

I. 定义

广义：凸目标，凸集约束的极小化问题。

-般优化问题： $\min f_0(x)$ 凸函数

凸集 $\left\{ \begin{array}{l} \text{δ-sublevel set } s.t. \underline{f_i(x) \leq 0}, i=1, \dots, m \text{ 凸函数} \\ \text{子空间 } \underline{a_i^T x = b_i \in [h_i(x) = 0]}, i=1, \dots, p \text{ 仿射函数} \end{array} \right\}$ 狹义凸問題

一些定义：① $x \in \mathbb{R}^n$. optimization variable

② $f_0: \mathbb{R}^n \rightarrow \mathbb{R}$, 目标函数 objective function,

损失函数 loss function, 效用函数 utility function

③ $f_i(x) \leq 0$ 不等式约束 Inequality constraint

$h_i(x) = 0$ 等式约束 equality constraint

④ $m=p=0$ 无约束 unconstrained

⑤ 优化问题的域 domain: $D = \bigcap_{i=1}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i$

⑥ 可行解集 feasible set 若 $f_i(x) \leq 0, i=1, \dots, m$

★ $x \in D$ 为可行解, $h_i(x) = 0, i=1, \dots, p$

$$X_f = \{x \text{ 为可行解}\}$$

⑦ 问题的最优值 optimal value: $p^* = \inf \{f_0(x) | x \in X_f\}$

若 X_f 为空集, $p^* = +\infty$.

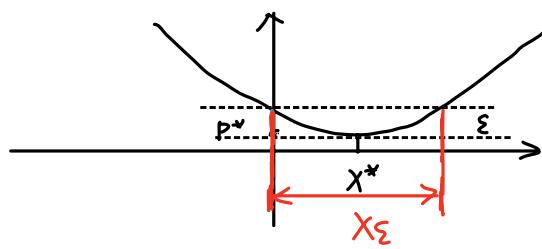
⑧ 最优解 optimal point/solution: 若 x^* 可行, 且 $f_0(x^*) = p^*$.

⑨ 最优解集 optimal set: $X_{opt} = \{x | x \in X_f, f_0(x) = p^*\}$

⑩ Σ : 次优解集. ε -suboptimal set

补充: satisfying (satisfy + suffice) solutions

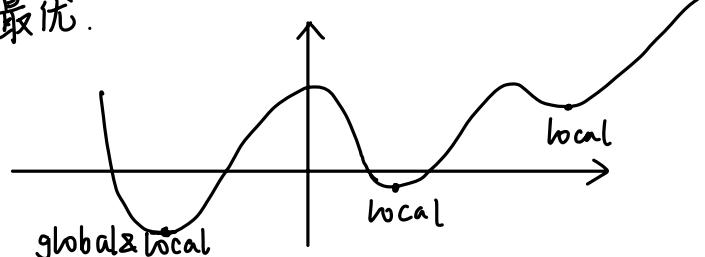
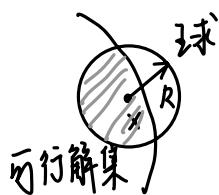
$$X_\varepsilon = \{x | x \in X_f, f_0(x) \leq p^* + \varepsilon\}$$



①局部最优解 locally optimal

$$f_0(x) = \inf \left\{ f_0(z) \mid \begin{array}{l} f_i(z) \leq 0, i=1, \dots, m \\ h_i(z) = 0, i=1, \dots, p \\ \|z-x\| \leq R \end{array} \right.$$

$\exists R > 0$ 使 x 局部最优.



X_{loc} (局部最优解集)

* 各种解之间的关系



* why $f_i(x) \leq 0$ rather than $f_i(x) < 0$?

若 $x \in X_f$, $f_i(x) = 0$, 则 $f_i(x) \leq 0$ 为活动约束 (active)

$f_i(x) < 0$, 则 $f_i(x) \leq 0$ 为不活动约束 (inactive)

eg: $\min -\text{utility}$

s.t. $\text{money} \leq 100$ (不能全都花完)

$-|\log(100 - \text{money})| \leq 100$ (使用域的方法)

2. 可行性优化问题 feasibility Problems

find $x \quad \min \Phi$

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

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

等价问题 例: Box Constraints

$$\min f_0(x)$$

$$\text{s.t. } l_i \leq x_i \leq u_i, i=1, \dots, n$$

$$\left\{ \begin{array}{l} l_i - x_i \leq 0, i=1, \dots, n \\ x_i - u_i \leq 0, i=1, \dots, n \end{array} \right.$$

例: $\min \varphi_0(f_0(x))$ (重纲标准化)

$$\text{s.t. } \varphi_i f_i(x) \leq 0, i=1, \dots, m, \varphi_i > 0$$

$$\varphi_i h_i(x) = 0, i=1, \dots, p, \varphi_i \neq 0$$

例: $\varphi_0: R \rightarrow R$ 单调递增.

$$\varphi_1, \dots, \varphi_m: R \rightarrow R \quad \varphi_i(u) \leq 0 \Leftrightarrow u \leq 0$$

$$\varphi_1, \dots, \varphi_m: R \rightarrow R \quad \varphi_i(u) = 0 \Leftrightarrow u = 0$$

$$\min \varphi_0(f_0(x))$$

$$\text{s.t. } \varphi_i(f_i(x)) \leq 0, i=1, \dots, m$$

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

例: $\min \|Ax - b\|_2$

$$\min \|Ax - b\|_2^2$$

例：消除等式约束。

$\{h_i(x) = 0, i=1, \dots, p\}$ 一组方程， $z \in R^k$, $\varphi: R^k \rightarrow R^n$, $x = \varphi(z)$ 代入

$$\min f_0(\varphi(z))$$

$$\text{s.t. } f_i(\varphi(z)) \leq 0, i=1, \dots, m$$

例：消除线性等式约束。 $Ax - b = 0$, $A \in R^{p \times n}$ $\Leftrightarrow x = \varphi(z)$ $\left\{ \begin{array}{l} \text{①无解} \rightarrow \text{原问题无解} \\ z \in R^{n-r} (\text{降维, 降秩}) \quad x = Fz + x_0 \quad \text{化零空间} \end{array} \right.$ $\left. \begin{array}{l} \text{由} z \text{决定的空间} \\ \text{②} A \text{可逆}, \quad x = A^{-1}b \end{array} \right.$

$$\min f_0(Fz + x_0) \quad \text{约束维数减少, 不再有等式约束}$$

$$\text{s.t. } f_i(Fz + x_0) \leq 0, i=1, \dots, m$$

例： $\min f_0(x)$ 是否关于 x 和 s 都是凸的（联合凸）

$$\text{s.t. } s_i \leq 0, i=1, \dots, m$$

$$f_i(x) - s_i = 0, i=1, \dots, m \quad \text{slack variable}$$

$$a_i^\top x = b_i, i=1, \dots, p$$

$$\min f_0(x) = x_1^2 + x_2^2$$

$$\text{s.t. } f_1(x) = \frac{x_1}{1+x_2^2} \leq 0$$

$$h_1(x) = (x_1 + x_2)^2 = 0$$

$$\min f_0(x) = x_1^2 + x_2^2$$

$$\text{s.t. } f_1(x) = x_1 \leq 0$$

$$\tilde{h}_1(x) = x_1 + x_2 = 0$$

☆ 区分 Quasi-convex optimization & Non-convex optimization

3. 重要性质：局部最优 = 全局最优

① 局部最优： $\exists R > 0, f_0(x) = \inf \{f_0(z) | z \text{可行}, \|x-z\|_2 \leq R\}$

证明：（反证法）设 x 不是全局最优， $\exists y$ 可行， $f_0(y) < f_0(x)$

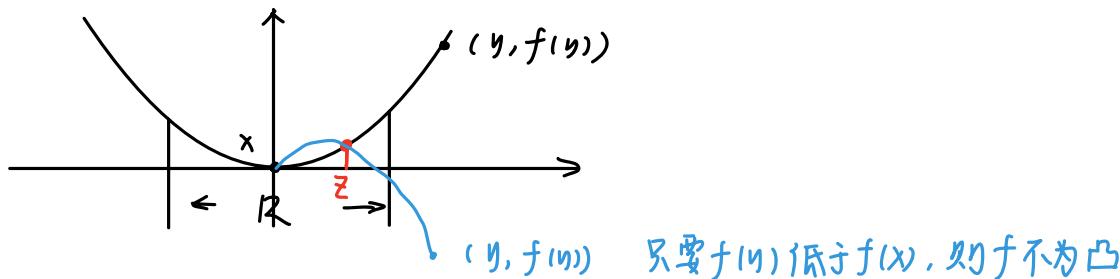
因 x 局部最优 $\|y-x\|_2 > R$

$z = (\theta)x + (1-\theta)y$, 取 $\theta = \frac{R}{2\|y-x\|_2}$ (由于 $\theta \in (0,1)$, 则 z 必为凸组合)

z 可行, 且 $f_0(z) \leq (\theta)f_0(x) + (1-\theta)f_0(y) < f_0(x)$

$\|z-x\|_2 = \theta\|x-y\|_2 = \frac{R}{2} < R$, then $f_0(x) \leq f_0(z)$

$f_0(y) < f_0(x) \leq f_0(z)$, 与上述条件矛盾.



② 可微目标函数情况下的最优性

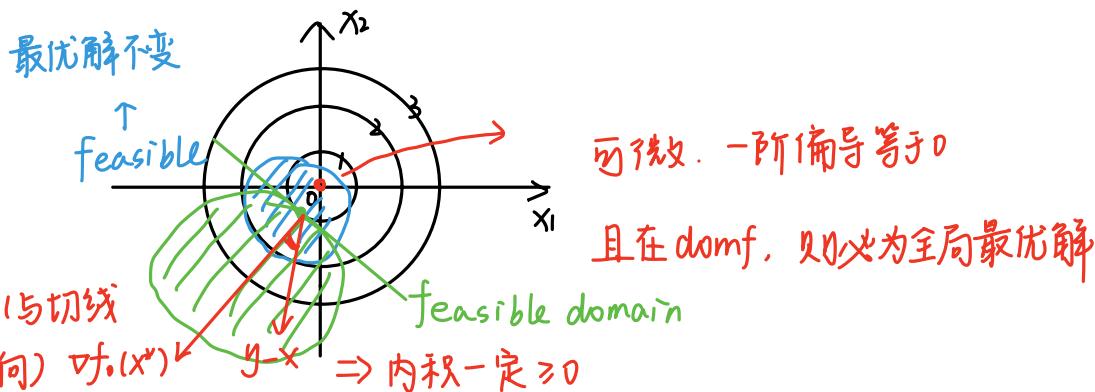
f_0 可微, 则 f_0 凸 $\Leftrightarrow \text{dom } f_0$ 为凸, $f_0(y) \geq f_0(x) + \nabla f_0^T(x)(y-x)$, $\forall x, y \in \text{dom } f_0$.

设凸问题可行域 $X_f = \left\{ x \mid \begin{array}{l} f_i(x) \leq 0, i=1, \dots, m \\ h_j(x) = 0, j=1, \dots, p \end{array} \right\} \cap \text{dom } f_0$

$x^* \in X_f$ 最优 $\Leftrightarrow \nabla f_0^T(x^*)(y-x^*) \geq 0, \forall y \in X_f$

直观解释：令 $\text{dom } f_0 \subseteq X_f$, 则 $\exists x^*, \forall y \in \text{dom } f_0$.

若 $f(x^*)$ 最小, 则 $f_0(y) \geq f_0(x^*)$, 则 $\nabla f_0^T(x^*)(y-x^*) \geq 0$



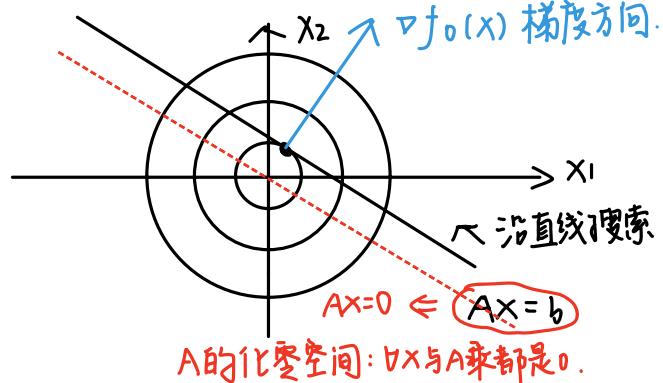
13) 约束仅为等式约束

$$\min f_0(x), \quad \text{dom } f_0 = \mathbb{R}^n$$

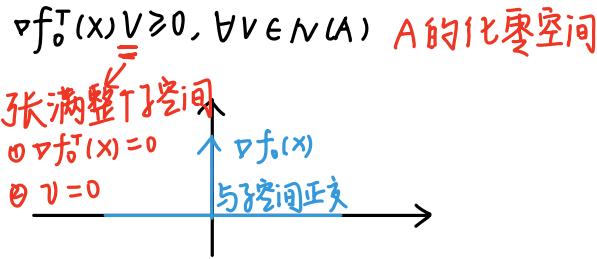
$$\text{s.t. } Ax = b$$

若 $\exists X, Ax = b, X$ 最优 $\Leftrightarrow \exists y, Ay = b$

$$\Rightarrow \nabla f_0^T(\lambda)(y - x) \geq 0$$



$$Ax = b, Ay = b \Rightarrow y = x + v, v \in N(A)$$



① $v = 0, Ax$ 可逆, $x = A^{-1}b$

② $\nabla f_0(x)$ 与 $N(A)$ 正交

13) 约束仅为非负约束

$$\min f_0(x)$$

$$\text{s.t. } x \geq 0$$

若 $\exists x \geq 0, x$ 最优 $\Leftrightarrow \forall y \geq 0, \nabla f_0^T(x)(y - x) \geq 0$

$$\text{即 } \nabla f_0^T(x)y - \nabla f_0^T(x)x \geq 0$$

① 若 $(\nabla f_0(x))_i < 0$, 则 $\nabla f_0^T(x)y$ 必可取到无穷小. 则必有 $\nabla f_0(x) \geq 0$

② $\forall y$ 均有 $\nabla f_0^T(x)(y - x) \geq 0 \Rightarrow \nabla f_0^T(x)x \leq 0$

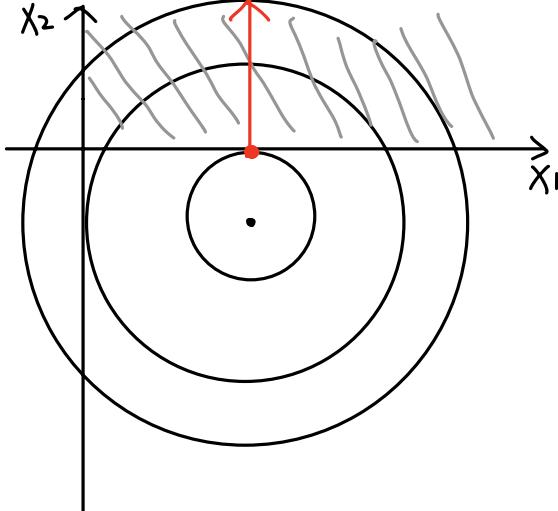
③ $\nabla f_0^T(x) \geq 0, x \geq 0 \Rightarrow \nabla f_0^T(x)x \geq 0$

Complementarity

互补条件

总结: 如果 x 是最优解, 则

$$\begin{cases} x \geq 0 \\ \nabla f_0(x) \geq 0 \\ (\nabla f_0(x))_i x_i = 0 \end{cases}$$



最优解的直观理解：能使等高线到达最低.

$$x_1 > 0, x_2 = 0$$

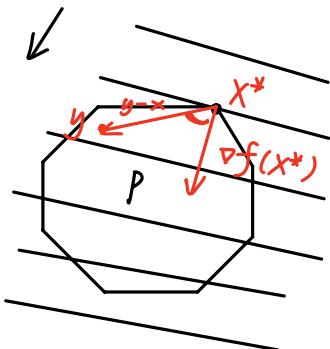
$$(\nabla f_0(x))_1 = 0, (\nabla f_0(x))_2 > 0$$

4. 典型凸问题

$$\begin{aligned} \text{(I) 线性规划} \quad & \min c^T x + d \quad c \in \mathbb{R}^n, d \in \mathbb{R} && \text{仿射} \\ \text{s.t.} \quad & Gx \leq h \quad G \in \mathbb{R}^{m \times n}, h \in \mathbb{R}^m && \text{仿射} \\ & Ax = b \quad A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^k && \text{仿射} \end{aligned} \quad \left. \begin{array}{l} \text{仿射} \\ \text{仿射} \\ \text{仿射} \end{array} \right\} \text{凸问题}$$

目标与约束均为线性 Linear program (LP问题)

历史：1930-1940 调度问题 Dantzig、Von Neumann、Kantorovich



$\nabla f(x^*)(y - x)$ 一定非负.

$$\begin{aligned} \text{(等价变换)} \quad & \min c^T x + d && \min c^T x^+ - c^T x^- + d \\ \text{s.t.} \quad & Gx + s = h &\Leftrightarrow& G_1 x^+ - G_1 x^- + s = h \\ & Ax = b && Ax^+ - Ax^- = b \\ & s \geq 0 && s \geq 0, x^+ \geq 0, x^- \geq 0 \end{aligned}$$

$$\text{eg: } x = [1, -1, 2, -2]^T, x^+ = [1, 0, 2, 0]^T, x^- = [0, 1, 0, 2]^T$$

$$x = x^+ - x^-, x^+, x^- \geq 0$$

$$\begin{cases} \min & C^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{cases} \quad (\text{linprog 函数的命名规则})$$

例：食谱问题。

一份健康的饮食包含 m 种不同的营养，每种至少需要 b_1, \dots, b_m 。我们可以从 n 种食物中选择非负的量 x_1, \dots, x_n 以构成一份食谱。单位第 j 种食品含有营养 i 的量为 a_{ij} ，而价格为 c_j 。我们希望设计出一份最便宜的满足营养需求的食谱。这一问题可以描述为线性规划

目标：单位食物价格 c_j ，找到总价最低的食谱。

$$\begin{array}{ll} \min & \sum_{j=1}^n c_j x_j \\ \text{s.t.} & \sum_{j=1}^n a_{ij} x_j \geq b_i, i=1, \dots, m \\ & x_j \geq 0, j=1, \dots, n \end{array} \xrightarrow{\text{向量形式}} \begin{array}{l} \min [c_1 \dots c_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ \text{s.t.} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \\ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \geq 0 \end{array}$$

(2) 线性分数规划 linear fractional programming

$$(P_0) \quad \min f_0(x) \quad f_0(x) = \frac{c^T x + d}{e^T x + f} \text{ 是拟凸函数, 不是凸函数}$$

$$\text{s.t. } Gx \leq h \quad \text{dom } f_0 = \{x \mid c^T x + f > 0\}$$

$$Ax = b$$

$$(P_1) \quad \min C^T x + d^T z$$

验证：① x^* 是 P_0 的 optimal $\Leftrightarrow y^*, z^*$ 是 P_1 的 optimal

$$\text{s.t. } Gy - hz \leq 0$$

② $\forall x$, 能找 y, z 对于 P_1 可行, 且函数值相等

$$Ay - bz = 0$$

$\forall y, z$, 能找到 x 对于 P_0 可行, 且函数值相等

$$e^T y + fz = 1$$

$$z \geq 0$$

证明: 1) 若 x 在 P_0 可行. 取 $y = \frac{x}{e^T x + f}$, $z = \frac{1}{e^T x + f}$

$$\begin{cases} Gx \leq h \\ Ax = b \\ e^T x + f > 0 \end{cases} \quad \begin{aligned} Gy - hz &= \frac{Gx - h}{e^T x + f} \leq 0 \\ Ay - bz &= \frac{Ax - b}{e^T x + f} = 0 \\ e^T x + fz &= 1 \end{aligned}$$

$$z \geq 0$$

目标函数值都 = $\frac{C^T x + d}{e^T x + f}$

2) 若 y, z 在 P_1 中可行

若 $z > 0$, 则 $x = \frac{y}{z}$, then x 在 P_0 中可行且两问题在目标函数相同.

若 $z = 0$, 设 x_0 为 P_0 可行解, 则 $\underline{x = x_0 + ty}$ 对 P_0 可行, $\forall t \geq 0$

$$Gy \leq 0, Ay = 0, e^T y = 1,$$

$$Gx = Gx_0 + tGy \leq h$$

$$Ax = Ax_0 + tAy = 1$$

$$e^T x + f = e^T x_0 + f + t e^T y > 0$$

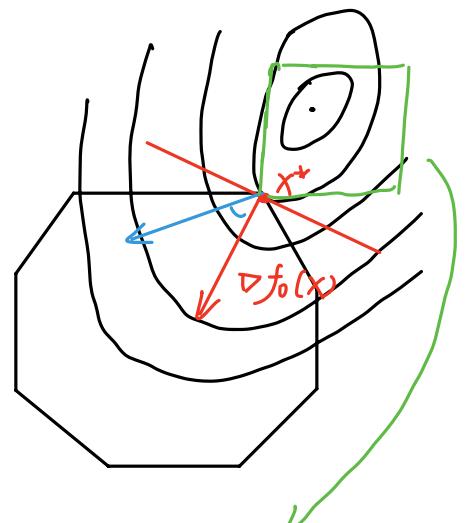
$$f_0(x) = f_0(x_0 + ty) = \frac{C^T x_0 + C^T ty + d}{e^T x_0 + e^T ty + f} \xrightarrow[t \rightarrow \infty]{} C^T y$$

(3) 二次规划 Quadratic Programming

$$\min \frac{1}{2} x^T P x + q^T x + r, \quad P \in S^n_+$$

$$\text{s.t. } Gx \leq h$$

$$Ax = b$$



最优解不一定在边缘 or 顶点

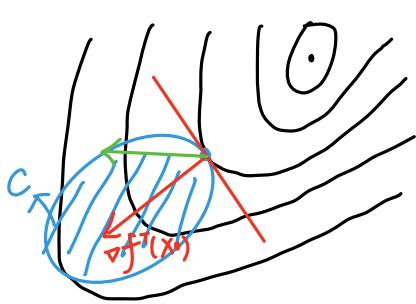
还可能在集合内部.

(4) 二次约束二次规划 QCQP

$$\min \frac{1}{2} x^T P x + q^T x + r$$

$$\text{s.t. } \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0$$

$$Ax = b$$



例：带噪声的测量系统 $b = Ax + e$ 待测量向量

$$\hat{x} = \arg \min_x \|b - Ax\|_2 = \arg \min_x \|b - Ax\|_2^2$$

$$= \arg \min_x x^T \underbrace{A^T A}_{半正定} x - 2b^T A x + b^T b \quad (\text{无约束的QP问题})$$

$$= (A^T A)^{-1} A^T b \quad (-\text{阶偏导} = 0 \text{ 来计算})$$

(推广) 若 X 稀疏 Lasso

$$\hat{x} = \arg \min_x \|b - Ax\|_2^2 + \lambda_0 \|x\|_0 \quad \text{非凸函数(用凸近似)}$$

$$\text{凸问题/QP} \Leftarrow = \arg \min_x \|b - Ax\|_2^2 + \lambda_1 \|x\|_1 \quad (L_1-\text{Regularized least squares})$$

$$(\text{技巧: } x = x^+ - x^-) \star = \arg \min_{x^+, x^-} \|b - Ax^+ + Ax^-\|_2^2 + \lambda_1 \|x^+ - x^-\|_1 + \lambda_2 \|x^+\|_1 + \lambda_2 \|x^-\|_1$$

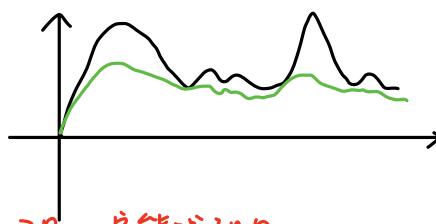
$$\text{非光滑} \rightarrow \text{光滑可微} \quad \text{s.t. } x^+, x^- \geq 0 \quad (\text{本质是二次规划问题})$$

(再推广) X 中元素幅度类似 $\star L_2\text{-regularized least square}$ \star 岭回归

判断二次项是否
一定是正定矩阵

$$\hat{x} = \arg \min_x \|b - Ax\|_2^2 + \lambda_2 \|x\|_2^2$$

Δ 半正定 $A^T A + \lambda_2 I$ $\lambda > 0, \text{ 正定}$
 \Rightarrow 确认二次规划



$$\text{QCQP问题} \Leftarrow = \arg \min_x \|b - Ax\|_2^2 \Leftrightarrow \text{转化: } \forall \lambda_2 \geq 0, \text{一定能找到} \star$$

$$\text{s.t. } \|x\|_2^2 \leq \underline{\theta}^{>0} \quad \text{使两个问题最优解一样}$$

例：投资组合问题

B	Bucket	
	beginning	ending
#1	x_1	$p_1 x_1$
:	:	:
#n	x_n	$p_n x_n$

$$\max \quad \underline{p}_1 x_1 + \cdots + \underline{p}_n x_n$$

$$\text{s.t.} \quad x_1 + \cdots + x_n \leq B$$

$$x_1, \dots, x_n \geq 0$$

eg: $\bar{p}^T = [1.05, 1.05, 1]$, $\Sigma = \begin{bmatrix} 1.2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$\min \quad x^T \Sigma^{-1} x \quad \text{(确定收益下限)} \quad (1952, \text{Markowitz}) \Rightarrow \text{出名要趁早.}$$

s.t. $\bar{p}^T x \geq r_{\min}$ 最小化风险

$$I^T x = B$$

$$x \geq 0$$

(5) 半正定规划 Semi-Definite Programming SDP

$\min \quad \text{tr}(Cx)$ 矩阵空间的线性规划 (优化矩阵)

s.t. $\text{tr}(A_i x) = b_i, i=1, \dots, p$

$$x \geq 0$$

$$x \in S_+^n, C \in R^{n \times n}, A_i \in R^{n \times n}, b_i \in R$$

eg: 特例. 对角矩阵 $\text{diag}\{x\}$



$$\min (\text{diag}\{c\})^T \text{diag}\{x\}$$

s.t. $(\text{diag}\{A_i\})^T \text{diag}\{x\} = b_i, i=1, \dots, p$

$$\text{diag}\{x\} \geq 0$$

$$\min C^T x \quad (\text{优化向量})$$

$$\text{s.t. } X_1 A_1 + \cdots + X_n A_n \preceq B \quad \text{半正定约束, 左-右=非正定矩阵}$$

$$X \in \mathbb{R}^k, B, A_1, \dots, A_n \in \mathbb{S}^k, C \in \mathbb{R}^n \quad \text{对称的} k \times k \text{维的矩阵}$$

$$\text{Def: } A(x) = A_0 + X_1 A_1 + \cdots + X_n A_n$$

$$A_i \in \mathbb{R}^{p \times q}, i=0, \dots, n, X \in \mathbb{R}^n$$

谱范数 $\|A(x)\|_2$, 即 $A(x)$ 最大的奇异值. $\rightarrow A^{-1}(x) A(x)$

$$\min \|A(x)\|_2$$

标量

$$\text{补充线代: } \|A(x)\|_2 \leq \sqrt{s}, s > 0 \Leftrightarrow A^T(x) A(x) - sI \preceq 0$$

(用对角矩阵直观理解)

$$\min \sqrt{s}$$

$$\text{s.t. } A^T(x) A(x) \preceq sI \quad \text{二次项系数是半正定矩阵}$$

$$\Leftrightarrow \frac{1}{2} \sqrt{s} = t$$

$$\min t$$

$$\text{s.t. } A^T(x) A(x) - t^2 I \preceq 0 \quad \Leftrightarrow \text{s.t. } \begin{bmatrix} tI & A(x) \\ A^T(x) & tI \end{bmatrix} \succeq 0$$

$t \geq 0$ 非凸, 很难

$$\min t$$

$t \geq 0$

$$\Leftrightarrow$$

$$\min t \quad \text{转化成标准的SDP问题}$$

$$\text{s.t. } Y = \begin{bmatrix} tI & A(x) \\ A(x) & tI \end{bmatrix} \succeq 0$$

$Y \succeq 0$ 半正定

$t \geq 0$

(6) 多目标优化问题

$$\min f_0(x)$$

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

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

$$f_0: \mathbb{R}^n \rightarrow \mathbb{R}^q, f_i: \mathbb{R}^n \rightarrow \mathbb{R}, h_i: \mathbb{R}^n \rightarrow \mathbb{R}$$

例：投资组合问题 $\min Risk$

$$\min -income$$

s.t. Resources

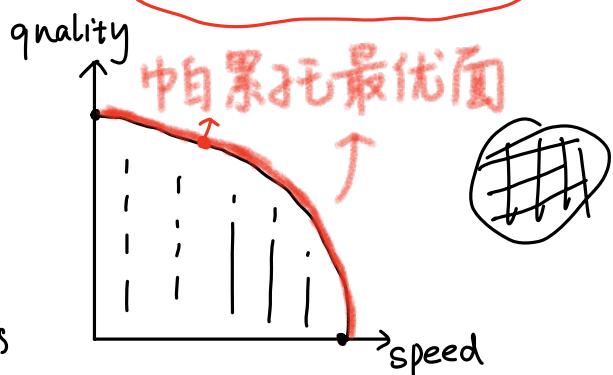
例：又好又快发展

$$\min -speed$$

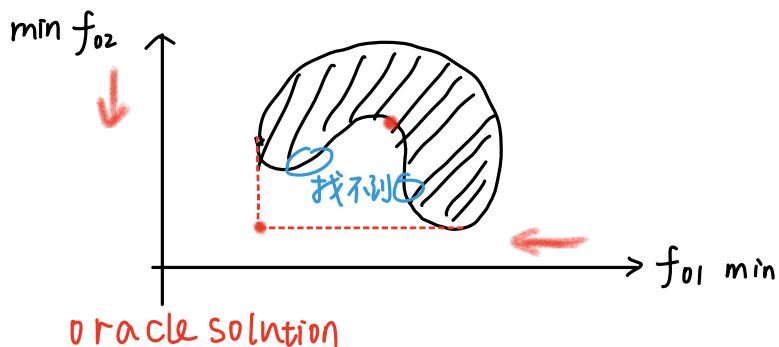
$$\min -quality$$

s.t. resources

包含两个目标函数值和一个最优策略（帕累托最优值、点）



{ Pareto optimal front: 任意点，若能找到另一解，在某指标更优，则其他更差。
 { Pareto optimal value / point: 对应的值和点。



不要试图解多目标优化

将其转化为单目标

若 $\{f_0(x)\}$ 在 \mathbb{R}^k 中为凸, $f_i(x)$ 为凸, $h_i(x)$ 为仿射, 则可以通过下述方法

求得 Pareto front 中一点

$$\min \sum_{i=1}^n \lambda_i f_{\theta_i}(x) \quad \lambda_i \geq 0$$

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

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

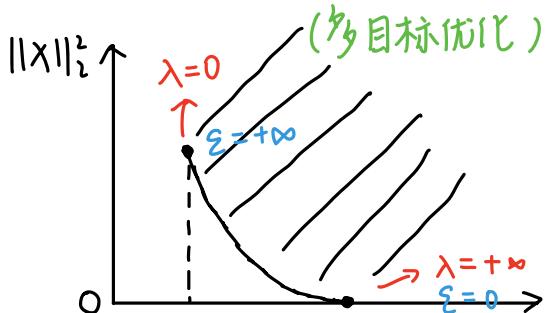
通过遍历 $\{\lambda_i\}$ 可找出所有点.

3: Ridge Regression

$$b = Ax + e, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, e \in \mathbb{R}^n, x \in \mathbb{R}^n$$

$$\min \|b - Ax\|_2^2 \quad \text{① 误差平方和最小} \quad \Leftrightarrow \min \|b - Ax\|_2^2 + \lambda \|x\|_2^2$$

$$\min \|x\|_2^2 \quad \text{② } x \text{ 的能量最小.} \quad (\text{标准岭回归})$$



☆ 通过拉格朗日来联系起来.

有约束 \rightarrow 无约束.

$$\min \|b - Ax\|_2^2 \quad (\text{岭回归的变形})$$

$$\text{s.t. } \|x\|_2^2 \leq \epsilon$$