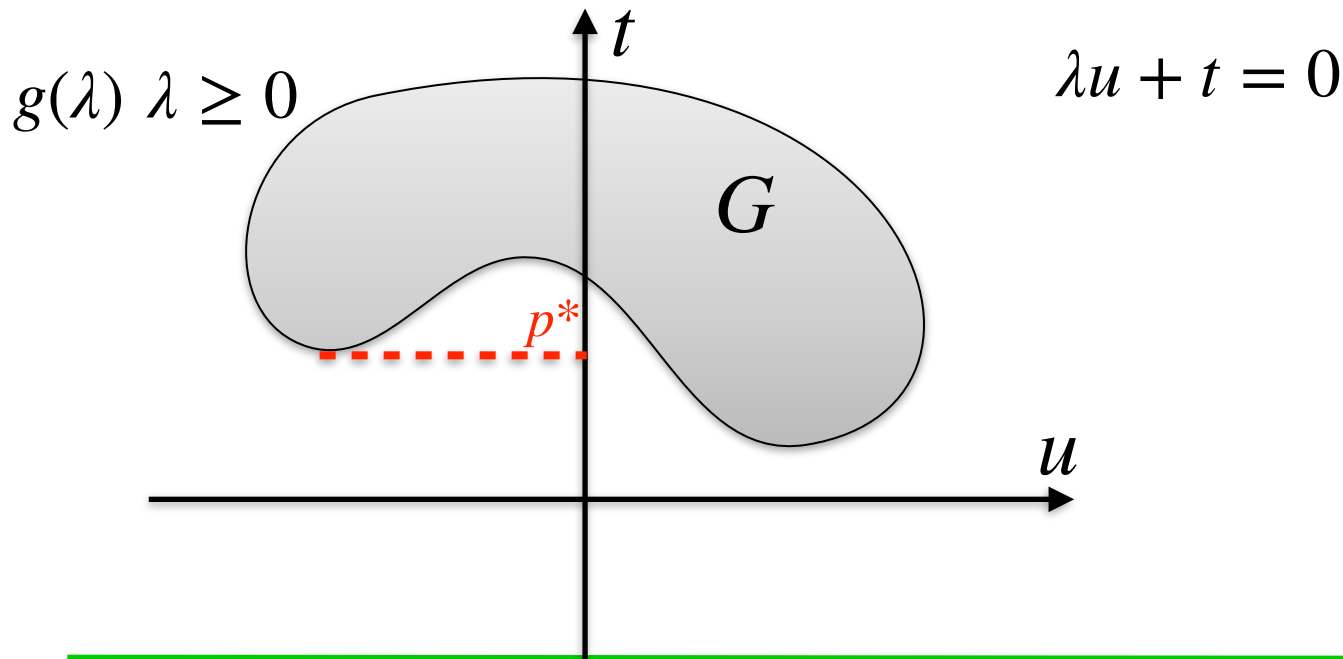




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

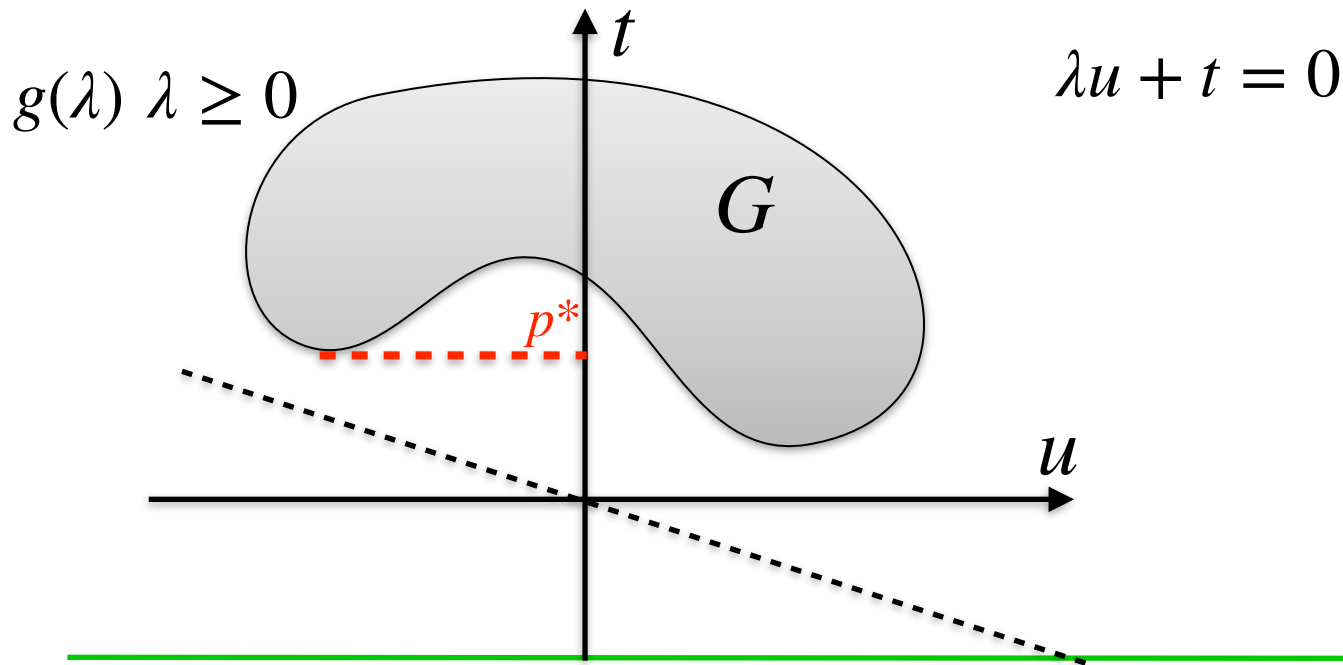




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

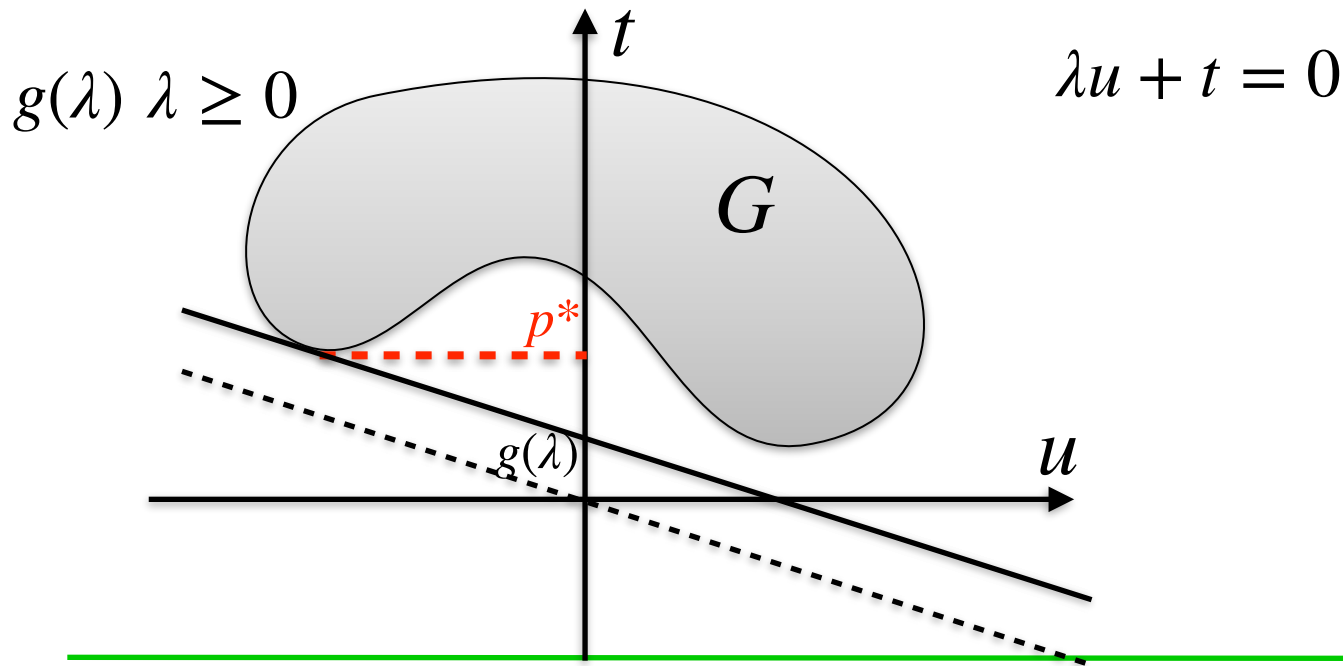




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$

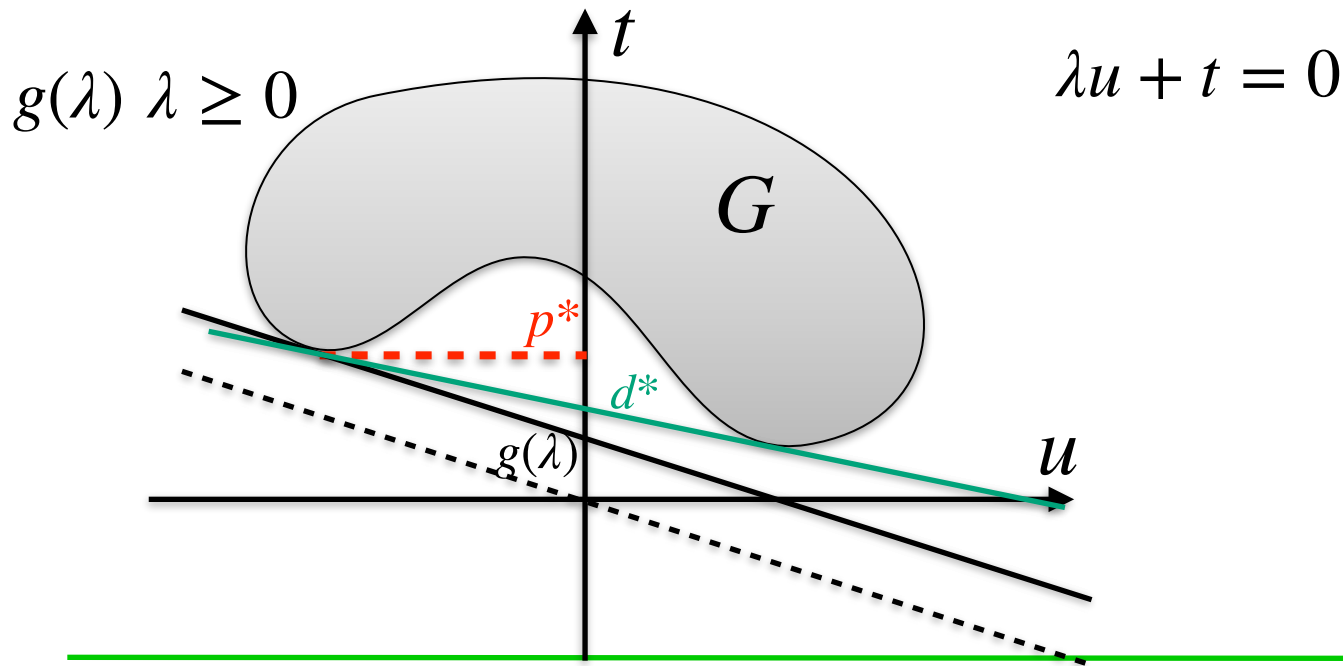




几何解释



- 考虑只含一个约束 $f_1(x) \leq 0$ 的问题
- 定义集合 $G = \{(f_0(x), f_1(x)) \mid x \in D\}$
- $p^* = \inf\{t \mid (u, t) \in G, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \mid (u, t) \in G\}$



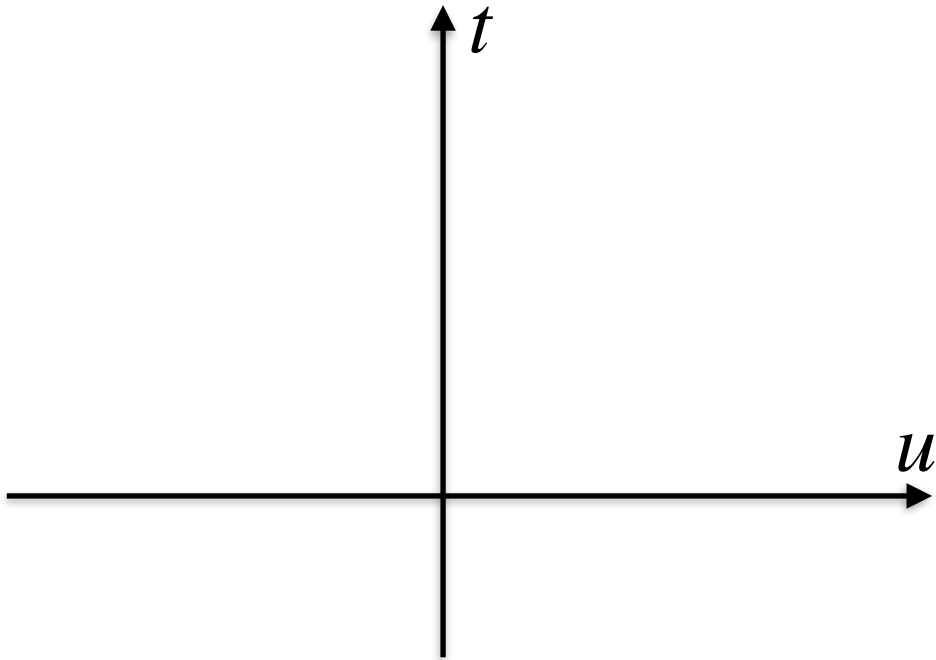


几何解释



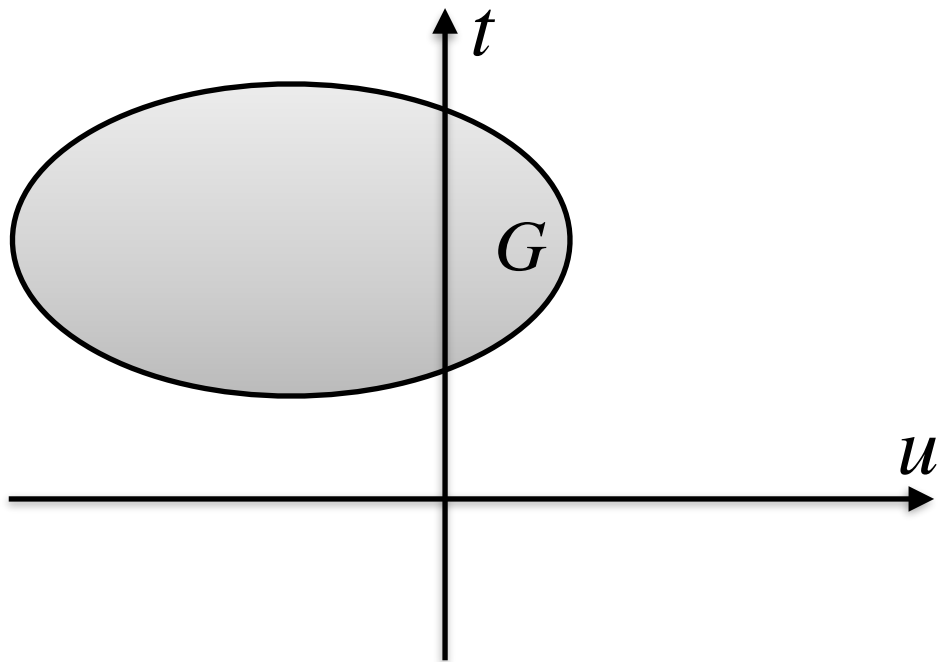


几何解释



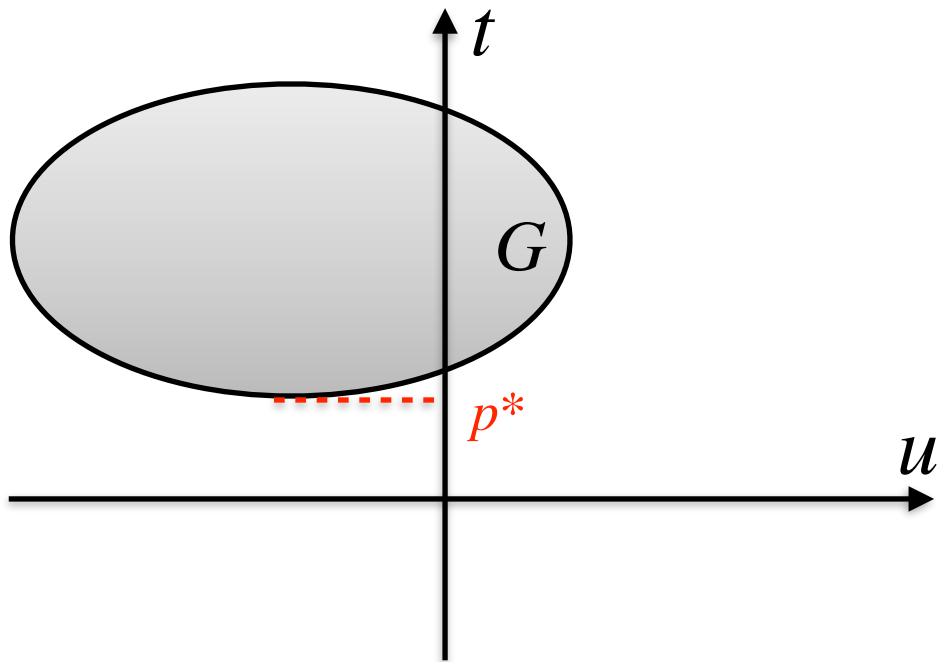


几何解释



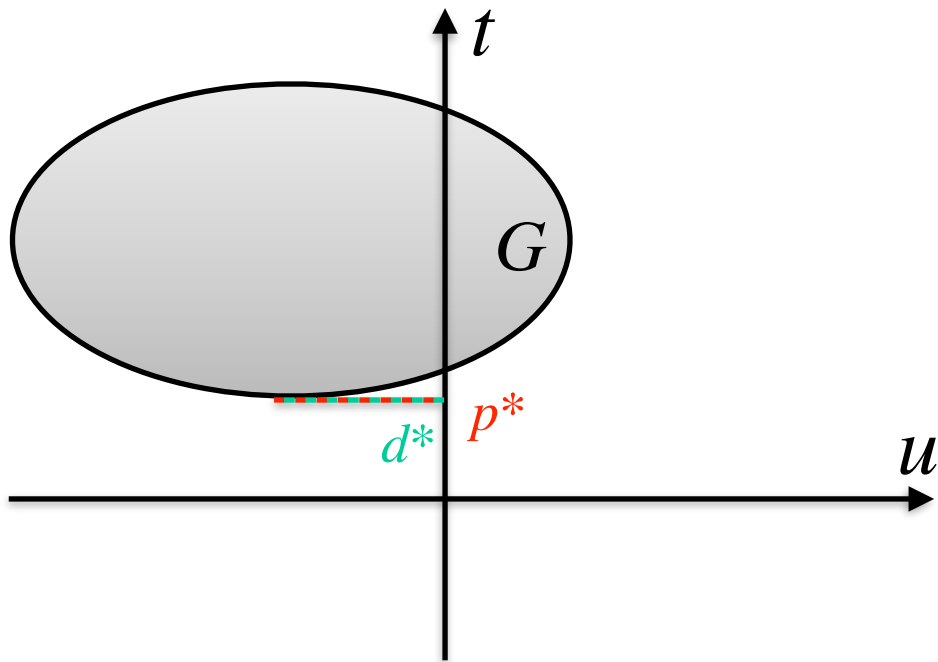


几何解释





几何解释





经济学解释





经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

○ x : 产品数量



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

- x : 产品数量
- $f_i(x) \leq 0$: 原材料



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

- x : 产品数量
- $f_i(x) \leq 0$: 原材料
- $-f_0(x)$: 利润



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

- x : 产品数量
- $f_i(x) \leq 0$: 原材料
- $-f_0(x)$: 利润
- $f_0(x)$: 损失



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

- x : 产品数量
- $f_i(x) \leq 0$: 原材料
- $-f_0(x)$: 利润
- $f_0(x)$: 损失
- p^* : 最优损失



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 $\lambda_i (\lambda_i \geq 0)$



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 λ_i ($\lambda_i \geq 0$)

$$\text{○ } \min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$$



经济学解释



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 λ_i ($\lambda_i \geq 0$)

$$\text{○ } \min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$$

$$\text{○ } d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$$



经济学解释



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 $\lambda_i (\lambda_i \geq 0)$

○ $\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$

○ $d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$

□ 何时 $d^* = p^*$?



经济学解释



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 $\lambda_i (\lambda_i \geq 0)$

○ $\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$

○ $d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$

□ 何时 $d^* = p^*$?

□ 若 $f_i(x^*) < 0$,



经济学解释



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 $\lambda_i (\lambda_i \geq 0)$

○ $\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$

○ $d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$

□ 何时 $d^* = p^*$?

□ 若 $f_i(x^*) < 0$,

□ 则 $\lambda_i = 0$



经济学解释



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

- x : 产品数量
- $f_i(x) \leq 0$: 原材料
- $-f_0(x)$: 利润
- $f_0(x)$: 损失
- p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 λ_i ($\lambda_i \geq 0$)

- $\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$

- $d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$

- 何时 $d^* = p^*$?
- 若 $f_i(x^*) < 0$,
 - 则 $\lambda_i = 0$
- 若 $\lambda_i^* > 0$,



经济学解释



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

○ x : 产品数量

○ $f_i(x) \leq 0$: 原材料

○ $-f_0(x)$: 利润

○ $f_0(x)$: 损失

○ p^* : 最优损失

□ 假设原材料可买可卖，第 i 种原材料价格为 λ_i ($\lambda_i \geq 0$)

○ $\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) = g(\lambda)$

○ $d^* = \max_{\lambda \geq 0} g(\lambda) \leq p^*$

- 何时 $d^* = p^*$?
- 若 $f_i(x^*) < 0$,
 - 则 $\lambda_i = 0$
- 若 $\lambda_i^* > 0$,
 - 则 $f_i(x^*) = 0$



鞍点解释





鞍点解释



□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



鞍点解释



□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:



鞍点解释



□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



鞍点解释

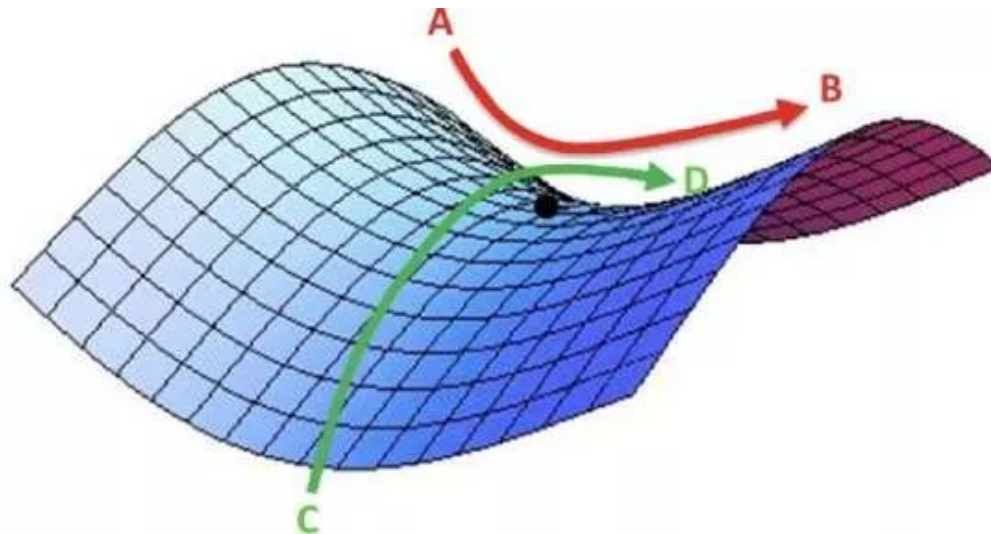


□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$





鞍点解释

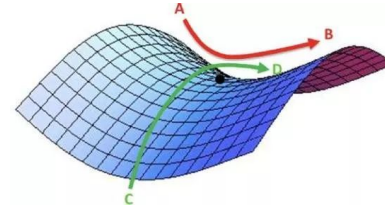


□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$





鞍点解释

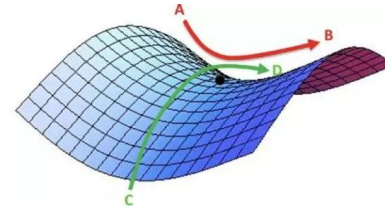


□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



○ $\arg_{z,w} \sup_{z \in Z} \inf_{w \in W} f(w, z) = \arg_{w,z} \inf_{w \in W} \sup_{z \in Z} f(w, z)$



鞍点解释

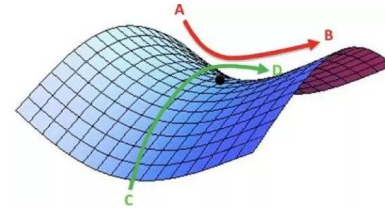


□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



○ $\arg_{z,w} \sup_{z \in Z} \inf_{w \in W} f(w, z) = \arg_{w,z} \inf_{w \in W} \sup_{z \in Z} f(w, z)$

○ 鞍点 (\tilde{w}, \tilde{z}) : $f(\tilde{w}, z) \leq f(\tilde{w}, \tilde{z}) \leq f(w, \tilde{z})$



鞍点解释

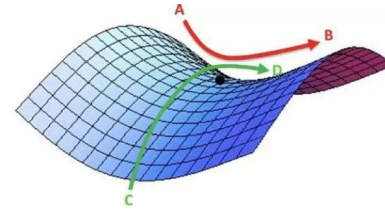


□ 极大极小不等式:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ 若等式成立:

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) = \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



○ $\arg_{z,w} \sup_{z \in Z} \inf_{w \in W} f(w, z) = \arg_{w,z} \inf_{w \in W} \sup_{z \in Z} f(w, z)$

○ 鞍点 (\tilde{w}, \tilde{z}) : $f(\tilde{w}, z) \leq f(\tilde{w}, \tilde{z}) \leq f(w, \tilde{z})$

○ $f(\tilde{w}, \tilde{z}) = \inf_{w \in W} f(w, \tilde{z}), f(\tilde{w}, \tilde{z}) = \sup_{z \in Z} f(\tilde{w}, z)$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

□ $d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

- $d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$
- $p^* = \inf_x \{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m\}$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

$$\square \quad d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$$

$$\square \quad p^* = \inf_x \{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m\}$$

$$\square \quad = \inf_x \begin{cases} f_0(x) & f_i(x) \leq 0, i = 1, \dots, m \\ \infty & \text{otherwise} \end{cases}$$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

$$\square d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$$

$$\square p^* = \inf_x \{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m\}$$

$$\square = \inf_x \begin{cases} f_0(x) & f_i(x) \leq 0, i = 1, \dots, m \\ \infty & \text{otherwise} \end{cases}$$

$$\square = \inf_x \sup_{\lambda \geq 0} (f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)) = \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

$$\square d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$$

$$\square p^* = \inf_x \{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m\}$$

$$\square = \inf_x \begin{cases} f_0(x) & f_i(x) \leq 0, i = 1, \dots, m \\ \infty & \text{otherwise} \end{cases}$$

$$\square = \inf_x \sup_{\lambda \geq 0} (f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)) = \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$$

$$\square p^* \leq d^*, \text{ 何时 } p^* = d^*?$$



鞍点解释



$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \leq \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

$$\square d^* = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$$

$$\square p^* = \inf_x \{f_0(x) \mid f_i(x) \leq 0, i = 1, \dots, m\}$$

$$\square = \inf_x \begin{cases} f_0(x) & f_i(x) \leq 0, i = 1, \dots, m \\ \infty & \text{otherwise} \end{cases}$$

$$\square = \inf_x \sup_{\lambda \geq 0} (f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)) = \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$$

$\square p^* \leq d^*$, 何时 $p^* = d^*$?

$\circ L(x, \lambda)$ 存在鞍点 (x^*, λ^*) : $\inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$



鞍点定理





鞍点定理



□ 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于



鞍点定理



- 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于
 - 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解



鞍点定理



- 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于
 - 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解
- 证明：→



鞍点定理



- 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于
 - 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解
- 证明：→
 - 存在鞍点 $\inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$ ，则对偶间隙为零



鞍点定理



- 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于
 - 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解
- 证明：→
 - 存在鞍点 $\inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$ ，则对偶间隙为零
 - $(\tilde{x}, \tilde{\lambda}) = \arg_{x, \lambda} \inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \arg_{x, \lambda} \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$



鞍点定理



□ 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于

○ 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解

□ 证明：→

○ 存在鞍点 $\inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$ ，则对偶间隙为零

○ $(\tilde{x}, \tilde{\lambda}) = \arg_{x, \lambda} \inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \arg_{x, \lambda} \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$

● $\tilde{x} = \arg_x \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$



鞍点定理



□ 若 $(\tilde{x}, \tilde{\lambda})$ 为 $L(x, \lambda)$ 的鞍点，等价于

○ 强对偶存在， $\tilde{x}, \tilde{\lambda}$ 为原问题和对偶问题最优解

□ 证明：→

○ 存在鞍点 $\inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$ ，则对偶间隙为零

○ $(\tilde{x}, \tilde{\lambda}) = \arg_{x, \lambda} \inf_x \sup_{\lambda \geq 0} L(x, \lambda) = \arg_{x, \lambda} \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$

● $\tilde{x} = \arg_x \inf_x \sup_{\lambda \geq 0} L(x, \lambda)$

● $\tilde{\lambda} = \arg_{\lambda} \sup_{\lambda \geq 0} \inf_x L(x, \lambda)$



鞍点定理





鞍点定理



□ 证明： ←



鞍点定理



□ 证明: ←

○ $f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$



鞍点定理



□ 证明: ←

○ $f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$

○ $f_0(\tilde{x}) = g(\tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\}$



鞍点定理



□ 证明: ←

○ $f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$

○ $f_0(\tilde{x}) = g(\tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\}$

○ $\leq f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) \leq f_0(\tilde{x})$



鞍点定理



□ 证明: ←

○ $f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$

○ $f_0(\tilde{x}) = g(\tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\}$

○ $\leq f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) \leq f_0(\tilde{x})$

□ $f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$



鞍点定理



□ 证明: ←

○ $f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$

○ $f_0(\tilde{x}) = g(\tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\}$

○ $\leq f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) \leq f_0(\tilde{x})$

□ $f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$

□ $f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$
$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$
$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda}_i \geq 0$$

$$\square \inf_x L(x, \tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} = f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$

$$\square \inf_x L(x, \tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} = f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) = \sup_{\lambda \geq 0} \{f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i f_i(\tilde{x})\} \leq f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$

$$\square \inf_x L(x, \tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} = f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) = \sup_{\lambda \geq 0} \{f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i f_i(\tilde{x})\} \leq f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) \geq L(\tilde{x}, \tilde{\lambda})$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$

$$\square \inf_x L(x, \tilde{\lambda}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} = f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) = \sup_{\lambda \geq 0} \{f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i f_i(\tilde{x})\} \leq f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) \geq L(\tilde{x}, \tilde{\lambda})$$

$$\square \sup_{\lambda \geq 0} L(\tilde{x}, \lambda) = L(\tilde{x}, \tilde{\lambda})$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$
$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$

$$\square \tilde{\lambda}_i f_i(\tilde{x}) = 0$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda} \geq 0$$

$$\square \tilde{\lambda}_i f_i(\tilde{x}) = 0$$

$$\square \text{若 } f_i(\tilde{x}) < 0,$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda}_i \geq 0$$

$$\square \tilde{\lambda}_i f_i(\tilde{x}) = 0$$

$$\square \text{若 } f_i(\tilde{x}) < 0,$$

$$\quad \bullet \text{ 则 } \tilde{\lambda}_i = 0$$



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$
$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda}_i \geq 0$$

- $\tilde{\lambda}_i f_i(\tilde{x}) = 0$
- 若 $f_i(\tilde{x}) < 0$,
 - 则 $\tilde{\lambda}_i = 0$
- 若 $\tilde{\lambda}_i > 0$,



鞍点定理



$$\square f_0(\tilde{x}) = f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) = L(\tilde{x}, \tilde{\lambda})$$

$$\square f_0(\tilde{x}) = \inf_x \{f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x)\} \quad f_i(\tilde{x}) \leq 0, \tilde{\lambda}_i \geq 0$$

$$\square \tilde{\lambda}_i f_i(\tilde{x}) = 0$$

$$\square \text{若 } f_i(\tilde{x}) < 0,$$

$$\circ \text{则 } \tilde{\lambda}_i = 0$$

$$\square \text{若 } \tilde{\lambda}_i > 0,$$

$$\circ \text{则 } f_i(\tilde{x}) = 0$$



互补松弛性





互补松弛性



minimize $f_0(x)$
subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $h_i(x) = 0, \quad i = 1, \dots, p$



互补松弛性



minimize $f_0(x)$
subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $h_i(x) = 0, \quad i = 1, \dots, p$

□ 对偶函数:
$$g(\lambda, \nu) = \inf_x \left\{ f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right\}$$



互补松弛性



minimize $f_0(x)$
subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $h_i(x) = 0, \quad i = 1, \dots, p$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$



互补松弛性



minimize $f_0(x)$
subject to $f_i(x) \leq 0, \quad i = 1, \dots, m$
 $h_i(x) = 0, \quad i = 1, \dots, p$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$

□ 若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, ν^* 须满足



互补松弛性



$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$

□ 若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, ν^* 须满足

- $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda^* \geq 0$



互补松弛性



$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$

□ 若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, ν^* 须满足

○ $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda^* \geq 0$

○ $f_0(x^*) = g(\lambda^*, \nu^*) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x)\}$



互补松弛性



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$

□ 若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, ν^* 须满足

○ $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda^* \geq 0$

○ $f_0(x^*) = g(\lambda^*, \nu^*) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x)\}$

○ $\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*) \leq f_0(x^*)$



互补松弛性



$$\max_{\lambda, v} \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

对偶函数: $g(\lambda, v) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p v_i h_i(x)\}$

对偶问题: $\max g(\lambda, v) \text{ s.t. } \lambda \geq 0$

若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, v^* 须满足

○ $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda_i^* \geq 0$

○ $f_0(x^*) = g(\lambda^*, v^*) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\}$

○ $\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \leq f_0(x^*)$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

$$\square \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

$$\square \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

$$\square \lambda_i^* f_i(x^*) \leq 0 \rightarrow \lambda_i^* f_i(x^*) = 0, i = 1, \dots, m$$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

$$\square \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

$$\square \lambda_i^* f_i(x^*) \leq 0 \rightarrow \lambda_i^* f_i(x^*) = 0, i = 1, \dots, m$$

$$\square \lambda_i^* > 0 \rightarrow f_i(x^*) = 0$$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

$$\square \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

$$\square \lambda_i^* f_i(x^*) \leq 0 \rightarrow \lambda_i^* f_i(x^*) = 0, i = 1, \dots, m$$

$$\square \lambda_i^* > 0 \rightarrow f_i(x^*) = 0$$

$$\square f_i(x^*) > 0 \rightarrow \lambda_i = 0$$



互补松弛性



$$\sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) = 0$$

$$\square \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

$$\square \lambda_i^* f_i(x^*) \leq 0 \rightarrow \lambda_i^* f_i(x^*) = 0, i = 1, \dots, m$$

$$\square \lambda_i^* > 0 \rightarrow f_i(x^*) = 0$$

$$\square f_i(x^*) > 0 \rightarrow \lambda_i = 0$$

□ 互补松弛



互补松弛性



$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

□ 对偶函数: $g(\lambda, \nu) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)\}$

□ 对偶问题: $\max g(\lambda, \nu) \text{ s.t. } \lambda \geq 0$

□ 若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, ν^* 须满足

○ $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda^* \geq 0$

○ $f_0(x^*) = g(\lambda^*, \nu^*) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x)\}$

○ $\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*) \leq f_0(x^*)$



互补松弛性



$$\inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\}$$

$$= f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*)$$

对偶问题: $\max g(\lambda, v)$ s.t. $\lambda \geq 0$

若 $p^* = d^*$ 且所有函数均可微, 则最优解 x^*, λ^*, v^* 须满足

○ $f_i(x^*) \leq 0, h_i(x^*) = 0, \lambda_i^* \geq 0$

○ $f_0(x^*) = g(\lambda^*, v^*) = \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\}$

○ $\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \leq f_0(x^*)$



稳定性



$$\begin{aligned} \square & \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\} \\ \square & = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \end{aligned}$$



稳定性



$$\begin{aligned} \square & \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\} \\ \square & = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \end{aligned}$$

$$\square L(x^*, \lambda^*, v^*) = \inf_x L(x, \lambda^*, v^*)$$



稳定性



$$\begin{aligned} \square & \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\} \\ \square & = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \end{aligned}$$

$$\square L(x^*, \lambda^*, v^*) = \inf_x L(x, \lambda^*, v^*)$$

$$\square \text{ 则 } \frac{\partial L(x, \lambda^*, v^*)}{\partial x} \Big|_{x=x^*} = 0$$



稳定性



$$\begin{aligned} \square & \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\} \\ \square & = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \end{aligned}$$

$$\square L(x^*, \lambda^*, v^*) = \inf_x L(x, \lambda^*, v^*)$$

$$\square \text{ 则 } \frac{\partial L(x, \lambda^*, v^*)}{\partial x} \Big|_{x=x^*} = 0$$

$$\circ \text{ 即 } \nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^p v_i^* \nabla h_i(x^*) = 0$$



稳定性



$$\begin{aligned} \square & \inf_x \{f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p v_i^* h_i(x)\} \\ \square & = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p v_i^* h_i(x^*) \end{aligned}$$

$$\square L(x^*, \lambda^*, v^*) = \inf_x L(x, \lambda^*, v^*)$$

$$\square \text{ 则 } \frac{\partial L(x, \lambda^*, v^*)}{\partial x} \Big|_{x=x^*} = 0$$

$$\circ \text{ 即 } \nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^p v_i^* \nabla h_i(x^*) = 0$$

□ 稳定性条件



KKT条件





KKT条件



- 一般可微优化问题，对偶间隙为**0**须满足**KKT**条件



KKT条件



- 一般可微优化问题，对偶间隙为0须满足**KKT**条件
- **Karush-Kuhn-Tucker (KKT)** 条件包含如下四个条件：



KKT条件



- 一般可微优化问题，对偶间隙为**0**须满足**KKT**条件
- **Karush-Kuhn-Tucker (KKT)** 条件包含如下四个条件：
 - 原始可行性：

$$f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$$



KKT条件



- 一般可微优化问题，对偶间隙为**0**须满足**KKT**条件
- **Karush-Kuhn-Tucker (KKT)** 条件包含如下四个条件：
 - 原始可行性：
$$f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$$
 - 对偶可行性：
$$\lambda \succeq 0$$



KKT条件



- 一般可微优化问题，对偶间隙为0须满足**KKT**条件
- **Karush-Kuhn-Tucker (KKT)** 条件包含如下四个条件：

- 原始可行性：

$$f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$$

- 对偶可行性：

$$\lambda \succeq 0$$

- 互补松弛性：

$$\lambda_i f_i(x) = 0, i = 1, \dots, m$$



KKT条件



- 一般可微优化问题，对偶间隙为0须满足**KKT**条件
- **Karush-Kuhn-Tucker (KKT)** 条件包含如下四个条件：

- 原始可行性：

$$f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$$

- 对偶可行性：

$$\lambda \succeq 0$$

- 互补松弛性：

$$\lambda_i f_i(x) = 0, i = 1, \dots, m$$

- **Lagrangian**梯度的定常方程式（稳定性条件）：

$$\nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + \sum_{i=1}^p \nu_i \nabla h_i(x) = 0$$



凸问题的**KKT**条件





凸问题的KKT条件



□ 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件
- 证明：



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件
- 证明：
 - $f_i(\tilde{x}) \leq 0, h_i(\tilde{x}) = 0, \tilde{\lambda} \geq 0$



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件
- 证明：
 - $f_i(\tilde{x}) \leq 0, h_i(\tilde{x}) = 0, \tilde{\lambda} \geq 0$
 - $L(x, \tilde{\lambda}, \tilde{v}) = f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \sum_{i=1}^p \tilde{v}_i h_i(x)$ 为凸函数



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件
- 证明：
 - $f_i(\tilde{x}) \leq 0, h_i(\tilde{x}) = 0, \tilde{\lambda} \geq 0$
 - $L(x, \tilde{\lambda}, \tilde{v}) = f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \sum_{i=1}^p \tilde{v}_i h_i(x)$ 为凸函数
 - $\frac{\partial L(x, \tilde{\lambda}, \tilde{v})}{\partial x} \Big|_{x=\tilde{x}} = 0$ ，则 \tilde{x} 为极小值点



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件
- 证明：
 - $f_i(\tilde{x}) \leq 0, h_i(\tilde{x}) = 0, \tilde{\lambda} \geq 0$
 - $L(x, \tilde{\lambda}, \tilde{v}) = f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \sum_{i=1}^p \tilde{v}_i h_i(x)$ 为凸函数
 - $\frac{\partial L(x, \tilde{\lambda}, \tilde{v})}{\partial x} \Big|_{x=\tilde{x}} = 0$ ，则 \tilde{x} 为极小值点
 - $g(\tilde{\lambda}, \tilde{v}) = \inf_x L(x, \tilde{\lambda}, \tilde{v}) = L(\tilde{x}, \tilde{\lambda}, \tilde{v})$



凸问题的KKT条件



- 若 $\tilde{x}, \tilde{\lambda}, \tilde{v}$ 满足凸问题的KKT条件，则是最优解
 - KKT条件为可微凸问题对偶间隙为0的充要条件

□ 证明：

- $f_i(\tilde{x}) \leq 0, h_i(\tilde{x}) = 0, \tilde{\lambda} \geq 0$

- $L(x, \tilde{\lambda}, \tilde{v}) = f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \sum_{i=1}^p \tilde{v}_i h_i(x)$ 为凸函数

- $\frac{\partial L(x, \tilde{\lambda}, \tilde{v})}{\partial x} \Big|_{x=\tilde{x}} = 0$ ，则 \tilde{x} 为极小值点

- $g(\tilde{\lambda}, \tilde{v}) = \inf_x L(x, \tilde{\lambda}, \tilde{v}) = L(\tilde{x}, \tilde{\lambda}, \tilde{v})$

- $= f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) + \sum_{i=1}^p \tilde{v}_i h_i(\tilde{x}) = f_0(\tilde{x})$



例





例



$$\begin{aligned} \min & \frac{1}{2} x^T P x + q^T x + r, P \in S_+^n \\ \text{s.t.} & Ax = b \end{aligned}$$



例



- $$\min \frac{1}{2} x^T P x + q^T x + r, P \in S_+^n$$
- **s.t.** $Ax = b$
- **KKT条件:**



例



- $\min \frac{1}{2} x^T P x + q^T x + r, P \in S_+^n$
- **s.t.** $Ax = b$
- **KKT条件:**
 - $Ax^* = b$



例



- $\min \frac{1}{2}x^T Px + q^T x + r, P \in S_+^n$
- **s.t.** $Ax = b$
- **KKT条件:**
 - $Ax^* = b$
 - $\frac{\partial}{\partial x} \left\{ \frac{1}{2}x^T Px + q^T x + r + (Ax - b)^T v^* \right\} \Big|_{x=x^*} = 0$
 $\iff Px^* + q + A^T v^* = 0$



例



- $\min \frac{1}{2}x^T Px + q^T x + r, P \in S_+^n$
- **s.t.** $Ax = b$
- **KKT条件:**
 - $Ax^* = b$
 - $\frac{\partial}{\partial x} \left\{ \frac{1}{2}x^T Px + q^T x + r + (Ax - b)^T v^* \right\} \Big|_{x=x^*} = 0$
 $\iff Px^* + q + A^T v^* = 0$
- $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$



例：注水





例：注水



$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^n \log(x_i + \alpha_i) \\ \text{subject to} & x \succeq 0, \quad \mathbf{1}^T x = 1 \end{array}$$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

○ $x^* \geq 0, \mathbf{1}^T x^* = 1$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$
- $\lambda^* \geq 0$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$
- $\lambda^* \geq 0$
- $\lambda_i^* x_i^* = 0, i = 1, \dots, n$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$
- $\lambda^* \geq 0$
- $\lambda_i^* x_i^* = 0, i = 1, \dots, n$
- $-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$

- $\lambda^* \geq 0$

- $\lambda_i^* x_i^* = 0, i = 1, \dots, n$

- $-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$

- 则 $\lambda_i^* = \nu^* - \frac{1}{\alpha_i + x_i^*}$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$

- $\lambda^* \geq 0$

- $\lambda_i^* x_i^* = 0, i = 1, \dots, n$

- $-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$

- 则 $\lambda_i^* = \nu^* - \frac{1}{\alpha_i + x_i^*}$

- $$\begin{cases} x_i^* \left(\nu^* - \frac{1}{\alpha_i + x_i^*} \right) = 0 \\ \nu^* \geq \frac{1}{\alpha_i + x_i^*} \end{cases}$$



例：注水



$$\begin{aligned} & \text{minimize} && -\sum_{i=1}^n \log(x_i + \alpha_i) \\ & \text{subject to} && x \succeq 0, \quad \mathbf{1}^T x = 1 \end{aligned}$$

□ KKT条件

- $x^* \geq 0, \mathbf{1}^T x^* = 1$

- $\lambda^* \geq 0$

- $\lambda_i^* x_i^* = 0, i = 1, \dots, n$

- $-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$

- 则 $\lambda_i^* = \nu^* - \frac{1}{\alpha_i + x_i^*}$

$$\begin{cases} \square & x_i^* \left(\nu^* - \frac{1}{\alpha_i + x_i^*} \right) = 0 \\ & \nu^* \geq \frac{1}{\alpha_i + x_i^*} \\ & \nu^* \geq 1/\alpha_i \rightarrow x_i^* = 0 \\ & \nu^* < 1/\alpha_i \rightarrow x_i^* > 0 \\ \square & x_i^* = 1/\nu^* - \alpha_i \end{cases}$$



例：注水



□ KKT条件

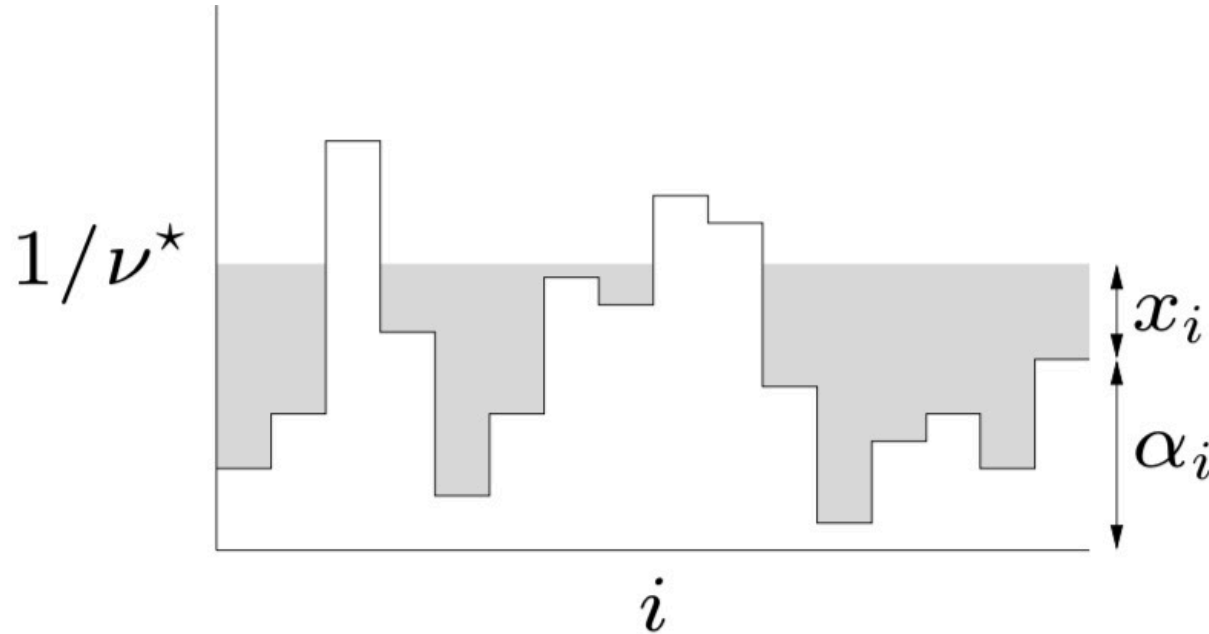
○ $x^* \geq 0, 1^T x^* = 1$

○ $\lambda^* \geq 0$

○ $\lambda_i^* x_i^* = 0, i = 1, \dots$

○ $\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$

□ 则 $\lambda_i^* = \nu^* - \frac{1}{\alpha_i + x_i^*}$



□
$$\begin{cases} \nu^* \geq 1/\alpha_i \rightarrow x_i^* = 0 \\ \nu^* < 1/\alpha_i \rightarrow x_i^* > 0 \\ x_i^* = 1/\nu^* - \alpha_i \end{cases}$$



例：注水



□ KKT条件

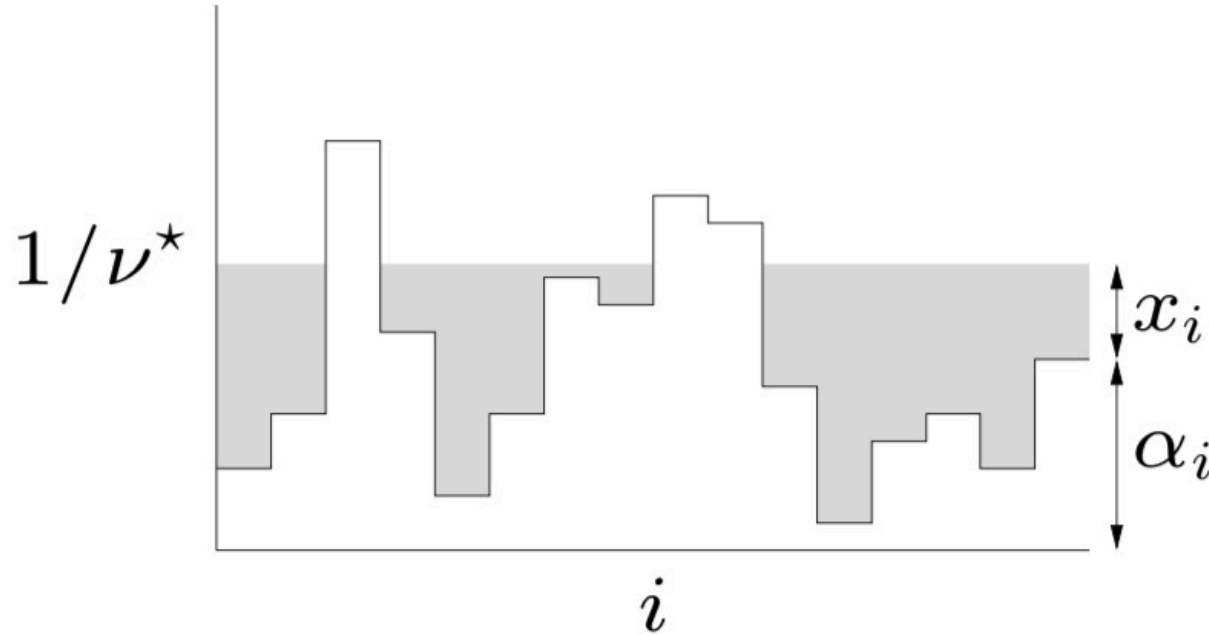
○ $x^* \geq 0, 1^T x^* = 1$

○ $\lambda^* \geq 0$

○ $\lambda_i^* x_i^* = 0, i = 1, \dots$

○ $\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0$

□ 则 $\lambda_i^* = \nu^* - \frac{1}{\alpha_i + x_i^*}$



□
$$\begin{cases} \nu^* \geq 1/\alpha_i \rightarrow x_i^* = 0 \\ \nu^* < 1/\alpha_i \rightarrow x_i^* > 0 \\ x_i^* = 1/\nu^* - \alpha_i \end{cases}$$



例





例



$$\square \min f_0(x) \text{ s.t. } x \geq 0$$



例



$$\square \min f_0(x) \text{ s.t. } x \geq 0$$

$$x \geq 0$$

$$\circ \quad \nabla f_0(x) \geq 0$$

$$x_i (\nabla f_0(x))_i = 0$$



例



$$\square \min f_0(x) \text{ s.t. } x \geq 0$$

$$x \geq 0$$

$$\circ \quad \nabla f_0(x) \geq 0$$

$$x_i (\nabla f_0(x))_i = 0$$

\square KKT条件



例



$$\square \min f_0(x) \text{ s.t. } x \geq 0$$

$$x \geq 0$$

$$\circ \quad \nabla f_0(x) \geq 0$$

$$x_i (\nabla f_0(x))_i = 0$$

\square KKT条件

$$x^* \geq 0$$

$$\lambda^* \geq 0$$

$$\square \quad \lambda_i^* (-x_i^*) = 0$$

$$\nabla f_0(x^*) + (-\lambda^*) = 0$$



对偶及问题重形式化





对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶



对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶
- 当对偶问题很难求解，重形式化原问题是一种有效的办法



对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶
- 当对偶问题很难求解，重形式化原问题是一种有效的办法
- 常用的形式化



对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶
- 当对偶问题很难求解，重形式化原问题是一种有效的办法
- 常用的形式化
 - 引入新的变量以及相应的等式约束



对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶
- 当对偶问题很难求解，重形式化原问题是一种有效的办法
- 常用的形式化
 - 引入新的变量以及相应的等式约束
 - 将显式约束隐式表达，反之亦然



对偶及问题重形式化



- 某一问题的等价形式可产生非常不同的对偶
- 当对偶问题很难求解，重形式化原问题是一种有效的办法
- 常用的形式化
 - 引入新的变量以及相应的等式约束
 - 将显式约束隐式表达，反之亦然
 - 变换目标或约束函数变换



引入新变量和等式约束





引入新变量和等式约束



minimize $f_0(Ax + b)$



引入新变量和等式约束



minimize $f_0(Ax + b)$

□ 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$



引入新变量和等式约束



minimize $f_0(Ax + b)$

- 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$
- 对偶问题: $\max p^*$



引入新变量和等式约束



minimize $f_0(Ax + b)$

- 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$
- 对偶问题: $\max p^*$
 - 有对偶性, 但无意义



引入新变量和等式约束



minimize $f_0(Ax + b)$

- 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$
- 对偶问题: $\max p^*$
 - 有对偶性, 但无意义
- 重形式化问题和对偶



引入新变量和等式约束



$$\text{minimize } f_0(Ax + b)$$

□ 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$

□ 对偶问题: $\max p^*$

○ 有对偶性, 但无意义

□ 重形式化问题和对偶

$$\begin{aligned} &\text{minimize } f_0(y) \\ &\text{subject to } Ax + b - y = 0 \end{aligned}$$



引入新变量和等式约束



$$\text{minimize } f_0(Ax + b)$$

□ 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$

□ 对偶问题: $\max p^*$

○ 有对偶性, 但无意义

□ 重形式化问题和对偶

$$\begin{aligned} &\text{minimize } f_0(y) \\ &\text{subject to } Ax + b - y = 0 \end{aligned}$$

□ 对偶函数 $g(\nu) = \inf_{x,y} (f_0(y) - \nu^T y + \nu^T Ax + b^T \nu)$

$$= \begin{cases} -f_0^*(\nu) + b^T \nu & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$



引入新变量和等式约束



$$\text{minimize } f_0(Ax + b)$$

□ 对偶函数为常函数: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^*$

□ 对偶问题: $\max p^*$

○ 有对偶性, 但无意义

□ 重形式化问题和对偶

$$\begin{aligned} &\text{minimize } f_0(y) \\ &\text{subject to } Ax + b - y = 0 \end{aligned}$$

$$\begin{aligned} &\text{maximize } b^T \nu - f_0^*(\nu) \\ &\text{subject to } A^T \nu = 0 \end{aligned}$$

□ 对偶函数 $g(\nu) = \inf_{x,y} (f_0(y) - \nu^T y + \nu^T Ax + b^T \nu)$

$$= \begin{cases} -f_0^*(\nu) + b^T \nu & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$



范数逼近问题





范数逼近问题



$$\text{minimize } \|Ax - b\|$$



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为

$$g(\nu) = \inf_{x,y} (\|y\| + \nu^T y - \nu^T Ax + b^T \nu)$$



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为

$$\begin{aligned} g(\nu) &= \inf_{x,y} (\|y\| + \nu^T y - \nu^T Ax + b^T \nu) \\ &= \begin{cases} b^T \nu + \inf_y (\|y\| + \nu^T y) & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为

$$\begin{aligned} g(\nu) &= \inf_{x,y} (\|y\| + \nu^T y - \nu^T Ax + b^T \nu) \\ &= \begin{cases} b^T \nu + \inf_y (\|y\| + \nu^T y) & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} b^T \nu & A^T \nu = 0, \quad \|\nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为

$$\begin{aligned} g(\nu) &= \inf_{x,y} (\|y\| + \nu^T y - \nu^T Ax + b^T \nu) \\ &= \begin{cases} b^T \nu + \inf_y (\|y\| + \nu^T y) & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} b^T \nu & A^T \nu = 0, \quad \|\nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

□ 范数逼近问题的对偶：



范数逼近问题



$$\text{minimize } \|Ax - b\|$$

$$\begin{aligned} &\text{minimize } \|y\| \\ &\text{subject to } y = Ax - b \end{aligned}$$

□ 对偶问题为

$$\begin{aligned} g(\nu) &= \inf_{x,y} (\|y\| + \nu^T y - \nu^T Ax + b^T \nu) \\ &= \begin{cases} b^T \nu + \inf_y (\|y\| + \nu^T y) & A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} b^T \nu & A^T \nu = 0, \quad \|\nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

□ 范数逼近问题的对偶：

$$\begin{aligned} &\text{maximize } b^T \nu \\ &\text{subject to } A^T \nu = 0, \quad \|\nu\|_* \leq 1 \end{aligned}$$



隐式约束





隐式约束



- 带框约束的线性规划



隐式约束



□ 带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$



隐式约束



□ 带框约束的线性规划

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b \\ & && -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} & \text{maximize} && -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ & \text{subject to} && c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & && \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$



隐式约束



□ 带框约束的线性规划

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && Ax = b \\ & && -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} & \text{maximize} && -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ & \text{subject to} && c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & && \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

□ 将框约束隐式表达



隐式约束



□ 带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} \text{maximize} \quad & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} \quad & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

□ 将框约束隐式表达

$$\begin{aligned} \text{minimize} \quad & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} \quad & Ax = b \end{aligned}$$



隐式约束



□ 带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} \text{maximize} \quad & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} \quad & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

□ 将框约束隐式表达

$$\begin{aligned} \text{minimize} \quad & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} \quad & Ax = b \end{aligned}$$

□ 对偶函数及对偶问题：



隐式约束



带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} \text{maximize} \quad & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} \quad & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

将框约束隐式表达

$$\begin{aligned} \text{minimize} \quad & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} \quad & Ax = b \end{aligned}$$

对偶函数及对偶问题:

$$g(\nu) = \inf_{-\mathbf{1} \preceq x \preceq \mathbf{1}} (c^T x + \nu^T (Ax - b))$$



隐式约束



带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} \text{maximize} \quad & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} \quad & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

将框约束隐式表达

$$\begin{aligned} \text{minimize} \quad & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} \quad & Ax = b \end{aligned}$$

对偶函数及对偶问题:

$$\begin{aligned} g(\nu) &= \inf_{-\mathbf{1} \preceq x \preceq \mathbf{1}} (c^T x + \nu^T (Ax - b)) \\ &= -b^T \nu - \|A^T \nu + c\|_1 \end{aligned}$$



隐式约束



带框约束的线性规划

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{aligned}$$

$$\begin{aligned} \text{maximize} \quad & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} \quad & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0 \end{aligned}$$

将框约束隐式表达

$$\begin{aligned} \text{minimize} \quad & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} \quad & Ax = b \end{aligned}$$

对偶函数及对偶问题:

$$\begin{aligned} g(\nu) &= \inf_{-1 \preceq x \preceq 1} (c^T x + \nu^T (Ax - b)) & \text{maximize} \quad & -b^T \nu - \|A^T \nu + c\|_1 \\ &= -b^T \nu - \|A^T \nu + c\|_1 \end{aligned}$$



扰动及灵敏度分析





扰动及灵敏度分析



□ 优化问题及其对偶问题



扰动及灵敏度分析



□ 优化问题及其对偶问题

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

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



扰动及灵敏度分析



□ 优化问题及其对偶问题

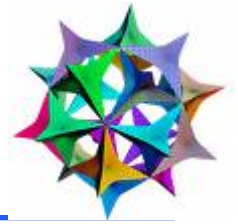
$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

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

□ 扰动问题及其对偶



扰动及灵敏度分析



□ 优化问题及其对偶问题

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ &&& h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

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

□ 扰动问题及其对偶

$$\begin{aligned} \text{min.} & f_0(x) \\ \text{s.t.} & f_i(x) \leq u_i, \quad i = 1, \dots, m \\ & h_i(x) = v_i, \quad i = 1, \dots, p \end{aligned}$$

$$\begin{aligned} \text{max.} & g(\lambda, \nu) - u^T \lambda - v^T \nu \\ \text{s.t.} & \lambda \succeq 0 \end{aligned}$$



扰动及灵敏度分析



□ 优化问题及其对偶问题

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

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

□ 扰动问题及其对偶

$$\begin{aligned} & \text{min.} && f_0(x) \\ & \text{s.t.} && f_i(x) \leq u_i, \quad i = 1, \dots, m \\ & && h_i(x) = v_i, \quad i = 1, \dots, p \end{aligned}$$

$$\begin{aligned} & \text{max.} && g(\lambda, \nu) - u^T \lambda - v^T \nu \\ & \text{s.t.} && \lambda \succeq 0 \end{aligned}$$

□ x 是原问题优化变量， u ， v 为参数



扰动及灵敏度分析



□ 优化问题及其对偶问题

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \\ &&& h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

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

□ 扰动问题及其对偶

$$\begin{aligned} &\text{min.} && f_0(x) \\ &\text{s.t.} && f_i(x) \leq u_i, \quad i = 1, \dots, m \\ &&& h_i(x) = v_i, \quad i = 1, \dots, p \end{aligned}$$

$$\begin{aligned} &\text{max.} && g(\lambda, \nu) - u^T \lambda - v^T \nu \\ &\text{s.t.} && \lambda \succeq 0 \end{aligned}$$

□ x 是原问题优化变量， u ， v 为参数

□ $p^*(u, v)$ 为最优值，为 u 和 v 的函数



扰动问题的性质





扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数



扰动问题的性质



- 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数
- 证明：



扰动问题的性质



- 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数
- 证明：
- $$p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$



扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数

□ 证明：

$$\square p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$

$$\square = \inf_x g(x, u, v)$$



扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数

□ 证明：

$$\square p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$

$$\square \quad \quad \quad = \inf_x g(x, u, v)$$

$$\square g(x, u, v) = f_0(x), \quad \mathbf{dom} \, g = \mathbf{dom} \, f_0 \cap D$$



扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数

□ 证明：

$$\square p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$

$$\square \quad \quad \quad = \inf_x g(x, u, v)$$

$$\square g(x, u, v) = f_0(x), \quad \mathbf{dom} \, g = \mathbf{dom} \, f_0 \cap D$$

$$\circ D = \{x \mid f_i(x) \leq 0, h_i(x) = 0\}$$



扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数

□ 证明：

$$\square p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$

$$\square = \inf_x g(x, u, v)$$

$$\square g(x, u, v) = f_0(x), \mathbf{dom} g = \mathbf{dom} f_0 \cap D$$

$$\circ D = \{x \mid f_i(x) \leq 0, h_i(x) = 0\}$$

○ $\mathbf{dom} g$ 为凸集



扰动问题的性质



□ 若原问题为凸，则 $p^*(u, v)$ 为 (u, v) 的凸函数

□ 证明：

$$\square p^*(u, v) = \inf_x \{f_0(x) \mid f_i(x) \leq 0, h_i(x) = 0\}$$

$$\square = \inf_x g(x, u, v)$$

$$\square g(x, u, v) = f_0(x), \text{ dom } g = \text{dom } f_0 \cap D$$

$$\circ D = \{x \mid f_i(x) \leq 0, h_i(x) = 0\}$$

○ $\text{dom } g$ 为凸集

○ 为 (x, u, v) 的凸函数



扰动问题的性质





扰动问题的性质



- 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 证：设 \tilde{x} 为干扰问题的最优解



扰动问题的性质



- 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则
 - $p^*(u, v) \geq p^*(0, 0) - \lambda^{*T} u - w^{*T} v$
- 证：设 \tilde{x} 为干扰问题的最优解
- 则 $f_i(\tilde{x}) \leq u_i, h_i(\tilde{x}) = u_i$



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 证：设 \tilde{x} 为干扰问题的最优解

□ 则 $f_i(\tilde{x}) \leq u_i, h_i(\tilde{x}) = u_i$

□ $p^*(0,0) = g(\lambda^*, w^*)$



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 证：设 \tilde{x} 为干扰问题的最优解

□ 则 $f_i(\tilde{x}) \leq u_i, h_i(\tilde{x}) = u_i$

□ $p^*(0,0) = g(\lambda^*, w^*)$

□
$$\leq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^{*T} f_i(\tilde{x}) + \sum_{i=1}^n w_i^{*T} h_i(\tilde{x})$$



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 证：设 \tilde{x} 为干扰问题的最优解

□ 则 $f_i(\tilde{x}) \leq u_i, h_i(\tilde{x}) = u_i$

□ $p^*(0,0) = g(\lambda^*, w^*)$

□
$$\leq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^{*T} f_i(\tilde{x}) + \sum_{i=1}^n w_i^{*T} h_i(\tilde{x})$$

□
$$\leq f_0(\tilde{x}) + \lambda^{*T}u + w^{*T}v$$



扰动问题的性质



□ 若原问题为凸，且对偶间隙为零， λ^*, w^* 为原问题对偶最优解，则

○ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 证：设 \tilde{x} 为干扰问题的最优解

□ 则 $f_i(\tilde{x}) \leq u_i, h_i(\tilde{x}) = u_i$

□ $p^*(0,0) = g(\lambda^*, w^*)$

□
$$\leq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^{*T} f_i(\tilde{x}) + \sum_{i=1}^n w_i^{*T} h_i(\tilde{x})$$

□
$$\leq f_0(\tilde{x}) + \lambda^{*T}u + w^{*T}v$$

□
$$= p^*(u, v) + \lambda^{*T}u + w^{*T}v$$



灵敏度分析





灵敏度分析



$$\square p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$$



灵敏度分析



□ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加



灵敏度分析



□ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加

□ 若 λ_i^* 比较小：若我们放松约束 $i\{u_i > 0\}$ ，最优值 p^* 不会减小太多



灵敏度分析



□ $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$

□ 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加

□ 若 λ_i^* 比较小：若我们放松约束 $i\{u_i > 0\}$ ，最优值 p^* 不会减小太多

□ 若 w_i^* 较大且为正，我们选择 $v_i < 0$ ，最优值 p^* 必然大幅增加



灵敏度分析



- $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$
- 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加
- 若 λ_i^* 比较小：若我们放松约束 $i\{u_i > 0\}$ ，最优值 p^* 不会减小太多
- 若 w_i^* 较大且为正，我们选择 $v_i < 0$ ，最优值 p^* 必然大幅增加
- 若 w_i^* 较大且为负，我们选择 $v_i > 0$ ，最优值 p^* 必然大幅增加



灵敏度分析



- $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$
- 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加
- 若 λ_i^* 比较小：若我们放松约束 $i\{u_i > 0\}$ ，最优值 p^* 不会减小太多
- 若 w_i^* 较大且为正，我们选择 $v_i < 0$ ，最优值 p^* 必然大幅增加
- 若 w_i^* 较大且为负，我们选择 $v_i > 0$ ，最优值 p^* 必然大幅增加
- 若 w_i^* 较小且为正，我们选择 $v_i > 0$ ，最优值 p^* 不会减小太多



灵敏度分析



- $p^*(u, v) \geq p^*(0,0) - \lambda^{*T}u - w^{*T}v$
- 若 λ_i^* 比较大：若我们加强约束 $i\{u_i < 0\}$ ，最优值 p^* 大幅增加
- 若 λ_i^* 比较小：若我们放松约束 $i\{u_i > 0\}$ ，最优值 p^* 不会减小太多
- 若 w_i^* 较大且为正，我们选择 $v_i < 0$ ，最优值 p^* 必然大幅增加
- 若 w_i^* 较大且为负，我们选择 $v_i > 0$ ，最优值 p^* 必然大幅增加
- 若 w_i^* 较小且为正，我们选择 $v_i > 0$ ，最优值 p^* 不会减小太多
- 若 w_i^* 较小且为负，我们选择 $v_i < 0$ ，最优值 p^* 不会减小太多



局部灵敏度分析





局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

□ 证明：根据全局灵敏度分析结果



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0, 0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i}$$

□ 证明：根据全局灵敏度分析结果

$$\frac{\partial p^*(0, 0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} \geq -\lambda_i^*$$

$$\frac{\partial p^*(0, 0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} \leq -\lambda_i^*$$



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

□ 证明：根据全局灵敏度分析结果

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^*$$

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \leq -\lambda_i^*$$

□ 若仅有一个不等式约束，



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

□ 证明：根据全局灵敏度分析结果

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^*$$

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \leq -\lambda_i^*$$

□ 若仅有一个不等式约束，

□ 最优值 $p^*(u)$ 的图像：



局部灵敏度分析



□ 若原问题为凸且强对偶，在 $p^*(u, v)$ 在 $(0,0)$ 可微，则有

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

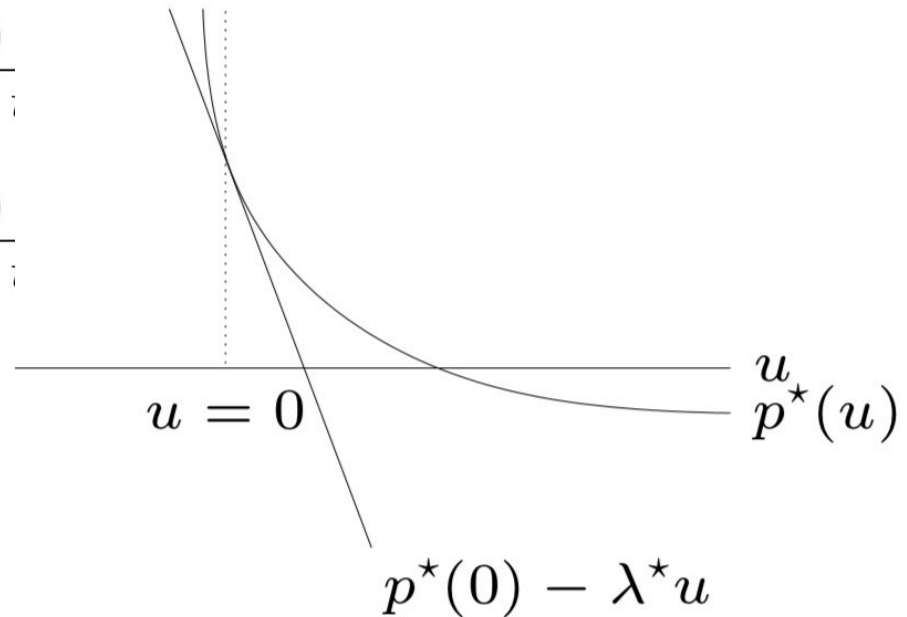
□ 证明：根据全局灵敏度分析结果

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t}$$

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t}$$

□ 若仅有一个不等式约束，

□ 最优值 $p^*(u)$ 的图像：





基于广义不等式的问题





基于广义不等式的问题



$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

□ 定义和标量形式是一致的



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

- 定义和标量形式是一致的
- $f_i(x) \preceq_{K_i} 0$ 的 **Lagrange** 乘子为向量 $\lambda_i \in \mathbf{R}^{k_i}$



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

- 定义和标量形式是一致的
- $f_i(x) \preceq_{K_i} 0$ 的**Lagrange**乘子为向量 $\lambda_i \in \mathbf{R}^{k_i}$
- **Lagrangian**函数 $L : \mathbf{R}^n \times \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$ 为



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

- 定义和标量形式是一致的
- $f_i(x) \preceq_{K_i} 0$ 的**Lagrange**乘子为向量 $\lambda_i \in \mathbf{R}^{k_i}$
- **Lagrangian**函数 $L : \mathbf{R}^n \times \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$
为
$$L(x, \lambda_1, \dots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i^T f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

- 定义和标量形式是一致的
- $f_i(x) \preceq_{K_i} 0$ 的**Lagrange**乘子为向量 $\lambda_i \in \mathbf{R}^{k_i}$
- **Lagrangian**函数 $L : \mathbf{R}^n \times \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$
为
$$L(x, \lambda_1, \dots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i^T f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$
- 对偶函数 $g : \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$



基于广义不等式的问题



$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & && h_i(x) = 0, \quad i = 1, \dots, p \end{aligned}$$

- 定义和标量形式是一致的
- $f_i(x) \preceq_{K_i} 0$ 的 **Lagrange** 乘子为向量 $\lambda_i \in \mathbf{R}^{k_i}$
- **Lagrangian** 函数 $L : \mathbf{R}^n \times \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$
为
$$L(x, \lambda_1, \dots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i^T f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$
- 对偶函数 $g : \mathbf{R}^{k_1} \times \dots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \rightarrow \mathbf{R}$
$$g(\lambda_1, \dots, \lambda_m, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu)$$



下界性质





下界性质



□ 若 $\lambda_i \succeq_{K_i^*} 0$, 则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$, 则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明: 若 \tilde{x} 为可行的, 且 $\lambda \succeq_{K_i^*} 0$ 则有



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$ ，则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明：若 \tilde{x} 为可行的，且 $\lambda \succeq_{K_i^*} 0$ 则有

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^p \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu) \\ &= g(\lambda_1, \dots, \lambda_m, \nu) \end{aligned}$$



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$ ，则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明：若 \tilde{x} 为可行的，且 $\lambda \succeq_{K_i^*} 0$ 则有

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^p \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu) \\ &= g(\lambda_1, \dots, \lambda_m, \nu) \end{aligned}$$

- 对所有可行解， \tilde{x} 为最优值，有 $p^* \geq g(\lambda_1, \dots, \lambda_m, \nu)$



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$ ，则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明：若 \tilde{x} 为可行的，且 $\lambda \succeq_{K_i^*} 0$ 则有

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^p \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu) \\ &= g(\lambda_1, \dots, \lambda_m, \nu) \end{aligned}$$

- 对所有可行解， \tilde{x} 为最优值，有 $p^* \geq g(\lambda_1, \dots, \lambda_m, \nu)$
- 对偶问题



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$ ，则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明：若 \tilde{x} 为可行的，且 $\lambda \succeq_{K_i^*} 0$ 则有

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^p \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu) \\ &= g(\lambda_1, \dots, \lambda_m, \nu) \end{aligned}$$

- 对所有可行解， \tilde{x} 为最优值，有 $p^* \geq g(\lambda_1, \dots, \lambda_m, \nu)$
- 对偶问题

$$\begin{aligned} &\text{maximize} && g(\lambda_1, \dots, \lambda_m, \nu) \\ &\text{subject to} && \lambda_i \succeq_{K_i^*} 0, \quad i = 1, \dots, m \end{aligned}$$



下界性质



- 若 $\lambda_i \succeq_{K_i^*} 0$ ，则 $g(\lambda_1, \dots, \lambda_m, \nu) \leq p^*$
- 证明：若 \tilde{x} 为可行的，且 $\lambda \succeq_{K_i^*} 0$ 则有

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^m \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^p \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \dots, \lambda_m, \nu) \\ &= g(\lambda_1, \dots, \lambda_m, \nu) \end{aligned}$$

- 对所有可行解， \tilde{x} 为最优值，有 $p^* \geq g(\lambda_1, \dots, \lambda_m, \nu)$

- 对偶问题

$$\begin{aligned} &\text{maximize} && g(\lambda_1, \dots, \lambda_m, \nu) \\ &\text{subject to} && \lambda_i \succeq_{K_i^*} 0, \quad i = 1, \dots, m \end{aligned}$$

- 弱对偶： $p^* \geq d^*$ ，强对偶： $p^* = d^*$



半定规划





半定规划



□ 原问题

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n \preceq G \end{array}$$



半定规划



- 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$
- **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$



半定规划



- 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$
- **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$
- **Lagrangian**函数



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$

□ 对偶函数



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$

□ 对偶函数

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(GZ) & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ -\infty & \text{otherwise} \end{cases}$$



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$

□ 对偶函数

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(GZ) & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ -\infty & \text{otherwise} \end{cases}$$

□ 半定规划的对偶



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$

□ 对偶函数

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(GZ) & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ -\infty & \text{otherwise} \end{cases}$$

□ 半定规划的对偶

$$\begin{aligned} & \text{maximize} && -\mathbf{tr}(GZ) \\ & \text{subject to} && Z \succeq 0, \quad \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \end{aligned}$$



半定规划



□ 原问题 minimize $c^T x$
subject to $x_1 F_1 + \cdots + x_n F_n \preceq G$

□ **Lagrange**乘子为矩阵 $Z \in \mathbf{S}^k$

□ **Lagrangian**函数

$$L(x, Z) = c^T x + \mathbf{tr}(Z(x_1 F_1 + \cdots + x_n F_n - G))$$

□ 对偶函数

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(GZ) & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \\ -\infty & \text{otherwise} \end{cases}$$

□ 半定规划的对偶

$$\begin{aligned} & \text{maximize} && -\mathbf{tr}(GZ) \\ & \text{subject to} && Z \succeq 0, \quad \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \dots, n \end{aligned}$$

□ 若半定规划是严格可行的, 则 $p^* = d^*$