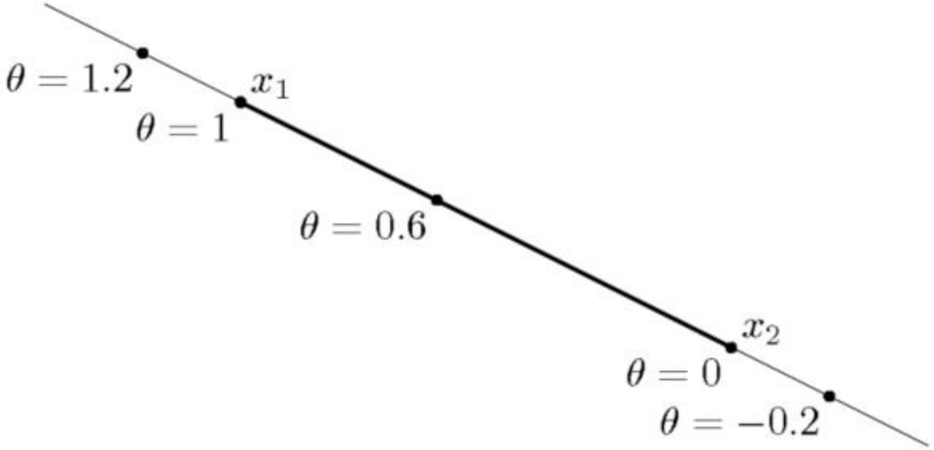
بهینه سازی
مجموعه کوژ- تابع کوژ

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

مجموعه افین

$$C \subseteq \mathbb{R}^n$$

$$\boldsymbol{x}_۱, \boldsymbol{x}_۲ \in C \overset{\theta\in\mathbb{R}}{\Longrightarrow} \boldsymbol{x}_\theta = \theta \boldsymbol{x}_۱ + (1-\theta)\boldsymbol{x}_۲ \in \mathrm{C}$$



مجموعه افین

خط در فضای $\mathbb{R}^n$

فضای $\mathbb{R}^n$

$\mathrm{C} = \{\mathbf{x} | \mathrm{A}\mathbf{x} = \mathbf{b}\}$

پوستهٔ افین

$$Aff(C) = \left\{\theta_۱ x_۱ + \theta_۲ x_۲ \cdots + \theta_n x_n | x_۱, x_۲, \cdots, x_n \in C, \sum_{i=۱}^{n} \theta_i = ۱\right\}$$

مجموعهٔ کوژ

مجموعهٔ کوژ

- $S \in \mathbb{R}^n$
- «قطعه» خطی مستقیم واصل دو نقطه از S
  - کل قطعه خط در S خواهد بود

$$\boldsymbol{x}_۱, \boldsymbol{x}_۲ \in S \xRightarrow{\forall\theta\in[0,۱]} \boldsymbol{x}_\theta = \theta\boldsymbol{x}_۱ + (1-\theta)\boldsymbol{x}_۲ \in S$$

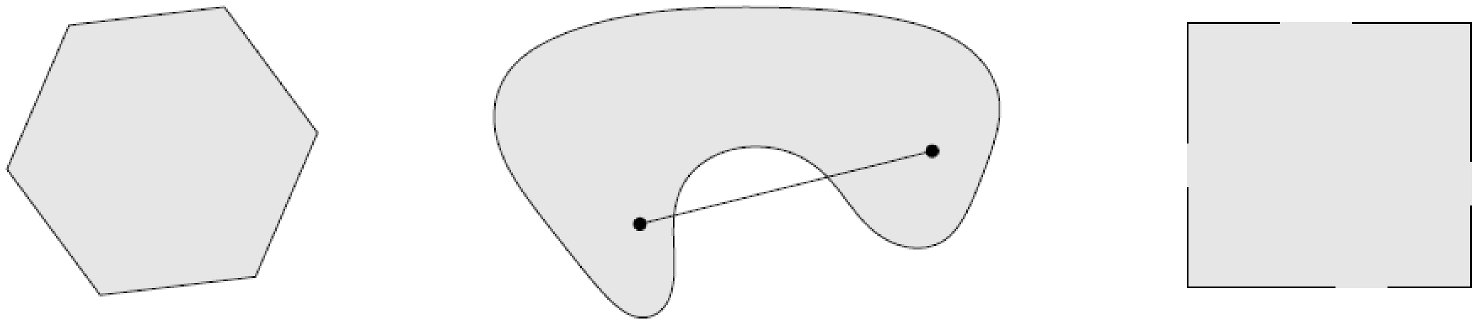
مجموعهٔ کوژ

مجموعهٔ کوژ

- $S \in \mathbb{R}^n$
- «قطعه» خطی مستقیم واصل دو نقطه از S
  - کل قطعه خط در S خواهد بود

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in S \xRightarrow{\forall\theta\in[0,1]} \boldsymbol{x}_\theta = \theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in S$$

چند نمونه کوژ و ناکوژ



مجموعهٔ کوژ

مجموعه تهی

تک نقطه

تمامی فضای $\mathbb{R}^n$

خط یا تکه خط

ابرصفحه (معادله خطی) $\boldsymbol{a}^T\boldsymbol{x} = b$

نیم‌فضا (نامساوی خطی) $\boldsymbol{a}^T\boldsymbol{x} \geq b$

گوی $\|\boldsymbol{x} - \boldsymbol{x}'\| \leq h$

مخروط کوژ

مثال-نیم‌فضا

$$C = \{\boldsymbol{x} | A\boldsymbol{x} \le \boldsymbol{b}\},$$

$$A \in R^{m \times n}, \boldsymbol{b} \in R^m$$

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in C \Longrightarrow A\boldsymbol{x}_1 \le \boldsymbol{b}, A\boldsymbol{x}_2 \le \boldsymbol{b}$$

$$A(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) = \theta A\boldsymbol{x}_1 + (1-\theta)A\boldsymbol{x}_2$$

$$\le \theta\boldsymbol{b} + (1-\theta)\boldsymbol{b}$$

$$= \boldsymbol{b}$$

$$\Longrightarrow \theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in C$$

مجموعه کوژ- ادامه

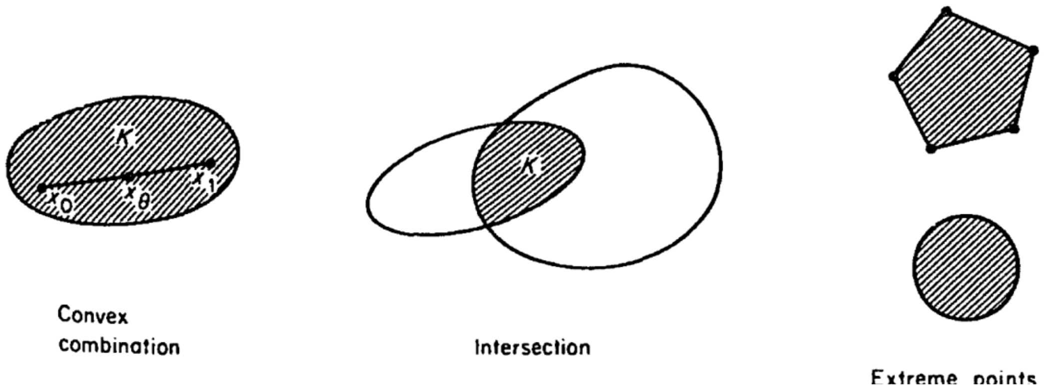
تعمیم تعریف مجموعه کوژ

$$\forall \boldsymbol{x}_i \in S, i = 1, \dots, m \Longrightarrow \boldsymbol{x}_\theta \in S\,, \boldsymbol{x}_\theta = \sum_{i=۱}^{m} \theta_i \boldsymbol{x}_i, \sum_{i=۱}^{m} \theta_i = ۱, \theta_i \geq 0$$

$\boldsymbol{x}_\theta$ ترکیب کوژ نقاط $\boldsymbol{x}_۱$ و $\boldsymbol{x}_۲$ و **…**

پوسته کوژ

- $\boldsymbol{x}_۱$ و $\boldsymbol{x}_۲$ و … و $\boldsymbol{x}_m$
- تمامی نقاط $\boldsymbol{x}_\theta$ مندرج در مجموعه بالا

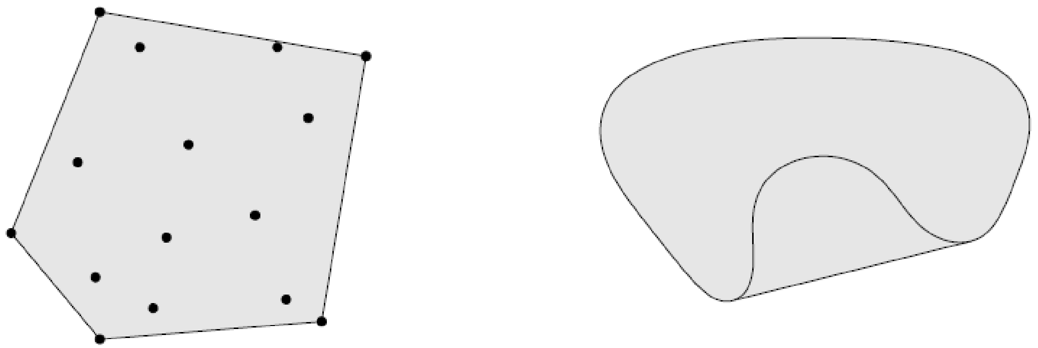


پوستهٔ کوژ

مجموعهٔ ترکیب کوژ تمامی نقاط مجموعه

- $\mathfrak{M} \in \mathbb{R}^n$

$$\text{Kož}\mathfrak{M} = \{\theta_۱ \boldsymbol{x}_۱ + \theta_۲ \boldsymbol{x}_۲ + \cdots + \theta_m \boldsymbol{x_m} | \boldsymbol{x}_۱, \boldsymbol{x}_۲, \dots, \boldsymbol{x_m} \in \mathfrak{M}, \theta_i \geq 0, \forall i = 1, \dots m, \theta_۱ + \theta_۲ + \cdots + \theta_m = ۱\}$$



مخروط

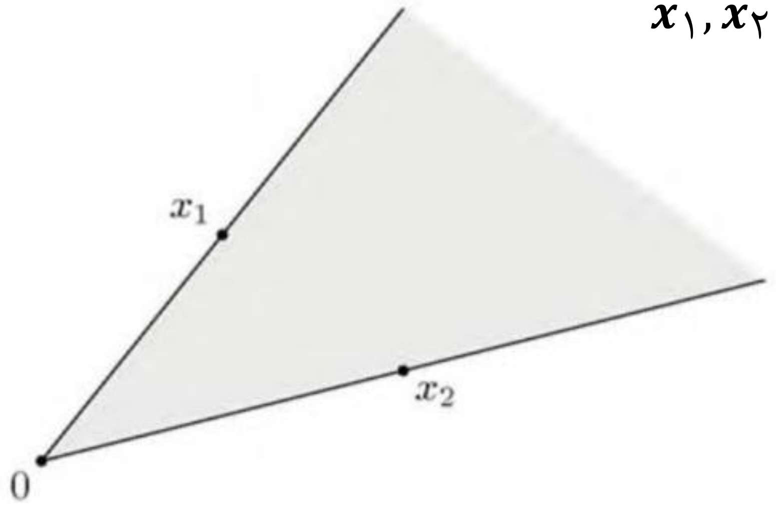
مجموعهٔ مخروط یا ناهمگن نامنفی!؟

- $S \subseteq \mathbb{R}^n$

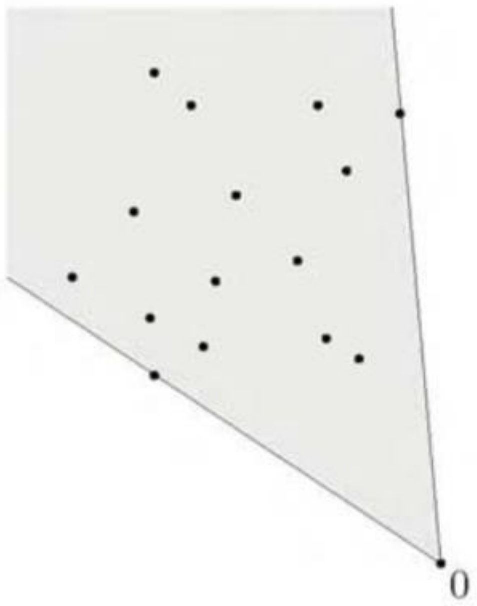
$$\boldsymbol{x} \in S \xRightarrow{\forall \theta \geq 0} \theta \boldsymbol{x} \in S$$

مخروط کوژ: هم مخروط و هم کوژ باشد.

$$\boldsymbol{x}_۱, \boldsymbol{x}_۲ \in C \xRightarrow{\forall \theta_۱, \theta_۲ \geq 0} \theta_۱ \boldsymbol{x}_۱ + \theta_۲ \boldsymbol{x}_۲ \in C$$



مخروط



ترکیب خطی نامنفی یا ترکیب مخروطی

$$\theta_1 x_1 + \cdots + \theta_k x_k \qquad \theta_1, \ldots, \theta_k \geq 0$$

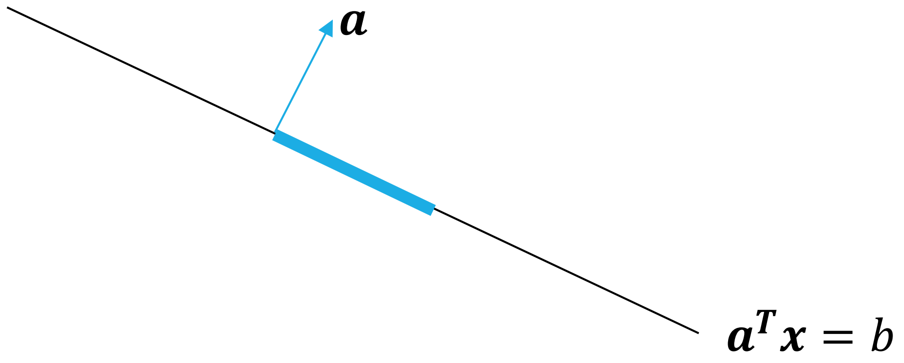
پوستهٔ مخروط

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C,\ \theta_i \geq 0,\ i = 1, \ldots, k\}$$

مثال

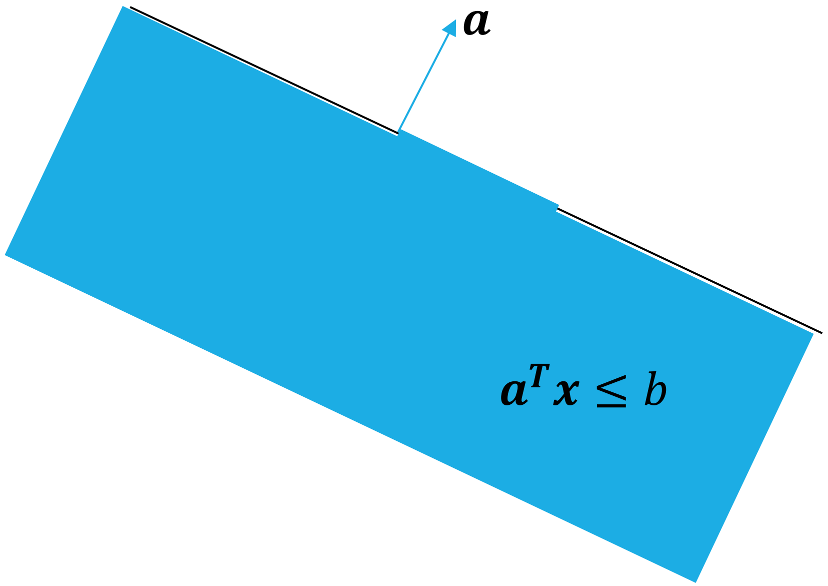
ابرصفحه

- $\{\boldsymbol{x} | \boldsymbol{a}^t \boldsymbol{x} = b, \boldsymbol{a} \neq 0\}$
- کوژ و افین



نیم‌صفحه

- $\{\boldsymbol{x} | \boldsymbol{a}^t \boldsymbol{x} \leq b, \boldsymbol{a} \neq 0\}$
- کوژ ولی غیر افین



مثال

گوی اقلیدسی

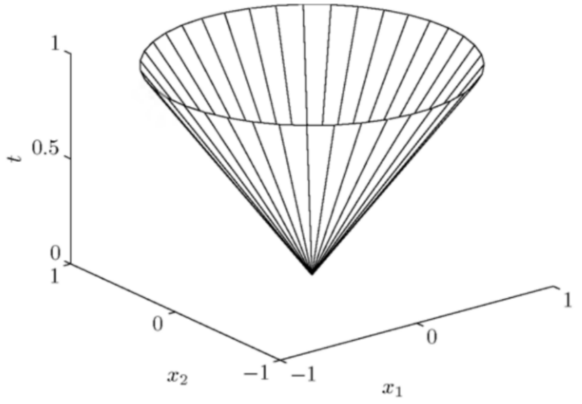
- $\{\mathbf{x} | \|\boldsymbol{x} - \boldsymbol{x}'\| \leq h\} = \{\mathbf{x} \big| (\boldsymbol{x} - \boldsymbol{x}')^T (\boldsymbol{x} - \boldsymbol{x}') \leq r^{۲}\}$

بیضی

- $\{\mathbf{x} \big| (\boldsymbol{x} - \boldsymbol{x}')^T P^{-۱} (\boldsymbol{x} - \boldsymbol{x}') \leq ۱\}$

مخروط نرمال

- $\{\mathbf{x} | \|\boldsymbol{x}\| \leq h\}$



مثال

چندضلعی

- $\{\mathbf{x}|A\boldsymbol{x} \leq b, C\boldsymbol{x} = d\}$

ماتریس مثبت نیمه‌معین

- $\{X \in \mathbb{R}^{n\times n}|X = X^T\}$
- $\{X \in \mathbb{R}^{n\times n}|X = X^T, X \geq 0\}$

بسته بودن خاصیت کوژی تحت عملیات‌ها

اشتراک

توابع افین (فرض کوژی مجموعهٔ S)

- مقیاس αS
- انتقال $S + a$
- تصویر: تصویر یا پراجکشن مجموعه کوژی به چند مورد از مختصاتش کوژ است.
  - $S \subseteq \mathbb{R}^m \times \mathbb{R}^n$ آن‌گاه کوژ بودن $T = \{x_1 \in \mathbb{R}^m | (x_1, x_2) \in S, x_2 \in \mathbb{R}^n\}$
- جمع
  - $S_1 + S_2 = \{x + y | x \in S_1, y \in S_2\}$
- $S_1 \times S_2$

بسته بودن خاصیت کوژی تحت عملیات‌ها

نامساوی خطی ماتریسی

$$f(x) = B + x_1 A_1 + \cdots + x_n A_n \geq 0$$

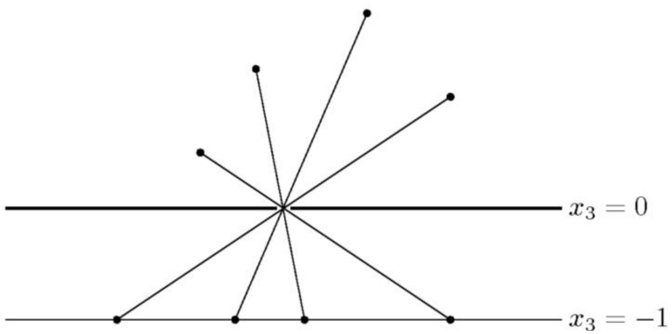
بسته بودن خاصیت کوژی تحت عملیات‌ها

توابع کسری خطی و پرسپکتیو

- تابع پرسپکتیو
- $P: \mathbb{R}^{n+1} \to \mathbb{R}^n$

$$P(\boldsymbol{z}, t) = \boldsymbol{z}/t$$

  - تغییر مقیاس و نرمال سازی بردارها به طوری که آخرین مولفه برابر یک و سپس حذف آن
  - مثال- دوربین $y = -\left(\frac{x_1}{x_3}, \frac{x_2}{x_3}\right) = -P(x)$



- تابع کسری-خطی: ترکیب تابع پرسپکتیو با تابع افین

$$g: \mathbb{R}^n \to R^{m+1}, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^n, d \in \mathbb{R}$$

- $g(\boldsymbol{x}) = \begin{bmatrix} A \\ c^T \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} b \\ d \end{bmatrix}$ تابعی افین
- $f: \mathbb{R}^n \to \mathbb{R}^m$ به طوری که دامنه آن $\{x | \boldsymbol{c}^T \boldsymbol{x} + d > 0\}$ و

$$f(\boldsymbol{x}) = \frac{(A\boldsymbol{x} + \boldsymbol{b})}{\boldsymbol{c}^T \boldsymbol{x} + d}$$

- یا برابر $\mathbb{R}^n$ در صورتی $c = 0$ و $d > 0$ و تبدیل تابع به تابع افین

بسته بودن خاصیت کوژی تحت عملیات‌ها

- تابع کسری-خطی: ترکیب تابع پرسپکتیو با تابع افین

$$g: \mathbb{R}^n \to R^{m+1}, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n, d \in \mathbb{R}$$

- $g(\boldsymbol{x}) = \begin{bmatrix} A \\ c^T \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} b \\ d \end{bmatrix}$ تابعی افین
- $f: \mathbb{R}^n \to \mathbb{R}^m$ به طوری که دامنه آن $\{x | \boldsymbol{c}^T \boldsymbol{x} + d > 0\}$ و

$$f(\boldsymbol{x}) = \frac{(A\boldsymbol{x} + \boldsymbol{b})}{\boldsymbol{c}^T \boldsymbol{x} + d}$$

- یا برابر $\mathbb{R}^n$ در صورتی $c = 0$ و $d > 0$ و تبدیل تابع به تابع افین
- تفسیر دیگر

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \in \mathbb{R}^{(m+1)\times(n+1)}$$

- عمگر روی (x,1) به طوری که منجر به $(A\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{c}^T \boldsymbol{x} + d)$
  - مناظر و مرایا؟! منجر به $(f(x),1)$

بسته بودن خاصیت کوژی تحت عملیات‌ها

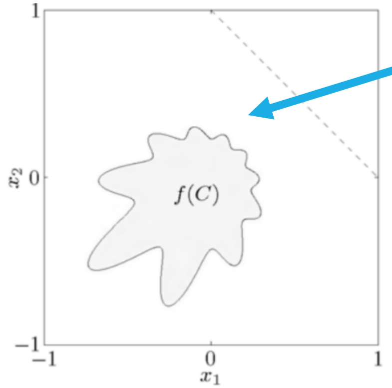
مثال- X و Y دو متغیر تصادفی با گرفتن مقدار از {1,…,n} و {1,…,m} و p_{ij} نمایشگر احتمال $p(X=i,Y=j)$

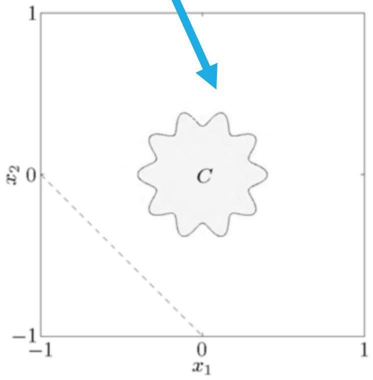
- $f_{ij}=p(X=i|Y=j)$ و برابر با

$$f_{ij}=\frac{p_{ij}}{\sum_{k=۱}^{n} p_{kj}}$$

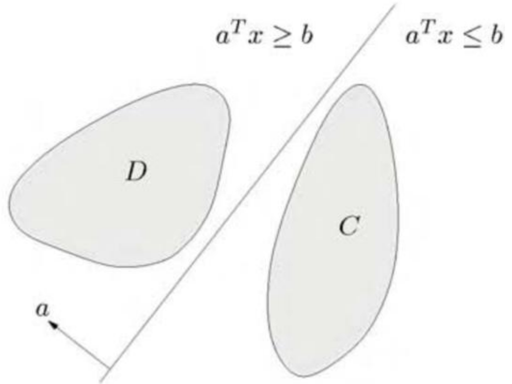
- f نگاشت خطی-کسری از p

مثال- $C\subseteq R^{۲}$ و $f(x)=\frac{x}{x_۱+x_۲+۱}$ و دامنه f برابر $\{(x_1,x_2)|x_1+x_2+1>0\}$، خطچین مرز دامنه





قضیه جدایی ابرصفحه



قضیه جدایی ابرصفحه: اگر دو مجموعه C و D کوژ با اشتراک تهی باشند، ابرصفحه‌ای وجود دارد که آنها را از یکدیگر جدا می‌کند. به دیگر سخن، $a \neq 0$ و b وجود دارد که $a^T x \leq b$ برای تمامی $x \in C$ و $a^T x \geq b$ به ازای تمامی $x \in D$.

فرض c و d دو نقطه در دو مجموعه با کمترین فاصله، آن‌گاه

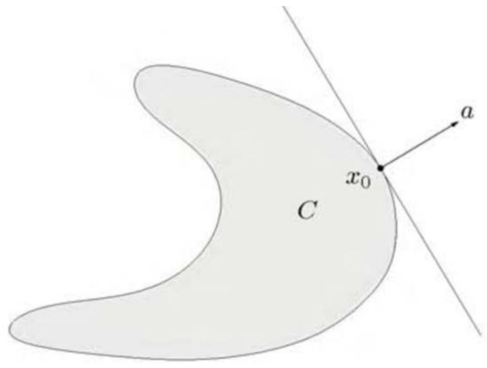
$$\boldsymbol{a} = \boldsymbol{d} - \boldsymbol{c}, b = \frac{\|\boldsymbol{d}\|_2^2 - \|\boldsymbol{c}\|_2^2}{2}$$

$$f(x) = \boldsymbol{a}^T \boldsymbol{x} - b = (\boldsymbol{d} - \boldsymbol{c})^T \left(\boldsymbol{x} - \left(\frac{1}{2}\right)(\boldsymbol{d} + \boldsymbol{c}) \right)$$

نامثبت روی C و نامنفی روی D، به دیگر سخن: جداساز C و D

قضیه پشتیبانی ابرصفحه

قضیه پشتیبانی ابرصفحه: x_0 نقطه‌ای مرزی روی C اگر $a \neq 0$ برآورده‌گر $\forall\ \mathrm{x} \in C: a^T x \leq a^T x_0$، آن‌گاه ابرصفحهٔ $\{x | a^T x = a^T x_0\}$ابرصفحهٔ پشتیبان C



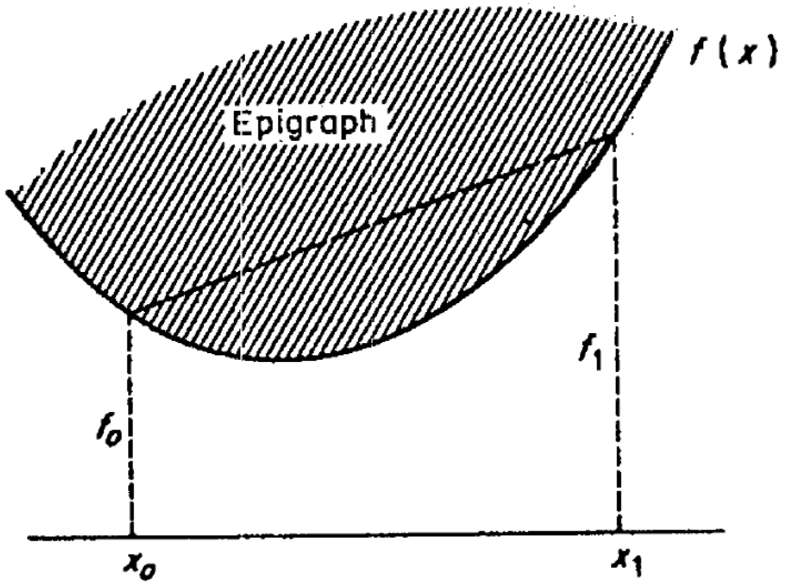
تابع کوژ

تابع کوژ $f: \mathbb{R}^n \to \mathbb{R}$

- دامنهٔ S کوژ و

$$\boldsymbol{x}_۱, \boldsymbol{x}_۲ \in S \Longrightarrow f(\theta \boldsymbol{x}_۱ + (۱-\theta)\boldsymbol{x}_۲) \leq \theta f(\boldsymbol{x}_۱) + (۱-\theta) f(\boldsymbol{x}_۲), \qquad \forall \theta \in [0,۱]$$

- مثال-
  - گوی کوژ $\{\mathbf{x} \in \mathbb{R}^n | \|\boldsymbol{x}\|_۲ \leq ۱\}$
  - چندضلعی‌های تعریفی با مساوی‌ها و نامساوی‌های خطی $\{\mathbf{x} \in \mathbb{R}^n | A\boldsymbol{x} = b, C\boldsymbol{x} \leq d\}$
  - تابع درجه دو کوژ $\mathrm{f}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathrm{H} \mathbf{x}$ با ماتریس مثبت نیمه معین



Graph below chord

- تابع اکیدا کوژ: نامساوی اکید با شرط $\boldsymbol{x}_۱ \neq \boldsymbol{x}_۲, \theta \in (0,۱)$

تابع کوژ گسترده

$$\tilde{f}: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$$

$$\tilde{f}(x) = \begin{cases} f(x), & x \in D(f) \\ \infty, & x \notin D(f) \end{cases}$$

$$\tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

مثال- تابع اندیس مجموعه کوژ

$$\tilde{I}(x) = \begin{cases} 0, x \in K \\ \infty, x \notin K \end{cases}$$

شرط مرتبهٔ اول

فرض f مشتق‌پذیر

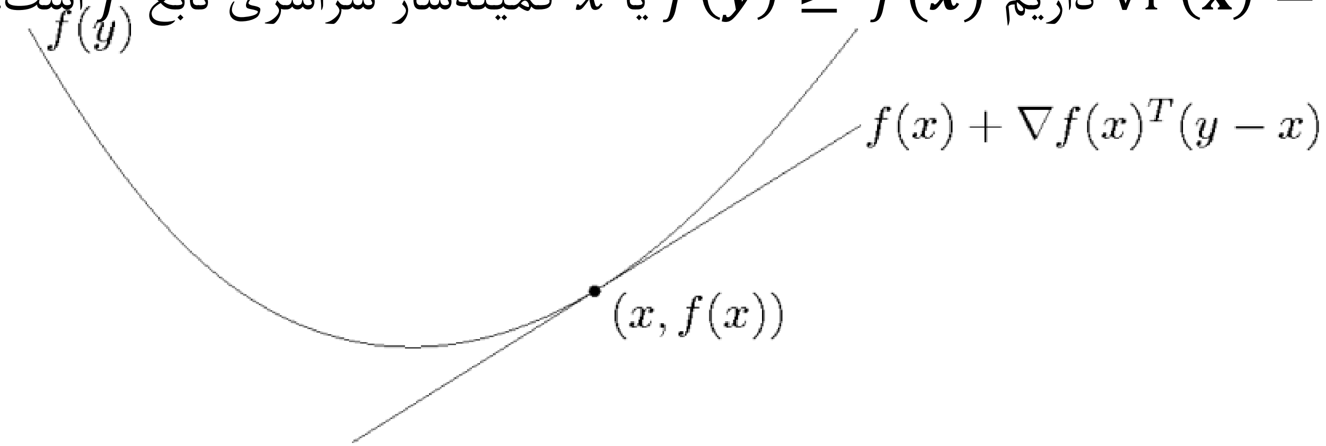
F کوژ است اگر و فقط اگر دامنهٔ **f** کوژ و

$$f(\boldsymbol{y}) \geq \mathrm{f}(\mathbf{x}) + \nabla \mathrm{f}(\mathbf{x})^{\mathrm{T}}(\mathbf{y}-\mathbf{x})$$

به دیگر سخن، تقریب مرتبه اول تیلور تابع زیرتخمین‌گر سراسری تابع کوژ و بالعکس!

در صورتی که $\nabla \mathrm{f}\,(\mathbf{x}) = \mathbf{0}$ داریم $f(\boldsymbol{y}) \geq f(\boldsymbol{x})$ یا x کمینه‌ساز سراسری تابع f است.

اثبات؟



شرط مرتبۀ اول

اثبات: شرط لازم

فرض **f** کوژ است و **x** و **y** عضو دامنۀ **f**

- به دلیل کوژی دامنه، داریم

$$0 \leq \theta \leq 1, x + \theta(y - x) \in D(f)$$

- طبق خاصیت کوژی

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

- جابجایی و تقسیم دو طرف بر θ

$$f(y) \geq f(x) + \frac{f\big(x+\theta(y-x)\big)-f(x)}{\theta}$$

- با $\theta \to 0$

$$f(y) \geq f(x) + f'(x)(y - x)$$

شرط مرتبهٔ اول

اثبات: شرط کافی

فرض f

$$f(y) \geq f(x) + f'(x)(y - x)$$

و انتخاب x و y عضو دامنهٔ f و $0 \leq \theta \leq 1$

- تعریف متغیر جدید:
- $z = \theta x + (1 - \theta)y$
- با توجه به برقراری شرط بالا:

$$f(x) \geq f(z) + f'(z)(x - z); \qquad f(y) \geq f(z) + f'(z)(y - z)$$

ضرب متقدم در θ و متأخر در $1 - \theta$ و جمع هر دو منجر به نتیجه

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

مثال‌های دیگری از توابع کوژ

تمامی توابع خطی و افین هم کوژ (و هم کاو)

e^{ax} بر اعداد حقیقی و به ازای هر a

x^a روی اعداد حقیقی مثبت و $1 \geq$ a و یا $0 \leq a$ در غیر این صورت؟

$|x|^p$ به ازای $۱ \geq p$، کوژ روی اعداد حقیقی

لگاریتم کاو روی اعداد حقیقی مثبت

تابع $xlogx$ (منفی تابع انتروپی)

- مشتق‌ها
- Log x + 1
- $۱/x$ و بزرگتر از صفر به ازای x مثبت
- پس منفی تابع انتروپی تابعی کوژ است!

مثال‌های دیگری از توابع کوژ

$$f(x) = \frac{۱}{۲}\,\boldsymbol{x}^T\,Q\boldsymbol{x} + \boldsymbol{q}^T\boldsymbol{x} + r$$

نرم‌ها

تابع بیش

$$\frac{x^۲}{y}$$

$$\ell og(e^{x_1} + e^{x_2} + \cdots + e^{x_n})$$

- تمرین اثبات کنید.

شرط مرتبهٔ دوم کوژی

تابع **f** با دامنهٔ کوژ و ماتریس هسی آن مثبت نیمه معین است.

$$\nabla^{۲} f(x) \geq 0$$

مجموعه زیرسطح

f تابع کوژ، آنگاه مجموعه زیرسطح زیرمجموعه‌ای کوژ است.

$$C_\alpha = \{x \in \mathrm{D}(f) | f(x) \le \alpha\}$$

- به ازای x و y عضو دامنه f

$$f(x) \le \alpha, f(y) \le \alpha$$

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) \le t\alpha + (1-t)\alpha = \alpha$$

دود اگر بالا نشیند، واصل کوژ

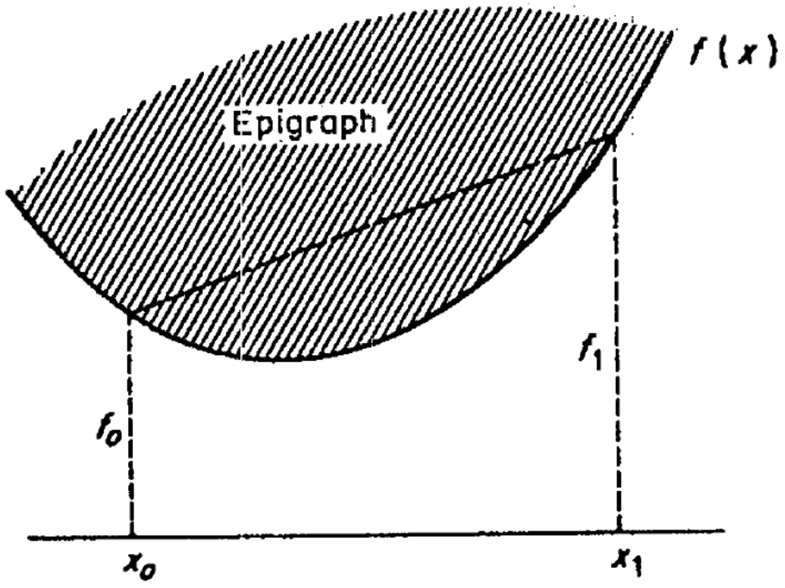
اپی‌گراف (بالای گراف)

$$epi\ f = \{(\boldsymbol{x}, t) | \boldsymbol{x} \in D(f), f(\boldsymbol{x}) \leq t\}$$

زیرمجموعه‌ای از $\mathbb{R}^{n+1}$

واصل مجموعه و تابع کوژ

- تابعی کوژ است اگر بالای گرافش مجموعه کوژ باشد.



Graph below chord

نامساوی‌ها

نامساوی ینسن

$$f\left(\frac{x+y}{۲}\right) \leq \frac{f(x)+f(y)}{۲}$$

$$f(\theta x + (۱-\theta)y) \leq \theta f(x) + (1-\theta)f(x)$$

$$f\left(\int_S p(x)dx\right) \leq \int_S f(x)p(x)dx$$

$$f(E[X]) \leq \mathrm{E}[\mathrm{f}(X)]$$

نامساوی‌ها

استفاده از نامساوی ینسن

$$f\left(\frac{x+y}{۲}\right) \leq \frac{f(x)+f(y)}{۲}$$

اعمال آن بر تابع کوژ $-\log x$

$$-\log\left(\frac{x+y}{۲}\right) \leq \frac{-\log x - \log y}{۲}$$

اعمال منفی بر دو طرف

$$\log\left(\frac{x+y}{۲}\right) \geq \frac{\log x + \log y}{۲}$$

$$\log\left(\frac{x+y}{۲}\right) \geq \log\sqrt{xy}$$

دو طرف را به توان نمایی رساندن منجر به

$$\frac{x+y}{۲} \geq \sqrt{xy};\mathrm{x,y} \geq 0$$

نامساوی‌ها

نامساوی هولدر

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$$

بسته بودن تابع کوژ تحت عملیات‌ها

جمع وزن‌دار نامنفی

- f کوژ، آنگاه αf نیز کوژ و $\alpha \geq 0$
- f_i کوژ و $\alpha_i \geq 0$، آن‌گاه $\sum_{i=1}^{m} \alpha_i f_i$ نیز کوژ

ترکیب با نگاشت افین: f کوژ در نتیجه g به صورت زیر هم کوژ، البته با شرط برقراری دامنه

$$f: \mathbb{R}^n \to \mathbb{R}$$
$$g: \mathbb{R}^m \to \mathbb{R}$$
$$A \in \mathbb{R}^{n \times m}, \boldsymbol{b} \in \mathbb{R}^m$$
$$g(\boldsymbol{x}) = f(A\boldsymbol{x} + \boldsymbol{b})$$

بسته بودن تابع کوژ تحت عملیات‌ها

بیش نقطه به نقطه

$$f(\boldsymbol{x}) = max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$$

- دامنه f برابر اشتراک دو دامنهٔ دو تابع $f_۱$ و $f_۲$

$$\begin{aligned} f(\theta x + (۱-\theta)y) &= \max\{f_۱(\theta x + (۱-\theta)y), f_۲(\theta x + (۱-\theta)y)\} \\ &\leq \max\{\theta f_۱(x) + (۱-\theta)f_۱(y), \theta f_۲(x) + (۱-\theta)f_۲(y)\} \\ &\leq \theta \max\{f_۱(x), f_۲(x)\} + (۱-\theta)\max\{f_۱(y), f_۲(y)\} \\ &= \theta f(x) + (۱-\theta)f(y) \end{aligned}$$

$$f(\boldsymbol{x}) = max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x})\}$$

بسته بودن تابع کوژ تحت عملیات‌ها

ترکیب

$$h: \mathbb{R}^k \to \mathbb{R}$$
$$g: \mathbb{R}^n \to \mathbb{R}^k$$
$$f = h\big(g(\boldsymbol{x})\big): \mathbb{R}^n \to \mathbb{R}$$

ترکیب تک‌عددی (اسکالر)

$$h: \mathbb{R} \to \mathbb{R}$$
$$g: \mathbb{R}^n \to \mathbb{R}$$

بسته بودن تابع کوژ تحت عملیات‌ها

ترکیب تک‌عددی (اسکالر)

$$h: \mathbb{R} \to \mathbb{R}$$
$$g: \mathbb{R}^n \to \mathbb{R}$$

تحویل مسئله کوژی f به $f'' \geq 0$. چرا؟

$$f''(x) = h''\big(g(x)\big)g'(x)^2 + h'\big(g(x)\big)g''(x)$$

فرض: g کوژ؟ $g'' \geq 0$، h کوژ و غیرنزولی؟ $h' \geq 0$ و $h'' \geq 0$

نتیجه: $f'' \geq 0$ و f کوژ

به طریق مشابه

- h کوژ و غیرنزولی و g کوژ، آن‌گاه f کوژ
- h کوژ و غیرصعودی و g کاو، آن‌گاه f کوژ
- h کاو و غیرنزولی و g کاو، آن‌گاه f کاو
- h کاو و غیرصعودی و g کوژ، آن‌گاه f کاو

امکان تعمیم به حالت گسترده h یا $\tilde{h}$

بسته بودن تابع کوژ تحت عملیات‌ها

مثال-

- g کوژ $\Leftarrow$ $e^{g(x)}$ کوژ
- g کاو و مثبت $\Leftarrow$ $\log\big(g(x)\big)$ کاو
- g کاو و مثبت$\Leftarrow$ $1/g(x)$ کاو
- g کوژ و نامنفی و $p \geq 1$ آن‌گاه $g(x)^p$ کوژ

بسته بودن تابع کوژ تحت عملیات‌ها

ترکیب برداری

$$h: \mathbb{R}^k \to \mathbb{R}$$
$$g_i: \mathbb{R}^n \to \mathbb{R}$$
$$f(x) = h\big(g(x)\big) = h\big(g_۱(\boldsymbol{x}), g_۲(\boldsymbol{x}), \cdots, g_k(\boldsymbol{x})\big): \mathbb{R}^n \to \mathbb{R}$$

مشتق دوم

$$f''(x) = g'(x)^T \nabla^2 h\big(g(x)\big) g'(x) + \nabla h\big(g(x)\big)^T g''(x)$$

به طریق مشابه

- h کوژ و غیرنزولی در هر یک از اجزای بردار و g_i کوژ، آن‌گاه f کوژ
- h کوژ و غیرصعودی در هر یک از اجزای بردار و g_i کاو، آن‌گاه f کوژ
- h کاو و غیرنزولی در هر یک از اجزای بردار و g_i کاو، آن‌گاه f کاو

بسته بودن تابع کوژ تحت عملیات‌ها

مثال- $z \in \mathbb{R}^k$

$h(z) = z_{[1]} + \cdots + z_{[r]}$ جمع r مولفهٔ بزرگ z

- h کوژ و غیرنزولی
- فرض g_1 تا g_k توابع کوژ روی R^n. آن‌گاه h(g(x)) جمع r تابع g_i بزرگ کوژ

مثال- $h(z) = \log\left(\sum_{i=1}^{k} e^{z_i}\right)$ کوژ و غیرنزولی در هر آرگومان در نتیجه کل تابع کوژ

بهینه‌سازی کوژ- کمترین مربعات

$$\text{کم } f_0(\boldsymbol{x}) = ||A\boldsymbol{x} - \boldsymbol{b}||_۲^۲ = \sum_{i=۱}^{k} (\boldsymbol{a}_i^T\boldsymbol{x} - b_i)^۲, A \in \mathbb{R}^{k\times n}, k \geq n, \boldsymbol{x} \in \mathbb{R}^n,$$

کمترین مربعات وزن‌دار

$$\sum_{i=۱}^{k} \omega_i (\boldsymbol{a}_i^T\boldsymbol{x} - b_i)^۲ , \omega_i > 0$$

- تعیین ω_i نمایشگر سطوح متفاوت تاثیر بر $\boldsymbol{a}_i^T\boldsymbol{x} - b_i$ متناظر
- در تخمین‌های آماری، تقریب بردار مجهول در اندازه‌گیری‌های خطی تحت تاثیر خطا

تنظیم regularization

$$\sum_{i=۱}^{k} (\boldsymbol{a}_i^T\boldsymbol{x} - b_i)^۲ + \rho \sum_{i=۱}^{n} x_i^۲ , \rho > 0$$

بهینه‌سازی کوژ- برنامه‌ریزی خطی

$$\text{کم}\ \boldsymbol{c}^T\boldsymbol{x}\,,$$
$$\boldsymbol{a}_i^T\boldsymbol{x} \le b_i, i = 1, \cdots, m$$

- خطی بودن تابع هدف و توابع قیود و $\boldsymbol{a}_i \in \mathbb{R}^n$ و $b_i \in \mathbb{R}$

بهینه‌سازی کوژ- برنامه‌ریزی خطی

$$\text{کم } \boldsymbol{c}^T\boldsymbol{x},$$
$$\boldsymbol{a}_i^T\boldsymbol{x} \le b_i, i = 1, \cdots, m$$

- خطی بودن تابع هدف و توابع قیود و $\boldsymbol{a}_i \in \mathbb{R}^n$ و $b_i \in \mathbb{R}$
- عدم وجود راه‌حل تحلیلی راحت

بهینه‌سازی کوژ- برنامه‌ریزی خطی

$$\text{کم}\ \boldsymbol{c}^T\boldsymbol{x}\,,$$
$$\boldsymbol{a}_i^T\boldsymbol{x} \leq b_i, i = 1,\cdots,m$$

- خطی بودن تابع هدف و توابع قیود و $\boldsymbol{a}_i \in \mathbb{R}^n$ و $b_i \in \mathbb{R}$
- عدم وجود راه‌حل تحلیلی راحت
- وجود روش‌های موثر برای حل آنها اعم از سیمپلکس دانتزیک، نقطه درونی،
- دارای پیچیدگی مجانبی $n^۲m$ با ثابتی دارای پیچش بیشتر نسبت به راه‌حل کمترین مربعات

کاربردهای برنامه‌ریزی خطی

- مثال مسئله تقریب چبیشف

$$\min \max_{i=۱,\cdots,k} \left|\boldsymbol{a}_i^T\boldsymbol{x} - b_i\right|$$

  - قرابت و تفاوت آن با کمترین مربعات؟
  - امکان حل آن با برنامه‌ریزی خطی؟

بهینه‌سازی کوژ- برنامه‌ریزی خطی

$$\text{کم } \boldsymbol{c}^T\boldsymbol{x},$$
$$\boldsymbol{a}_i^T\boldsymbol{x} \leq b_i, i = 1,\cdots,m$$

- خطی بودن تابع هدف و توابع قیود و $\boldsymbol{a}_i \in \mathbb{R}^n$ و $b_i \in \mathbb{R}$
- عدم وجود راه‌حل تحلیلی راحت
- وجود روش‌های موثر برای حل آنها اعم از سیمپلکس دانتزیک، نقطه درونی،
- دارای پیچیدگی مجانبی $n^۲m$ با ثابتی دارای پیچش بیشتر نسبت به راحل کمترین مربعات

کاربردهای برنامه‌ریزی خطی

- مثال مسئله تقریب چبیشف

$$\min \max_{i=۱,\cdots,k} |\boldsymbol{a}_i^T\boldsymbol{x} - b_i|$$

- قرابت و تفاوت آن با کمترین مربعات؟
- امکان حل آن با برنامه‌ریزی خطی

$$\text{کم } t,$$
$$\boldsymbol{a}_i^T\boldsymbol{x} - t \leq b_i, i = 1,\cdots,k$$
$$-\boldsymbol{a}_i^T\boldsymbol{x} - t \leq -b_i, i = 1,\cdots,k$$
$$t \in R$$

بهینه‌سازی کوژ

$$\text{کم } f_0(\boldsymbol{x})\,,$$
$$f_i(\boldsymbol{x}) \le b_i, i = ۱, \cdots, m$$

تمامی توابع محدب -- واجد شرایط زیر

$$f_i(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \le \alpha f_i(\boldsymbol{x}) + (1-\alpha) f_i(\boldsymbol{y}), \alpha \ge 0$$

حل مسائل بهینه‌سازی کوژ

- نبود تدوین تحلیل عمومی حل مسائل بهینه‌سازی کوژ
- وجود روش‌های موثر برای حل چنین مسائلی
- روش‌های نقطه داخلی
- ادامه تحقیقات در این مسئله بر خلاف برنامه‌ریزی خطی و کمترین مربعات
- پیچیدگی تشخیص تابع کوژ و سختی بیشتر روش‌های تبدیل به بهینه‌سازی کوژ
- تحقیق: روش‌ها و بهبودهای متأخر روش نقطه درونی

بهینه‌سازی کوژ- برنامه‌ریزی درجه دو

$$\min \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x}$$
$$A\boldsymbol{x} = \boldsymbol{b}$$
$$G\boldsymbol{x} \le \boldsymbol{d}$$

تابع هدف درجه دو کوژ به دلیل Q مثبت نیمه معین

مجموعه شدنی: چندضلعی و در نتیجه مجموعه کوژ

$\boldsymbol{x}$

بهینه‌سازی کوژ- برنامه‌ریزی درجه دو با محدودیت درجه دو

$$\min \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x}$$

$$\frac{1}{2}\boldsymbol{x}^T Q_i\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} \le r_i$$

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$G\boldsymbol{x} \le \boldsymbol{d}$$

تابع هدف درجه دو کوژ به دلیل Q مثبت نیمه معین

مجموعه شدنی: بیضی در صورت مثبت نیمه معین بودن Q_i

بهینه‌سازی کوژ- برنامه‌ریزی مخروط درجه دو

$$\min \boldsymbol{c}^T \boldsymbol{x}$$
$$\|A_i x + b_i\| \leq c_i x + d_i$$
$$F\boldsymbol{x} = \boldsymbol{g}$$

نامساوی‌ها: قیدهای مخروط درجه دو

عمومی تر از برنامه‌ریزی درچه دو مقید به تابع درجه دو

بهینه‌سازی کوژ- برنامه‌ریزی نیمه معین

$$\min \boldsymbol{c}^T \boldsymbol{x}$$
$$x_۱ F_۱ + x_۲ F_۲ + \cdots + x_n F_n + F_0 \leq \boldsymbol{0}$$

نامساوی‌های خطی ماتریسی

عمومی‌تر از برنامه‌ریزی خطی، درجه دو و مخروطی درجه دو

کاربرد: کمینه‌سازی مقدار ویژه، کمینه‌سازی نرم ماتریس

شرط لیپشوتز

$$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|$$

بهینه‌سازی کوژ با قید مساوی

$$\min_{\boldsymbol{x}} f(\boldsymbol{x})$$
$$A\boldsymbol{x} = \boldsymbol{b}$$

f کوژ و دو بار مشتق‌پذیر و $A \in \mathbb{R}^{m\times n}$ با رتبهٔ A برابر p

شرط بهینگی

$$\nabla f(\boldsymbol{x}^*) + A^T \nu^* = \boldsymbol{0}$$

بهینه‌سازی کوژ - درجه دو با قید تساوی

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x}$$
$$A\boldsymbol{x} = \boldsymbol{b}$$

f کوژ و دو بار مشتق‌پذیر و $A \in \mathbb{R}^{m\times n}$ با رتبهٔ A برابر p

شرط بهینگی

$$Q\boldsymbol{x}^* + \boldsymbol{c} + A^T\nu^* = \boldsymbol{0}$$
$$A\boldsymbol{x}^* = \boldsymbol{b}$$

یا

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{v}^* \end{bmatrix} = \begin{bmatrix} -\boldsymbol{c} \\ \boldsymbol{b} \end{bmatrix}$$

ماتریس ک‌ک‌ت $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}$

بهینه‌سازی کوژ - درجه دو با قید تساوی

لم- $A \in \mathbb{R}^{m\times n}$ با رتبهٔ کامل و Qمثبت معین است. آنگاه ماتریس ک‌ک‌ت $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}$ ناتکین و دارای پاسخ یگانه است $\begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{v}^* \end{bmatrix}$ است.

بهینه‌سازی کوژ با قید نامساوی

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{q}^T \boldsymbol{x}$$
$$A\boldsymbol{x} = \boldsymbol{a}$$
$$B\boldsymbol{x} \leq \boldsymbol{b}$$

روش مجموعه فعال

- آغاز از نقطه دلخواه x_0
- یافتن مقدار جدید در هر تکرار $x_{k+1} = x_k + \alpha_k d_k$
  - α_k طول گام و d_k جهت جستجو

موضوعات

برنامه‌ریزی مخروط

بهینگی پرتو

روش‌ها و بهبودهای متأخر روش نقطه درونی

منابع

[نازه‌دل]

[آنتونیو]

[جهان‌شاه‌لو]

[فلچر]

[بوید]