$\min f_0(x)$

s.t. $f_i(x) \leq 0, \ i=1,\cdots,m$

$h_i(x)=0, \ i=1,\cdots,p$

$x \in R^n, \ D=\bigcap_{i=1}^{m} \mathrm{dom} f_i \cap \bigcap_{i=1}^{p} \mathrm{dom} h_i$

p^* optimal value

1. n个基本概念

对偶变量

① 拉格朗日函数：$L(x,\lambda,v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{n} v_i h_i(x)$

② 拉格朗日对偶函数：$g(\lambda,v) = \inf_{x\in D} L(x,\lambda,v)$

性质：1) 必为凹函数（L.分段线性凸函数，inf 则必为凹函数）

2) $\forall \lambda \geq 0, \ \forall v, \ g(\lambda,v) \leq p^*$.（为原问题最优值提供下界）

证明：设 x^* 是原问题的最优解，则必可行

则 $f_i(x^*) \leq 0, \ h_i(x^*)=0$

当 $\forall \lambda \geq 0, \ \forall v$，有 $\sum_{i=1}^{m} \lambda f_i(x^*) + \sum_{i=1}^{p} v h_i(x^*) \leq 0$

$L(x^*,\lambda,v) = f_0(x^*) + \sum_{i=1}^{m} \lambda f_i(x^*) + \sum_{i=1}^{p} v h_i(x^*) \leq p^*$

作用：求解凸问题 $\sup g(\lambda,v)$. 能找到最好的下界.

例：$\min x^T x$, s.t. $Ax=b$, $x \in R^n, \ b \in R^p, \ A \in R^{p\times n}$（二次规划）

$\Rightarrow L(x,v) = x^T x + v^T(Ax-b)$

$\Rightarrow g(v) = \inf_{x\in D} L(x,v) = \inf x^T x + v^T A x - v^T b$

$= 2x + A^T v = 0 \Rightarrow x = -\frac{A^T v}{2}$

半负定矩阵 → 凹函数

代入，得 $g(v) = \frac{v^T A A^T v}{4} - \frac{v^T A A^T v}{2} - v^T b = -\frac{1}{4} v^T A A^T v - b^T v$

例：$\min \ c^T x$（线性规划）

s.t. $Ax=b$ $\Rightarrow$ $Ax-b=0$

$x \geq 0$ $\qquad -x \leq 0$

$$\Rightarrow L(x,\lambda,v) = c^Tx - \lambda^Tx + v^T(Ax-b)$$

$$= -b^Tv + (c+A^Tv-\lambda)^Tx$$

$$\Rightarrow g(\lambda,v) = \inf_x L(x,\lambda,v)$$

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

既凸又凹 扩展 ⇒ 凹函数

例：$\min \ x^TWx$

s.t. $x_i = \pm 1, \ i=1,\cdots,m$ 非凸约束（离散形式）

$x_i^2 - 1 = 0, \ i=1,\cdots,m$

$$\Rightarrow L(x,v) = x^TWx + \sum_{i=1}^{n} v_i(x_i^2-1)$$

$$= x^T(W+\mathrm{diag}(v))x - 1^Tv$$

$$\Rightarrow g(v) = \inf_x x^T(W+\mathrm{diag}(v))x - 1^Tv$$

$$= \begin{cases} -1^Tv, & W+\mathrm{diag}(v)\succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

凸集定义证明

③ 函数的共轭 f^* 是 $f: R^n \to R$ 的共轭，若 $f^*(y) = \sup_{x\in \mathrm{dom}f} (y^Tx - f(x))$

例：$\min f(x)$

s.t. $x=0$

$\Rightarrow L(x,v) = f(x) + v^Tx$，$\mathrm{dom}L = \mathrm{dom}f \times R^n$

$$\Rightarrow g(v) = \inf_{x\in \mathrm{dom}f} (f(x)+v^Tx) = -\sup_{x\in\mathrm{dom}f}(-v^Tx - f(x)) = -f^*(-v)$$

例：$\min \ f_0(x)$

$$\text{s.t. } Ax \preceq b$$
$$cx = d$$

$$\Rightarrow L(x,\lambda,v) = f_0(x) + \lambda^T(Ax-b) + V^T(cx-d)$$
$$= f_0(X) + (\lambda^T A + v^T c)X - \lambda^T b - v^T d$$
$$\Rightarrow g(\lambda,v) = \inf_{X\in \text{dom}f} L(X,\lambda,V) = -f_0^*(-(\lambda^T A + V^T c)X - \lambda^T b - v^T d)$$

④原问题和对偶问题

$$(D)\begin{cases} \max\ g(\lambda,v) \\ \text{s.t. } \lambda \succeq 0 \end{cases} \qquad (P)\begin{cases} \min\ f_0(x) \\ \text{s.t. } f_i(x)\le 0,\ i=1,\cdots,m \\ h_i(x)=0,\ i=1,\cdots,p \end{cases}$$

性质：1) $d^* \le p^*$

2) λ^*, v^* optimal larange multiplier

例：

$$\boxed{\begin{array}{l} \min\ c^T x \quad P_1 \\ \text{s.t. } Ax = b \\ \qquad x \succeq 0 \end{array}}$$

n维问题
p维约束

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

$$(D) \Rightarrow \max\ g(\lambda,v) \quad \text{s.t. } \lambda \succeq 0 \iff \max\ -b^T v \quad \text{s.t. } \lambda \succeq 0,\ A^T v - \lambda + c = 0$$

$$\rightarrow \boxed{\begin{array}{l} \min\ b^T v \quad D_1 \\ A^T v + c \succeq 0 \end{array}}$$

P维问题
n维约束

最优解一样，但最优值不一样

例：

$$\boxed{\begin{array}{l} \min\ c^T x \quad P_2 \\ \text{s.t. } Ax \preceq b \end{array}}$$

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

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

(D) $\max \; -b^T\lambda$ P_2

s.t. $A^T\lambda + c = 0$

$\lambda \geq 0$

☆☆☆ 比较问题的"n个约束的n维优化问题"

☆ 比较两个线性规划的例子 ⇒ 对偶的对偶不一定是自身（共轭的共轭不是自身）

什么情况下等价？

2. 对偶问题的性质.

(1) 对偶问题必为凸优化问题（因为对偶函数是凹函数）

(2) $d^* \leq p^*$. 什么时候 $d^* = p^*$?

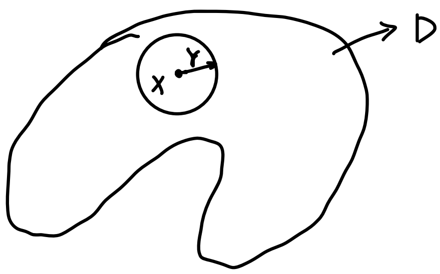
弱对偶 (weak duality) & 强对偶 (strong duality)

$p^* - d^*$ duality gap

定义 D(域) 的 Relative Interior：

$$\text{Relint } D = \{x \in D \mid B(x,r) \cap \text{aff} D \subseteq D, \exists r > 0\}$$

覆盖整个平面



直观看，去掉D的边缘则都成立.

Slater's condition ($p^* = d^*$的充分条件) 不是充要条件.

若有凸问题 $\min f_0(x)$

s.t. $f_i(x) \leq 0 \quad i=1,\cdots,m$

$Ax = b$

没有等于号

当 $\exists x \in \text{relint} D$，使 $f_i(x) < 0$，$i=1,\cdots,m$，$Ax = b$ 满足时，$p^* = d^*$.

A weak Slater's Condition

$cx+d=0$（超平面）

$cx+d\le 0$（半空间、多面体）　（充分条件）

若不等式约束为仿射约束时，只要可行域非空，必有 $p^*=d^*$.

$$\text{relint}\ D = \text{relint}\{\text{dom} f_0 \cap \bigcap_i \text{dom} f_i\}$$

$$\text{relint}\ D = \text{relint}\{\text{dom} f_0\}$$

线性规划问题，若可行，必有 $p^*=d^*$.

例：$\min\ x^T x$ 　 $p^*=d^*$ $\Longleftrightarrow$ 　 $\max\ -\frac{1}{4}v^T A A^T v - b^T v$

s.t. $Ax=b$ 　　 unbounded

例：QCQP　$\min\ \frac{1}{2}x^T P_0 x + q_0^T x + r_0$

s.t. $\frac{1}{2}x^T P_i x + q_i x + r_i$, $i=1,\cdots,m$

$P_0 \in S_{++}^n$, $P_i \in S_+^n$

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

$$= \frac{1}{2}x^T(\overbrace{P_0 + \sum_{i=1}^m \lambda_i P_i}^{P})x + (\overbrace{q_0 + \sum_{i=1}^m \lambda_i q_i}^{q})^T x + (\overbrace{r_0 + \sum_{i=1}^m \lambda_i r_i}^{r})$$

当 $\lambda\ge 0$ 时，$g(\lambda) = \inf_x L(x,\lambda) = -\frac{1}{2}q^T(\lambda)P^{-1}(\lambda)q(\lambda) + r(\lambda)$　凹函数

例：$\max\ -\frac{1}{2}q^T(\lambda)P^{-1}(\lambda)q(\lambda) + r(\lambda)$

s.t. $\lambda \ge 0$

$\exists x \in D = R^n$, $\frac{1}{2}x^T P_i x + q_i^T x + r_i < 0$, $i=1,\cdots,m$，则 $p^*=d^*$.

若 $q_i=0, r_i=0$，$\frac{1}{2}x^T P_i x < 0$，可以找到 x. (P)与(D)的 Duality gap 为 0.

3. 几种解释 $p^* = d^*$

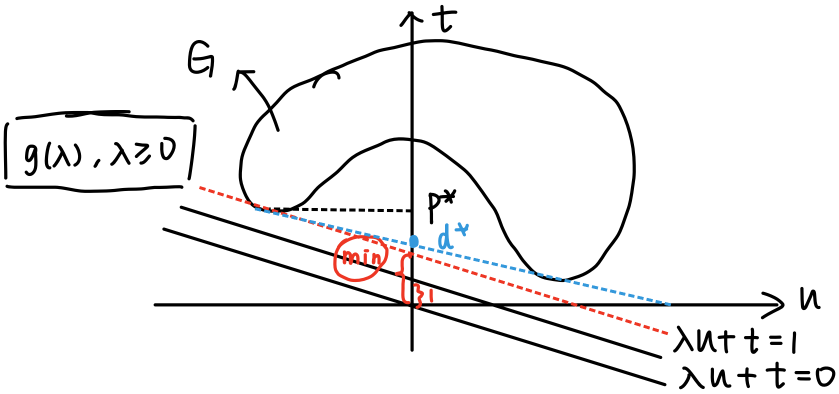
$$\min \ f_0(x)$$
$$\text{s.t.} \ f_i(x) \leq 0, \ i=1,\cdots,m, \ x \in D$$

(1) 几何解释

$$\min \ f_0(x)$$
$$\text{s.t.} \ f_1(x) \leq 0$$

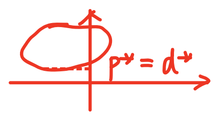
定义 $G = \{(f_1(x), f_0(x)) \mid x \in D\}$, $p^* = \inf\{t \mid (u,t) \in G, u \leq 0\}$

$g(\lambda) = \inf\{\lambda u + t \mid (u,t) \in G\}$ 拉格朗日函数求极小 = 对偶函数



在该问题只有弱对偶，没有强对偶.

在凸集中有强对偶.



d^* (改变 λ，使 $g(\lambda)$ 最大的点)

(2) 鞍点的解释 Saddle point

$\inf\limits_{x \in D} \sup\limits_{\lambda \geq 0} L(x,\lambda) = \sup\limits_{\lambda \geq 0} \inf\limits_{x \in D} L(x,\lambda)$ ，则称有鞍点. (x^*, λ^*) 即为鞍点.

$$d^* = \sup_{\lambda \geq 0} \inf_{x \in D} L(x,\lambda)$$

$p^* = \inf\limits_{x \in D} \sup\limits_{\lambda \geq 0} L(x,\lambda)$, $\left(L(x,\lambda) = f_0(x) + \sum\limits_i \lambda_i f_i(x)\right) \begin{cases} \forall i, 若 f_i(x) \leq 0, & f_0(x) \\ \exists i, 若 f_i(x) > 0, & +\infty \end{cases}$

如果有鞍点，则 $p^* = d^*$.

(3) 多目标优化的解释.

$\min\{f_0(x), f_1(x), \cdots\cdots, f_m(x)\}$, $\{\lambda_i\} = \min \underbrace{\boxed{f_0(x) + \sum\limits_i \tilde{\lambda}_i f_i(x)}}_{L(x,\lambda)}$

$$\min \ f_0(x)$$
$$\text{s.t.} \ f_i(x) \leq 0, \ i=1,\cdots,m$$

首先，给定λ，$\min_{x\in D} L(x,\lambda)$，$\max_{\lambda\succeq 0}$ $\overset{\tilde{x}}{\lambda}$，$\overset{g(\lambda)}{\tilde{\lambda}}$ （不是很理解）

(4) 经济学解释

$\min f_0(x)$ 计划经济

s.t. $f_i(x)\le 0,\ i=1,\cdots,m$

x：产品数量，$-f_0(x)$：利润，$f_0(x)$：损失，$f_i(x)$ 原材料

设原材料可交易，价格 $\lambda_i\ge 0$

$\min_x\ f_0(x)+\sum_{i=1}^m \lambda_i f_i(x)$，记为 $g(\lambda)$ 损失函数

(对偶) 市场经济

$$\boxed{\begin{array}{l}\max\ g(\lambda)\\ \text{s.t. } \lambda\succeq 0\end{array}}$$

最大损失 $d^*=\max_{\lambda\succeq 0} g(\lambda)$

$d^*\le p^*$（可交易原材料的损失小于不可以的损失）

x^*：最优生产方案

λ^*：市场价格应如何调节（对偶最优解）

$d^*=p^*$ 影子价格，供需平衡时的价格.

☆ $\begin{cases} f_i(x^*)<0 \Rightarrow \lambda_i^*=0 \\ \lambda_i^*>0 \Rightarrow f_i(x^*)=0\end{cases}$

(5) 鞍点解释（重新讲）

最小里找最大　　最大里找最小

$f(w,z)$ 在集合 $w\in S_w$，$z\in S_z$ 上，$\sup_{z\in S_z}\inf_{w\in S_w} f(w,z)\le \inf_{w\in S_w}\sup_{z\in S_z} f(w,z)$

$\arg_{w,z}\left\{\sup_{z\in S_z}\inf_{w\in S_w} f(w,z)\right\} = \arg_{w,z}\left\{\inf_{w\in S_w}\sup_{z\in S_z} f(w,z)\right\}$

$(\tilde{w},\tilde{z})$ 使等式成立称为鞍点，$f(\tilde{w},z)\le f(\tilde{w},\tilde{z})\le f(w,\tilde{z})$，$\forall z\in S_z$，$\forall w\in S_w$

$\Big\downarrow$ 推广到优化问题

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

$$\sup_{\lambda\succeq 0}\{L(x,\lambda)\}=\begin{cases} f_0(x), & f_i(x)\le 0,\ i=1,\cdots,m \\ +\infty, & \text{otherwise}\end{cases}$$

$p^* = \min_x \{ f_0(x) \mid f_i(x) \le 0,\ i=1,\cdots,m \} = \inf_x \sup_{\lambda \ge 0} \{ L(x,\lambda) \}$

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

$\Rightarrow p^* \ge d^*$. 且鞍点 $(\hat{x},\hat{\lambda})$ 能使 $p^* = d^*$.

适用于一般问题，不一定凸

证明：鞍点定理（若 $(\hat{x},\hat{\lambda})$ 为 $L(x,\lambda)$ 的鞍点 $\Leftrightarrow$ 强对偶存在，且 $\hat{x},\hat{\lambda}$ 为 Primal 与 Dual 的最优解）& 对偶间隙为 0

[充分性] $\sup_{\lambda \ge 0} \inf_x L(x,\lambda) = \inf_x \sup_{\lambda \ge 0} L(x,\lambda)$ 强对偶存在

$\hat{\lambda} = \arg_\lambda \sup_{\lambda \ge 0} \inf_x L(x,\lambda)$ $\hat{\lambda}$ 是 Dual 最优解

$\tilde{x} = \arg_x \inf_x \sup_{\lambda \ge 0} L(x,\lambda)$ $\hat{x}$ 是 Primal 最优解

[必要性] $f_i(\hat{x}) \le 0,\ i=1,\cdots,m$

$\hat{\lambda} \ge 0$

强对偶存在，则 $f_0(\hat{x}) = g(\hat{\lambda}) = \inf_x \{ f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) \}$

$= f_0(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i f_i(\hat{x}) \overset{=}{\le} f_0(\hat{x})$

所以 $\inf_x L(x,\hat{\lambda}) = L(\hat{x},\hat{\lambda})$

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

$= f_0(\hat{x}) = L(\hat{x},\hat{\lambda})$ 鞍点.

4. KKT 条件（Karush、Kuhn、Tucker）学生：Gale, Minsky, Nash

$\min f_0(x)$

s.t. $f_i(x) \le 0,\ i=1,\cdots,m$

$h_i(x) = 0,\ i=1,\cdots,p$

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 $\max\ g(\lambda,\nu),\quad s.t.\ \lambda \geq 0$

假设：① $p^* = d^*$ ② 所有函数可微（不一定凸） ③已经找到最优解 x^*, λ^*, ν^*.

$$\begin{cases} f_i(x^*) \leq 0,\ i=1,\cdots,m \\ h_i(x^*) = 0,\ i=1,\cdots,p \\ \lambda^* > 0 \end{cases}$$

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

$\inf_x L(x,\lambda^*,\nu^*)$

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

$L(x^*,\lambda^*,\nu^*)$

$$= \leq f_0(x^*)$$

1) $\sum_{i=1}^m \lambda_i^* f_i(x^*) = 0,\quad \lambda_i^* f_i(x^*) = 0,\ \forall i=1,\cdots,m$

$$\begin{cases} \lambda_i^* > 0 \Rightarrow f_i(x^*) = 0 \\ f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0 \end{cases}$$

Complementarity Slackness

互补松弛条件

2) $\left.\dfrac{\partial L(x,\lambda^*,\nu^*)}{\partial x}\right|_{x=x^*} = 0$

(对于非凸问题，必要条件) $\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^p \nu_i^* \nabla h_i(x^*) = 0$ stationary

☆ KKT: primal feasibility, dual feasibility, c.s., stationary

☆ 重要结论：若原问题为凸问题，各个函数可微，对偶间隙为0，则KKT为充要条件.

证明：(充分性) 设 $(\tilde{x},\tilde{\lambda},\tilde{\nu})$ 满足KKT条件，则必有 $(\tilde{x},\tilde{\lambda},\tilde{\nu})$ 为最优解.

只需证 $g(\tilde{\lambda},\tilde{\nu}) = f_0(\tilde{x})$ (①鞍点定义 ②鞍点定理)

$f_i(\tilde{x}) \leq 0,\ i=1,\cdots,m$

$h_i(\tilde{x}) = 0,\ i=1,\cdots,p$

$\lambda_i \geq 0,\ i=1,\cdots,m$

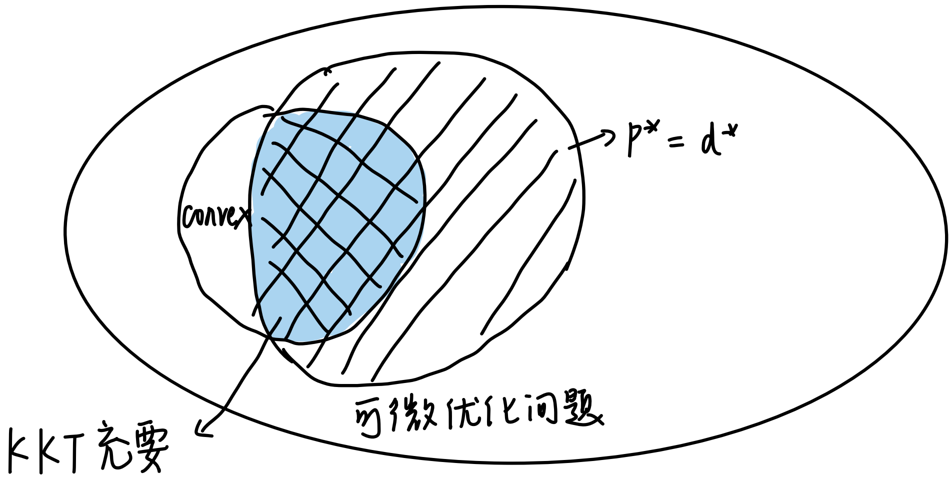
$$L(x,\tilde{\lambda},\tilde{\nu}) = f_0(x) + \underbrace{\sum_{i=1}^m \tilde{\lambda}_i f_i(x)}_{\text{非负加权和}} + \sum_{i=1}^p \tilde{\nu}_i h_i(x) \Rightarrow \text{凸函数}$$

$\left.\dfrac{\partial L(x,\tilde{\lambda},\tilde{\nu})}{\partial x}\right|_{x=\tilde{x}} = 0$ 得到的 $\tilde{x}$ 是全局最优.

$$g(\hat{\lambda},\hat{\nu}) = \inf_x L(x,\hat{\lambda},\hat{\nu}) = L(\tilde{x},\hat{\lambda},\hat{\nu})$$

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

$$= f_0(\tilde{x})$$



例: $\min \ \frac{1}{2}x^T P x + q^T x + r, \quad P \in S_+^n$

s.t. $Ax = b$

① Primal feasibility: $Ax^* = b$

② Dual feasibility: 无

③ Complementarity Slackness: 无

④ $\frac{\partial}{\partial x}\left\{\frac{1}{2}x^T P x + q^T x + r + (Ax-b)^T \nu^*\right\}\Big|_{x=x^*} = 0$

$\Leftrightarrow Px^* + q + A^T\nu^* = 0$

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

例: Water-filling Problem (通信领域)

$\min_x \ -\sum_{i=1}^n \log(\alpha_i + x_i) \qquad x \in R^n$ log取反 ⇒ 凸问题

s.t. $x \succeq 0 \qquad \alpha \in R^n$

$\mathbf{1}^T x = 1 \qquad\qquad a \succeq 0$

KKT条件：① $x^* \succeq 0$

② $\mathbf{1}^T x^* = 1$

③ $\lambda^* \succeq 0$

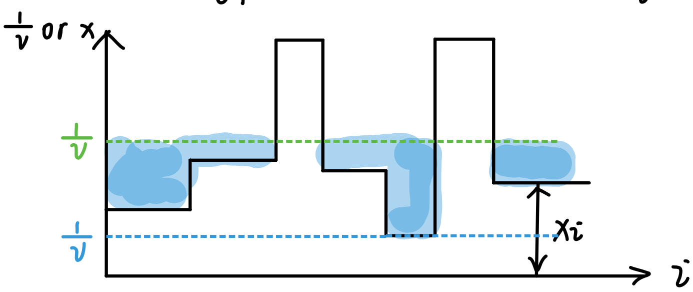
④ $\lambda_i^* x_i^* = 0,\ i=1,\cdots,n$

⑤ $-\dfrac{1}{a_i + x_i^*} - \lambda_i^* + \nu^* = 0 \Rightarrow \lambda_i^* = \nu^* - \dfrac{1}{a_i + x_i^*}$

将⑤代入③④可得：$\begin{cases} x_i^*\left(\nu^* - \dfrac{1}{a_i + x_i^*}\right) = 0, & i=1,\cdots,n \\ \nu^* \geq \dfrac{1}{a_i + x_i^*}, & i=1,\cdots,n \end{cases}$

若 $\nu^* \geq \dfrac{1}{a_i}$，则 $x_i^* = 0$ 池子里的水的高度

若 $\nu^* < \dfrac{1}{a_i}$，则 $x_i^* > 0$，$x_i^* = \dfrac{1}{\nu^*} - a_i$



- KKT条件与凸函数一阶条件

例：$\min\ f_0(x)$ $\Longleftrightarrow$ 最优性条件：$x \succeq 0$

s.t. $x \succeq 0$

$\nabla f_0(x) \succeq 0$ Complementarity

$x_i(\nabla f_0(x))_i = 0,\ i=1,\cdots,n$

互补（偏导）

$\Longleftrightarrow$ ① $x^* \succeq 0$

② $\lambda^* \succeq 0$

③ $\lambda_i^*(-x_i^*) = 0,\ i=1,\cdots,n$ 互补松弛（λ）

④ $\nabla f_0(x^*) + (-\lambda^*) = 0$

$\Longleftrightarrow$ $\nabla f_0(x^*) \succeq 0$

$x_i^* (\nabla f_0(x^*))_i = 0, \quad i=1,\cdots,n$

$\inf\limits_y (f_0(y) - v^T y)$

$\sup\limits_y (v^T y - f_0(y))$

例: $\min \ f_0(Ax+b) \iff \begin{array}{l} \min \ f_0(y) \\ \text{s.t.} \ Ax+b = y \end{array}$

$$\boxed{\begin{array}{l} L(x) = f_0(Ax+b) \\ g = \inf\limits_x f_0(Ax+b) \\ (D) \ \max g \end{array}}$$

(本质都是给原问题提供界)

$$L(x,y,v) = f_0(y) + v^T(Ax+b-y)$$

$$= f_0(y) - v^T y + v^T A x + v^T b$$

$$g(v) = \inf_{x,y} L(x,y,v) = \begin{cases} -f_0^*(v) + v^T b, & v^T A = 0 \\ -\infty, & v^T A \neq 0 \end{cases}$$

$(D) \ \max \ v^T b - f_0^*(v)$

$\quad \text{s.t.} \ v^T A = 0$ 求共轭函数

例: $\min \ \|Ax-b\| \iff \begin{array}{l} \min \|y\| \\ \text{s.t.} \ y = Ax-b \end{array}$

$$L(x,y,v) = \|y\| + v^T(Ax-b-y) = \|y\| - v^T y + v^T A x - v^T b$$

$$g(v) = \inf_{x,y} L(x,y,v) = \begin{cases} -v^T b - \|v\|_*, & v^T A = 0 \\ -\infty, & v^T A \neq 0 \end{cases}$$

$(D) \ \max \ -v^T b - \|v\|_* \iff \max \ v^T b - \|v\|_*$

$\quad \text{s.t.} \ v^T A = 0 \qquad\qquad \text{s.t.} \ v^T A = 0$

$\min \ \frac{1}{2}\|y\|^2$ (在一定意义等价. 最优解一样, 但最优值不一样)

$\text{s.t.} \ Ax - b = y$

$$\Rightarrow L(x,y,v) = \frac{1}{2}\|y\|^2 + v^T(Ax-b-y)$$

$$= \frac{1}{2}\|y\|^2 - v^T y + v^T A x - v^T b$$

$$\Rightarrow g(v) = \begin{cases} -v^T b - \frac{1}{2}\|v\|_*^2 & v^T A = 0 \\ -\infty \ (\text{看不清}) & v^T A \neq 0 \end{cases}$$

$(D) \ \max \ -v^T b - \frac{1}{2}\|v\|_*^2$

$\quad \text{s.t.} \ v^T A = 0$

例：带框约束的线性规划问题

$$\min \ c^T x$$

$$\text{s.t.} \ Ax = b$$

$$l \preceq x \preceq u$$ 按每个元素比较大小

$$\Rightarrow L(x,\lambda_1,\lambda_2,v) = c^T x + v^T(Ax-b) + \lambda_1^T(l-x) + \lambda_2^T(x-u)$$

$$= (c + A^T v - \lambda_1 + \lambda_2)^T x - v^T b + \lambda_1^T l - \lambda_2^T u$$

$$\Rightarrow g(v) = \begin{cases} -v^T b + \lambda_1^T l - \lambda_2^T u, & c + A^T v + \lambda_1 - \lambda_2 = 0 \\ -\infty & c + A^T v + \lambda_1 - \lambda_2 \neq 0 \end{cases}$$

$$(D) \ \max \ -b^T v - \lambda_1^T u + \lambda_2^T l$$

$$\text{s.t.} \ A^T v + \lambda_1 - \lambda_2 + c = 0$$

$$\lambda_1, \lambda_2 \succeq 0$$

$$\min \ f_0(x) \qquad \text{定义} \ f_0(x) = \begin{cases} c^T x & l \preceq x \preceq u \\ \infty & \text{otherwise} \end{cases}$$

$$\text{s.t.} \ Ax = b$$

$$\Rightarrow L(x,v) = f_0(x) + v^T(Ax-b)$$

$$\Rightarrow g(v) = \inf_x f_0(x) + v^T(Ax-b)$$

$$= \inf_{l \preceq x \preceq u} c^T x + v^T A x - v^T b$$

$$= \inf_{l \preceq x \preceq u} (A^T v + c)^T x - v^T b$$ 没有等式约束，只有框约束

$$= -v^T b + l^T (A^T v + c)^+ - u^T (A^T v + c)^-$$

5. 敏感性分析

$$\min\ f_0(x)\ \text{(原问题)} \qquad \Rightarrow \qquad \min\ f_0(x)\ \text{(干扰问题)}$$

$$\text{s.t.}\ f_i(x)\le 0,\ i=1,\cdots,m \qquad\qquad \text{s.t.}\ f_i(x)\le u_i,\ i=1,\cdots,m$$

$$h_i(x)=0,\ i=1,\cdots,p \qquad\qquad h_i(x)=w_i,\ i=1,\cdots,p$$

☆ $p^*(u,w)$ 最优值是关于 u 和 w 的函数，$p^*(0,0)=p^*$

$p^*(u,w)$ 的性质：

① 若原问题为凸，则 $p^*(u,w)$ 为 (u,w) 的凸函数.

证明：$p^*(u,w) = \inf\limits_{x}\left\{f_0(x)\ \middle|\ \begin{array}{l} f_i(x)\le u_i,\ i=1,\cdots,m \\ h_i(x)=w_i,\ i=1,\cdots,p \end{array}\right\}$ (记为 D)

$= \inf\limits_{x} g(x,u,w)$

$(\mathrm{dom} f_0, R^m, R^p)$ 凸集

凸函数 仿射函数 $\}$ 凸集
$f_i(x)-u_i\le 0$ $\}$

$g(x,u,w)\triangleq f_0(x)$，$\mathrm{dom}\,g = \mathrm{dom} f_0 \cap D$ $\quad h_i(x)-w_i=0$ (同理)

所以 $g(x,u,w)$ 为关于 x,u,w 的凸函数.

回忆：$\sup\limits_{y(x)} f(x,y)$ 是关于 x 的凸函数，如果 $f(x,y)$ 是关于 x 的凸函数

所以 $p^*(u,w)$ 是凸函数得证.

② 若原问题为凸，对偶间隙为0，λ^*, ν^* 为原问题对偶最优解.

$$p^*(u,w) \ge p^*(0,0) - \lambda^{*T}u - \nu^{*T}w$$

证明：设 $\tilde{x}$ 为干扰问题的最优解

$f_i(\tilde{x})\le u_i,\ i=1,\cdots,m$; $h_i(\hat{x})=w_i,\ i=1,\cdots,p$

$$p^*(0,0) = g(\lambda^*,\nu^*)$$
$$\le f_0(\tilde{x}) + \sum_{i=1}^{m}\lambda_i^* f_i(\tilde{x}) + \sum_{i=1}^{p}\nu_i h_i(\hat{x})$$
$$\le f_0(\tilde{x}) + \lambda^{*T}u + \nu^{*T}w$$
$$= p^*(u,w) + \lambda^{T*}u + \nu^{*T}w$$

举例：☆ 若 λ_i^* 很大，且加紧第 i 次不等式约束．$u_i<0$，性能↓↓↓ 最优值↑

经济学解释：资源价格贵，cost 会上升.

☆2) 若 v_i^* 很大正值，使 $w_i<0$. 或 $|v_i^*|$ 很大负值，使 $w_i<0$.

则最优值↑

3) 若 λ_i^* 很小，且 $u_i>0$，则最优值下降不大.

4) 若 v_i^* 很小正值，使 $w_i>0$，或 $|w_i|$ 很小负值，使 $w_i<0$

则最优值变化不大.

③ (局部敏感性) 若原问题为凸，对偶间隙为0，且 $p^*(u,w)$ 在 $(u,w)=(0,0)$ 处可微，则 $\lambda_i^* = -\dfrac{\partial p^*(0,0)}{\partial w_i}$ ，$v_i^* = -\dfrac{\partial p^*(0,0)}{\partial w_i}$

Taylor展开：$p^*(u,w) = p^*(0,0) - \lambda^{*T} u - v^{*T} w$

(用性质二和可微性证明)

例：Boolen LP问题 $\min \ c^T x$

$\text{s.t.} \ Ax \leq b$

提供下界

$x_i \in \{0,1\}, \ i=1,\cdots,n$

LP松弛问题(松弛) $\min \ c^T x$

$\text{s.t.} \ Ax \leq b$

$0 \leq x_i \leq 1 \quad , i=1,\cdots,n$

Boolen LP等价问题 $\min \ c^T x$

$\text{s.t.} \ Ax \leq b$

不是凸问题，因为等式约束非仿射 ⇐ $x_i(x_i-1)=0, \ i=1,\cdots,n$

$$L(x,\lambda,v) = c^T x + \lambda^T(Ax-b) + \sum_{i=1}^n v_i x_i^2 - \sum_{i=1}^n v_i x_i$$

$$= \sum_{i=1}^n v_i x_i^2 + (c + A^T\lambda - v)^T x - \lambda^T b$$

$$g(\lambda,\nu) = \inf_{x} L(x,\lambda,\nu) = \begin{cases} -\lambda^T b - \frac{1}{4}\sum_{i=1}^{n} (c_i + a_i^T\lambda - \nu_i)^2/\nu_i \ , & \nu \geq 0 \\ -\infty \ , & \text{otherwise} \end{cases}$$

$$(\text{Dual}) \quad \max \quad -\lambda^T b - \frac{1}{4}\sum_{i=1}^{n} (c_i + a_i^T\lambda - \nu_i)^2/\nu_i$$

$$\text{s.t.} \quad \lambda \geq 0,\ \nu \geq 0$$

用到概念 $\max\limits_{\lambda,\nu} f(\lambda,\nu) = \max\limits_{\lambda} \max\limits_{\nu} f(\lambda,\nu)$

$$\min_{\nu_i \geq 0} \frac{(\nu_i - \beta_i)^2}{\nu_i}$$

$$\forall \lambda, \quad \max \quad -\lambda^T b - \frac{1}{4}\sum_{i=1}^{n} (\overbrace{c_i + a_i^T\lambda}^{\text{记为}\beta_i} - \nu_i)^2/\nu_i \ , \quad \nu \geq 0$$

$$= \begin{cases} c_i + a_i^T\lambda & c_i + a_i^T\lambda \leq 0 \\ 0 & c_i + a_i^T\lambda > 0 \end{cases}$$

$$= \min\{0,\ c_i + a_i^T\lambda\}$$

$$\Longleftrightarrow \quad \max_{\lambda} \quad -\lambda^T b + \sum_{i=1}^{n} \underbrace{\min\{0,\ c_i + a_i^T\lambda\}}_{w_i}$$

$$\text{s.t.} \quad \lambda \geq 0$$

$$\Longleftrightarrow \quad \max_{\lambda} \quad -\lambda^T b + 1^T w$$

$$\text{s.t.} \quad \lambda \geq 0$$

$$w_i \leq a_i^T\lambda + c_i$$

$$w_i \leq 0$$

Dual Problem:

Lagrangian Relaxation

(和LP松弛的对偶)

$$L(x,u,v,w) = c^T x + u^T(Ax - b) - v^T x + w^T(x - 1)$$

$$= (c + A^T u - v + w)^T x - b^T u - 1^T w$$

$$\Rightarrow \max_{\lambda} \quad -b^T u - 1^T w$$

$$\text{s.t.} \quad c + A^T u - v + w = 0 \qquad (\text{等价于}) \quad c + A^T u + w \geq 0$$

$$u \geq 0,\ v \geq 0,\ w \geq 0 \qquad\qquad u \geq 0,\ w \geq 0$$

例：带等式约束的可微凸优化问题

$$\min f_0(x) \text{ 凸可微} \xrightarrow{\text{罚函数方法}} \min f_0(x)+\frac{\partial}{2}\|Ax-b\|_2^2 \Rightarrow \hat{x}$$

$$\text{s.t. } Ax-b=0$$

$\partial=\infty$ 一定等价. 固定∂求解, 然后不断增加∂.

但是为了数值稳定性. 收敛到 (P) optimal

$$\nabla f_0(\hat{x})+\partial A^T(A\hat{x}-b)=0$$

$$\min f_0(x)+\partial(A\hat{x}-b)^T(Ax-b)$$ [前后互相等价]

$$L(x,v)=f_0(x)+v^T(Ax-b)$$

$$g(v)=\inf_x\{f_0(x)+v^T(Ax-b)\}$$, 取 $v=\partial(A\hat{x}-b)$, 则与 $\max_v g(v)$ 等价.

$$g(\partial(A\hat{x}-b))=\inf_x\{f_0(x)+\partial(A\hat{x}-b)^T(Ax-b)\}$$

$$=f_0(\hat{x})+\partial\|A\hat{x}-b\|_2^2$$

$$f_0(x^*)=P^*=d^*\geq g(\partial(A\hat{x}-b))=f_0(\hat{x})+\partial\|Ax-b\|_2^2\geq \underline{f_0(\hat{x})}$$

惩罚函数最优解对应的目标函数值.

当$\partial=0$时, $\arg\min f_0(x)$

当$\partial\to\infty$时, $f(x^*)=f(\hat{x})$

超平面上

思路: 允许约束不满足, 但慢慢迭代后可以再回到约束, 进而趋近最优解.

例: 带线性不等式约束的可微凸优化问题 (log-barrier 方法)

$$\min f_0(x) \qquad x\in R^n,\ A\in R^{m\times n},\ b\in R^m$$

$$\text{s.t. } Ax\geq b$$

$u=0$时计算稳定性不好

(数值计算不可能有等号)

$$\text{Log-barrier } \min f_0(x)\left[-\sum_{i=1}^m u_i\left[\log(a_i^Tx-b_i)\right]\right]^{+\infty}$$

"离我远一点!"

固定$u=\infty$时求解, 然后慢慢减小.

设$\hat{x}$为罚函数最优解, $\nabla f_0(\hat{x})-\sum_{i=1}^m u_i\frac{a_i}{a_i^T\hat{x}-b_i}=0$

$$\hat{x}=\arg\min_x f_0(x)-\sum_{i=1}^m u_i\frac{a_i^Tx-b_i}{a_i^T\hat{x}-b_i}$$ ($\hat{x}$同时是两个问题的最优解)

$$L(x,\lambda)=f_0(x)+\sum_{i=1}^m\lambda_i(b_i-a_i^Tx)$$

$$g(\lambda)=\inf_x\{f_0(x)+\sum_{i=1}^m\lambda_i(b_i-a_i^Tx)\}$$, 取 $\lambda_i=\frac{u}{a_i^T\hat{x}-b_i}$

原问题 $\longrightarrow$ 惩罚，惩罚因子 λ, u

$\updownarrow$ Lagrange multiplier

原问题 $\searrow$ 对偶函数

作业：(4) 3, 9, 22, 24, 59, 62, (5) 5, 20, 27