1/

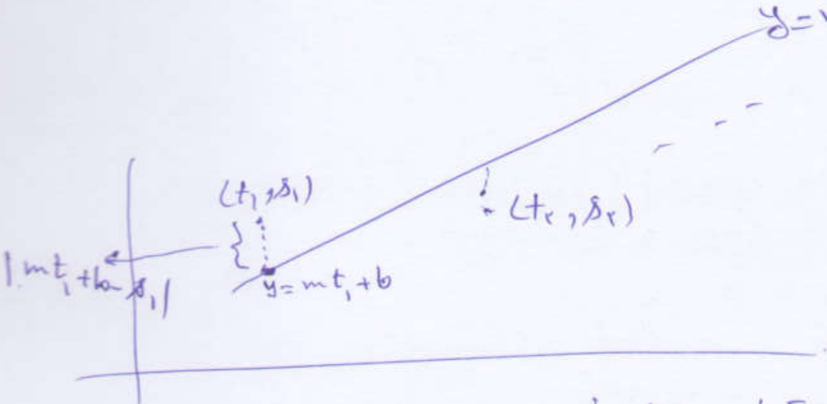
مقدمه‌ای از آنالیز محدب

اطلاعاتی از روش تصویری برای حل معادلات و نامعادلات مانند KKM و Nadler

از فصل ۲ و ~~آنالیز محدب~~ تصویر و فصل ۳ روشها

1.2.1 خط مناسب

مثالی در نظر بگیرید که نقاط $(t_1,s_1), \dots, (t_M,s_M)$ در صفحه داده شده (میتوانند مجموعه‌ای از اطلاعات باشند) میخواهیم مناسب‌ترین خطی را بیابیم که نزدیکترین حالت به این نقاط باشد



فرض کنید مربع فاصله نقطه i ام تا خط را d_i بنامیم پس $d_i=(mt_i+b-s_i)^2$

در این صورت مجموع خطا $\sum_{i=1}^{M} d_i = \sum_{i=1}^{M}(mt_i+b-s_i)^2$

پس مسئله به یافتن خط‌هایی که با (m,b) مشخص می‌گردند است به طوری که عبارت فوق مینیمم شود

$$\begin{cases} \text{minimize} \sum_{i=1}^{M}(mt_i+b-s_i)^2 \\ (m,b)\in\mathbb{R}^2 \end{cases}$$

میتوانیم بجای مربع با قدر مطلق کار کنیم در اینصورت مسئله:

$$\begin{cases} \text{minimize} \sum_{i=1}^{M}|mt_i+b-s_i| \\ (m,b)\in\mathbb{R}^2 \end{cases}$$

که جواب مسئله به کمترین مجموع مربعات مشهور است.

۲/

مثال بعدی: مسائل حمل و نقل است.

از محصولات شیمیایی

فرض کنید دو کارخانه F_1, F_2 محصولات $R_1, \ldots, R_m$ تولید می کند

هر کارخانه که محصولات متفاوتی با یکدیگر تولید می کند. فرض کنید c_i ظرفیت برحسب تن از ($i=1,2$)

دو کارخانه یک محصول و

محصول F_i و $i=1,2$ و d_j برحسب تن تقاضای هفتگی بازار از فروشگاه R_j.

$P_{1j} =$ مبلغ فروش یک تن از محصول F_1 در R_j

$P_{2j} =$ — — — F_2 در R_j

سوال این است که، چقدر محصول از کارخانه های F_1 و F_2 توسط فروشگاه های $R_1, \ldots, R_m$ خریداری شود تا هزینه مینیمم شود. قرار دهید

$x_{1j} =$ مقدار برحسب تن از فروشگاه R_j از محصول F_1

$x_{2j} =$ — — — F_2 —

$\Rightarrow \quad x_{1j}, x_{2j} \geq 0$

جمع هزینه $= P_{1j} x_{1j} + \cdots + P_{mj} x_{mj}$

تقاضای بازار : $x_{1j} + x_{2j} \geq d_j$

$$\sum_{j=1}^{m} x_{1j} \leq c_1 \quad , \quad \sum_{j=1}^{m} x_{2j} \leq c_2$$

بنابراین مسئله:

$$\begin{cases} \text{Minimiz} \sum_{i=1}^{m} P_{ij} x_{ij} \\ x_{1j} + x_{2j} \geq d_j \\ \sum_{j=1}^{m} x_{ij} \leq c_i \\ x_{ij} \geq 0 \end{cases}$$

مثال
موقعیت دسترسی

مسئله قرار گرفتن در جائی مناسب است تا بتوان به راحتی به مکان‌های مشخص رسید. یا به عبارتی از آن نقطه به نقاط دیگر کوتاه‌ترین شود.

اگر برای دو نقطه (a,b) و (x,y) داریم

$$d_1\big((a,b),(x,y)\big) = |a-x|+|b-y|$$

فاصله مستطیلی

حال برای نقاط مشخص $(a_1,b_1),\dots,(a_m,b_m)$ منظور یافتن

$$\underset{(x,y)\in\mathbb{R}^2}{\text{Minimize}} \left(\max |a_j-x|+|b_j-y| \right)$$

فصل ۲ (کتاب) نظریه

— به یاد می‌آوریم برای هر دو نقطه $x,y\in\mathbb{R}^n$، پاره خط واصل این دو نقطه با معادله پارامتری $\lambda x+(1-\lambda)y$ داده می‌شود که $\lambda\in[0,1]$

یا به صورت $\lambda x+\eta y$ که $\begin{cases} 0\le\lambda,\eta \\ \lambda+\eta=1 \end{cases}$ داده می‌شود.

$\{\lambda x+(1-\lambda)y \mid 0\le\lambda\le 1\}$ پاره خط واصل x به y است.

در حالت کلی اگر $x_1,\dots,x_m\in\mathbb{R}^n$ منظور از یک ترکیب محدب از این نقاط نقطه $\sum_{i=1}^{m}\lambda_i x_i$ که در آن

$$\begin{cases} 0\le\lambda_i & i=1,\dots,m \\ \sum_{i=1}^{m}\lambda_i=1 \end{cases}$$

اگر شرط $0\le\lambda_i$ ، $i=1,\dots,m$ را حذف کنیم به هر $\sum_{i=1}^{m}\lambda_i x_i$ یک ترکیب آفین از نقاط گویند. مجموعه‌ای آفین است هرگاه شامل هر ترکیب آفین از نقاطش باشد.

$\emptyset \neq C \subseteq \mathbb{R}^n$ را محدب نامیم هرگاه، شامل پاره خط واصل بین هر دو نقطه خود باشد.

یعنی اگر $u, v \in C$ و $0 \le \lambda \le 1$ آنگاه، $\lambda u + (1-\lambda) v \in C$ ،

تمرین: ۱) نشان دهید C محدب است اگر و تنها اگر شامل هر ترکیب محدب از نقاطش باشد.

۲) زیرمجموعه‌های آفین و زیرفضاها محدب اند.

چند مثال:

در ابتدا برای $x=(x_1,\dots,x_n)$ ، $y=(y_1,\dots,y_n) \in \mathbb{R}^n$ و $x \le y$ تعریف

$i=1,\dots,n$ ، $x_i \le y_i$.

a) اگر $A \in \mathbb{R}^{m\times n}$ یک ماتریس $m\times n$ و $b \in \mathbb{R}^m$ باشند مجموعه‌های ذیل محدب اند

$$L_1 = \{x \in \mathbb{R}^n \mid Ax = b\}$$

$$L_2 = \{x \in \mathbb{R}^n \mid Ax \le b,\ 0 \le x\}$$

(تمرین)

b) خط: اگر $0 \neq v \in \mathbb{R}^n$ ، $x_0 \in \mathbb{R}^n$ آنگاه خط گذرنده از x_0 در جهت v که با

$$L(v,x_0) = \{x \in \mathbb{R}^n \mid x = \lambda v + x_0 \quad \exists \lambda \in \mathbb{R}\}$$

محدب است (تمرین)

c) نیم خط یا اشعه

برای v و x_0 در حالت قبل قرار می‌دهیم

$$L_+(v,x_0) = \{x \in \mathbb{R}^n \mid \exists \lambda \ge 0;\ x = \lambda v + x_0\}$$

نیم خط با شروع از x_0 در جهت v محدب است (تمرین)

5/

د) ابرصفحه: برای $0 \neq a \in \mathbb{R}^n$ و $\gamma \in \mathbb{R}$،

$$H(a,\gamma) := \{x \in \mathbb{R}^n \mid \langle a,x\rangle = \gamma\}$$

ابرصفحه با دو مولفه معلوم a,γ تاسیس می‌شود و a را بردار نرمال آن نامند.

این مجموعه محدب است (تمرین)

ه) نیم‌فضا: با مفروضات حالت قبل قرار دهید

$$H(a,\gamma)^+ := \{x \in \mathbb{R}^n \mid \langle a,x\rangle \geq \gamma\}$$
$$H(a,\gamma)^- := \{x \in \mathbb{R}^n \mid \langle a,x\rangle \leq \gamma\}$$

که نیم‌فضاهای معرفی شده از $H(a,\gamma)$ تاسیس می‌شوند محدب‌اند (تمرین).

و) گوی‌های باز و بسته در $\mathbb{R}^n$ محدب‌اند که در آن

گوی باز $B(x_0,r) := \{x \in \mathbb{R}^n \mid d(x,x_0) < r\}$

گوی بسته $B[x_0,r] := \{x \in \mathbb{R}^n \mid d(x,x_0) \leq r\}$

که در آن اگر $x=(x_1,\dots,x_n)$، $y=(y_1,\dots,y_n)$ آنگاه

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}.$$

تمرین:

یک) اگر $(C_i)_{i\in I}$ خانواده‌ای از مجموعه‌های محدب باشد آنگاه $\bigcap_{i\in I} C_i$ محدب است.

(بنابراین کوچکترین مجموعه محدب شامل مجموعه A وجود دارد و آن $\bigcap_{\substack{C \text{ محدب} \\ A\subseteq C}} C$ است)

دو) اگر C محدب و $a\in\mathbb{R}$ آنگاه
$aC=\{ax \mid x\in C\}$ محدب است

سه) اگر C_1 و C_2 محدب باشند آنگاه

$$C_1+C_2=\{x_1+x_2 \mid x_1\in C_1,\ x_2\in C_2\}$$

محدب است.

چهار) اگر $F:\mathbb{R}^n\to\mathbb{R}^m$ تابع آفین باشد یعنی $F(x)=L(x)+b$ که در آن L تابعی خطی و $b\in\mathbb{R}^m$ آنگاه $F(C)=\{F(x)\mid x\in C\}$ محدب است.
برای C محدب مجموعه

و اگر $G:\mathbb{R}^m\to\mathbb{R}^n$ آفین باشد $G^{-1}(C)=\{x\in\mathbb{R}^m\mid G(x)\in C\}$ محدب است

پنج) برای C گزاره های ذیل معادلند

i) C آفین است

ii) برای هر $x_0\in C$، $C-x_0=\{x-x_0\mid x\in C\}$ زیرفضا است

iii) یک $x_0\in C$ وجود دارد که $C-x_0$ زیرفضا باشد.

شش) فرض کنید $C_1\subseteq\mathbb{R}^n$ و $C_2\subseteq\mathbb{R}^m$ محدب باشند. در این صورت حاصلضرب دکارتی آنها
$$C_1\times C_2=\left\{(x,y)\,\middle|\,\begin{matrix}x\in\mathbb{R}^n\\ y\in\mathbb{R}^m\end{matrix}\right\}$$
محدب است.

هفت) $C\subseteq\mathbb{R}^n$ مخروط به رأس 0 نامیده می‌شود هرگاه برای هر $x\in C$ و هر $0\le a$ رابطه $ax\in C$ برقرار باشد.

نشان دهید اشتراک هر خانواده از مخروطهای به رأس 0 خود مخروطی به رأس 0 است.

اجتماع

جمع دو مخروط

اگر C یک مخروط و $a\in\mathbb{R}$ آنگاه aC مخروط است.

هشت) گزاره های ذیل معادلند

i) C یک مخروط محدب است (یعنی یک مخروط است که محدب نیز می‌باشد)

ii) $C+C\subseteq C$.

غلاف محدب

همانطور که دیدیم غلاف محدب از یک مجموعه A کوچکترین مجموعه محدب شامل A است یعنی

$$\bigcap_{\substack{C \text{ محدب} \\ A\subseteq C}} C$$

که آن را با $Co(A)$ نشان میدهیم.

مثال: برای $\{x_0,\dots,x_p\}\subseteq\mathbb{R}^n$ این نقاط را مستقل آفینی نامند هرگاه $\{x_1-x_0,\dots,x_p-x_0\}$ مستقل خطی باشند. در این صورت به غلاف محدب $\{x_0,\dots,x_p\}$ یک k-سادک گفته می‌شود و آن را سادک k-بعدی نامند. $k=p$

مثلاً

x_1-x_0, x_2-x_0 مستقل خطی

پس ۲-سادک است و غلاف محدب آن مثلث ساخته شده از این سه نقطه است (درون و روی).

قضیه 2.1.1 صفحه ۱۴ از: اگر $A\subseteq\mathbb{R}^n$ آنگاه:

$Co(A)=B=$ مجموعه تمام ترکیبات محدب از A

اثبات: به استقراء میتوان نشان داد که اگر A محدب باشد هر ترکیب محدب از عناصر A به خود A تعلق دارد (تمرین).

بنابراین اگر $A\subseteq M$ محدب باشد پس $B\subseteq M$ (در اینصورت)

$B\subseteq Co(A)$.

از طرفی ~~Co(A)~~ B مجموعه‌ای محدب است و شامل A پس B یکی از اعضای شرکت کننده در تعریف $Co(A)$ است پس $Co(A)\subseteq B$.

علت محدب بودن B: اگر

$z=\lambda_1a_1+\dots+\lambda_pa_p \qquad 0\le\lambda_i\le1,\ \sum\lambda_i=1$

$w=\eta_1b_1+\dots+\eta_sb_s \qquad 0\le\eta_i\le1,\ \sum\eta_i=1$

میتوان فرض کرد $s=p$ در غیر اینصورت حالت نداشته از یکی را آنقدر صفر اضافه میکنیم تا به طول دیگری شود

روی هر جبر فاکتور فون نویمان $\mathcal{A}$ که $\mathcal{A} \subseteq B(\mathcal{H})$ یک نگاشت جمعی $*$- مشتق است. که در آن

$$A \diamond B = AB + BA^*.$$

نویسندگان در مرجع[۱۱] مفهوم مشتق سه تایی کج لی را معرفی کردند. اگر به ازای هر $A, B, C \in \mathcal{A}$ که $[A, B]_* = AB - BA^*$ نگاشت $\Phi : \mathcal{A} \to \mathcal{A}$ یک نگاشت مشتق سه تایی کج لی غیر خطی است اگر $\Phi([[A, B]_*, C]_*) = [[\Phi(A), B]_*, C]_* + [[A, \Phi(B)]_*, C]_* + [[A, B]_*, \Phi(C)]_*$. آن ها نشان دادند اگر نگاشت Φ در شرایط فوق روی جبر فون نویمان صدق کند آن گاه Φ یک نگاشت $*$- جمعی است.

$\lambda \in [0,1]$ دلخواه

$$\lambda z + (1-\lambda) w = \lambda \sum_{i=1}^{p} \lambda_i a_i + (1-\lambda) \sum_{i=1}^{p} \eta_i b_i$$

$$= \sum_{1}^{p} \lambda \lambda_i a_i + \sum_{1}^{p} (1-\lambda) \eta_i b_i$$

ملاحظه می‌کنیم ضرایب مثبت‌اند و جمع ضرایب

$$\sum_{1}^{p} \lambda \lambda_i + \sum_{1}^{p} (1-\lambda) \eta_i = \lambda \sum_{1}^{p} \lambda_i + (1-\lambda) \sum_{1}^{p} \eta_i = \lambda + (1-\lambda) = 1$$

پس ترکیب محدب از عناصر A شد لذا B محدب است.

تابع محدب:

برای $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ داریم

$$\mathrm{dom}\, f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}.$$

تمرین: $\mathrm{dom}(f_1 + f_2) = ?$ اگر $f_1, f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

- تابع $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ محدب نامیده می‌شود هرگاه

$$\forall x, y \in \mathrm{dom}\, f,\ \forall \lambda \in [0,1] \Rightarrow f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$$

- تابع $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ مقعر نامیده می‌شود هرگاه $-f$ محدب باشد.

تمرین: اگر f محدب باشد آنگاه $\mathrm{dom}\, f$ محدب می‌شود.

تمرین (i) اگر $F:\mathbb{R}^n \to \mathbb{R}$ آفین باشد آنگاه محدب است.

(ii) $\|x\| := \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$ محدب است.

(iii) F آفین است اگر و تنها اگر F و $-F$ محدب باشند.

تمرین. فرض کنید $C \subseteq \mathbb{R}^n$. آنگاه گزاره‌های زیر معادلند

i) C محدب است

ii) تابع $\delta_C : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ محدب است که در آن

$$\delta_C(x) = \begin{cases} 0 & x \in C \\ +\infty & x \notin C. \end{cases}$$

لم (2.1.1) ص ۱۴

یک) تابع مشتق‌پذیر $\varphi : \mathbb{R} \to \mathbb{R}$ محدب است اگر و تنها اگر φ' صعودی باشد.

دو) « « دوبار « « « « « $0 \le \varphi''(t)$ $\forall t \in \mathbb{R}$.

اثبات:

یک) اگر φ محدب باشد

$$\varphi(\underbrace{t+\alpha(s-t)}_{\alpha s + (1-\alpha)t}) \le \alpha\varphi(s) + (1-\alpha)\varphi(t) = \alpha[\varphi(s) - \varphi(t)] + \varphi(t)$$

$$\Rightarrow \varphi(t+\alpha(s-t)) - \varphi(t) \le \alpha[\varphi(s) - \varphi(t)]$$

فرض کنید $t<s$

$$\Rightarrow \frac{\varphi(t+\alpha(s-t)) - \varphi(t)}{\alpha} \le \varphi(s) - \varphi(t)$$

$$\Rightarrow \frac{\varphi(t+\alpha(s-t)) - \varphi(t)}{\alpha(s-t)} \le \frac{\varphi(s) - \varphi(t)}{s-t}$$

$$\alpha \to 0^+ \Rightarrow \varphi'(t) \le \frac{\varphi(s) - \varphi(t)}{s-t}$$

$$\xRightarrow[s \to t^+]{} \varphi'(t) \le \varphi'(s).$$

حالت $s<t$ نیز ثابت می‌شود.

حال اگر φ' صعودی باشد

$$\frac{\varphi(t) - \varphi(s)}{t-s} \underset{\exists z}{\overset{\text{مقدار میانگین}}{=}} \varphi'(z) \ge \varphi'(s) \Rightarrow \varphi(t) - \varphi(s) \ge \varphi'(s)(t-s)$$

بین s و t

$s<z<t$

صعودی بودن

$$\frac{\varphi(s)-\varphi(t)}{s-t} = \frac{\varphi(t)-\varphi(s)}{t-s} = \varphi'(z) \le \varphi'(t)$$

منفی $s-t$ ضرب $\Rightarrow$ $\boxed{\varphi(s)-\varphi(t) \ge \varphi'(t)(s-t)}$

$$0 \le \left(\varphi'(s)-\varphi'(t)\right)(s-t)$$

دو رابطه را جمع

رابطه صعودی بودن.

بالعکس: اگر φ' صعودی باشد

$s,t \in \mathbb{R}$

$\alpha \in (0,1)$, $t_\alpha := \alpha t + (1-\alpha)s$

$$\begin{cases} \alpha & \varphi(t_\alpha) \underset{\text{میانگین}}{=} \varphi(t) + \varphi'(\theta)\,(t_\alpha - t) \\ 1-\alpha & \varphi(t_\alpha) = \varphi(s) + \varphi'(\eta)\,(t_\alpha - s) \end{cases} \qquad \exists\theta: \ t \le \theta \le t_\alpha \le \eta \le s$$

$$\varphi'(\theta) \le \varphi'(\eta)$$

جمع

$$\varphi(t_\alpha) = \alpha\varphi(t) + (1-\alpha)\varphi(s) + \underbrace{\alpha(t_\alpha - t)}_{\alpha(1-\alpha)(s-t)}\varphi'(\theta) + \underbrace{(1-\alpha)(t_\alpha - s)}_{=-\alpha(1-\alpha)(s-t)}\varphi'(\eta)$$

$$= \alpha\varphi(t) + (1-\alpha)\varphi(s) + \underbrace{\alpha(1-\alpha)(s-t)}_{\text{مثبت}}\underbrace{[\varphi'(\theta) - \varphi'(\eta)]}_{\text{منفی}}$$

$$\le \alpha\varphi(t) + (1-\alpha)\varphi(s) + 0$$

پس φ محدب.

اثبات مشتق دوم هم از نتایج مقدمات ریاضی

2.2 عملیات حافظ تحدب از توابع محدب

. یادآوری از مشتق و مشتق‌پذیری

فرض کنید $S \subseteq \mathbb{R}^n$ و $f: \mathbb{R}^n \to \mathbb{R}$.

$S^\circ := \operatorname{int}(S)$

اگر $x_0 \in S^\circ$ گوییم f در x_0 مشتق‌پذیر است اگر یک بردار $\nabla f(x_0)$

که گرادیان f در x_0 نامیده می‌شود و تابع $R: \mathbb{R}^n \to \mathbb{R}$ موجود باشد که

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + R(x - x_0) \quad \forall x \in S$$

که $\lim_{x \to x_0} \frac{R(x-x_0)}{\|x-x_0\|} = 0$

11/

و اگر روی هر نقطه از مجموعه باز S مشتق‌پذیر باشد آنگاه f را روی S مشتق‌پذیر گوییم.

ب) $\ell(x) := f(x_0) + \langle \nabla f(x_0), x-x_0 \rangle$ تقریب مرتبه اول f حول x_0 نامیم.

و ب) R باقیمانده گوییم.

ج) $\frac{\partial}{\partial v} f(x_0) := \lim_{t\to 0} \frac{f(x_0+tv) - f(x_0)}{t}$ مشتق جهتی f در جهت v در نقطه x_0 نامیم.

زیرا که

$$f(x_0+tv) \overset{\text{از تعریف}}{\underset{\text{مشتق}}{=}} f(x_0) + \langle \nabla f(x_0), tv \rangle$$

$$\Rightarrow \quad \frac{f(x_0+tv) - f(x_0)}{t} = \langle \nabla f(x_0), v \rangle$$

$$\Rightarrow \frac{\partial}{\partial v} f(x_0) = \langle \nabla f(x_0), v \rangle .$$

در صورتی که این حد موجود نباشد دنبال حد را با علامت نقیض ⊖ مشتق جهتی و را فراهم بود.

تمرینات 2-2.8 انجام گردد.

تعریف: $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ را اکیداً محدب گوییم اگر برای هر زوج $x, y \in dom f$, $x \neq y$ و $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y)$$

~~$x \neq y$~~ .

تمرین: هر تابع آفین $F: \mathbb{R}^n \to \mathbb{R}$ اکیداً محدب نیست.

لم ۲.۲.۳ فرض کنید $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ و $0 \neq v \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$ و $\varphi(t) = f(x_0 + tv)$ آنگاه:

یک) اگر f محدب $\Leftarrow$ φ محدب.

دو) φ محدب $\Leftarrow$ f محدب.

سه) دو قسمت قبل برای اکیداً محدب نیز درست است.

(کفایت)

$$\varphi(at_1+(1-a)t_2) = f(x_0+(at_1+(1-a)t_2)v)$$
$$= f(a(x_0+t_1v)+(1-a)(x_0+t_2v)) \le a f(x_0+t_1v)+(1-a)f(x_0+t_2v)$$
$$= a\varphi(t_1)+(1-a)\varphi(t_2).$$

بنابراین φ محدب است.

(⇐) اگر $\varphi(t)=f(y+t(x-y))$

$$f(ax+(1-a)y) = \varphi(a) = \varphi(a\cdot 1+(1-a)\cdot 0)$$
$$\le a\varphi(1)+(1-a)\varphi(0)$$
$$= af(x)+(1-a)f(y)$$ ✓

گرادیان و هسیان

قضیه: $f:\mathbb{R}^n\to\mathbb{R}$ که روی یک مجموعه باز $C \subseteq \mathbb{R}^n$ دوبار مشتق‌پذیر است آنگاه:

f محدب $\iff$ $\nabla^2 f(x)$ نیمه معین مثبت باشد

$$\left(\langle \nabla^2 f(x)y, y\rangle \ge 0\right)$$

اثبات:

f محدب $\iff \left(\forall x_0\in\mathbb{R}^n,\ v\ne 0 \in \mathbb{R}^n;\ \varphi(t)=f(x_0+tv) \text{ محدب}\right)$

f دوبار مشتق‌پذیر $\Rightarrow$ φ دوبار مشتق‌پذیر

φ محدب $\Rightarrow \varphi''(t)\ge 0 \quad \forall t\in\mathbb{R}$

$$\varphi''(t)=\langle \nabla^2 f(x_0+tv)v, v\rangle \ge 0 \quad \forall t.$$

کافیست برای $t=0$ که ✓

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

الف) f محدب $\iff f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle \quad \forall y$

و هر x که f در x مشتق‌پذیر

ب) f اکیداً محدب $\iff$ با نابرابری اکید.

اثبات: الف)

$(\Rightarrow)$

اگر $y=x$ با برقرار است. فرض کنید $y\neq x$ باشد.

$$\frac{\partial f(x)}{\partial (y-x)} = \lim_{\alpha\to 0^+} \frac{f(x+\alpha(y-x))-f(x)}{\alpha} = \langle \nabla f(x), y-x \rangle$$

f محدب $\Rightarrow$ $f(x+\alpha(y-x)) \leq \alpha f(y) + (1-\alpha) f(x)$

$$\Downarrow$$

$$\frac{f(x+\alpha(y-x))-f(x)}{\alpha} \leq f(y)-f(x)$$

$\alpha\to 0^+ \Rightarrow \langle \nabla f(x), y-x \rangle \leq f(y)-f(x)$ ✓

$z := \alpha x + (1-\alpha) y$: $(\Leftarrow)$

$$\begin{array}{c|l} \alpha & f(x) \geq f(z) + \langle \nabla f(z), x-z \rangle \\ 1-\alpha & f(y) \geq f(z) + \langle \nabla f(z), y-z \rangle \end{array}$$

جمع.

$\alpha f(x) + (1-\alpha) f(y) \geq f(z) = f(\alpha x + (1-\alpha) y)$ ✓

ب) مشابه است

~~کاربرد در بهینه‌سازی~~:

(P) $\min_{x\in C} f(x)$

f محدب

C محدب

قضیه 2.2.3 $f:\mathbb{R}^n\to\mathbb{R}$ مشتق‌پذیر.

محدب $\iff \langle \nabla f(x)-\nabla f(y), x-y\rangle \geq 0 \quad \forall x,y\in\mathbb{R}^n$

$\varphi(t) := f(x_0+tv)$.

$(\Rightarrow)$ $f(x) \geq f(y) + \langle \nabla f(y), x-y\rangle$

$f(y) \geq f(x) + \langle \nabla f(x), y-x\rangle$

جمع

$$\cancel{f(x)+f(y)} \geq \cancel{f(x)+f(y)} + \langle \nabla f(y)-\nabla f(x), x-y\rangle$$

$$\Rightarrow \langle \nabla f(y)-\nabla f(x), y-x\rangle \geq 0.$$

$(\Leftarrow)$ کافیست φ' صعودی باشد.

$\varphi'(t) = \langle \nabla f(x_0+tv), v\rangle$

از نامساوی فرض

$$0 \leq \langle \nabla f(x_0+tv) - \nabla f(x_0+sv), (t-s)v\rangle$$

$$= (t-s)\langle \nabla f(x_0+tv) - \nabla f(x_0+sv), v\rangle$$

$$= (t-s)(\varphi'(t)-\varphi'(s))$$

این قضیه به طور کلی برای Gهایی که الزاماً گرادیان نیستند برقرار است.

در تذکر 2.2.3 $G:\mathbb{R}^n\to\mathbb{R}^n$ که

$\langle G(x)-G(y), x-y\rangle \geq 0 \quad \forall x,y$ را یکنوا نامند.

2.3 اپی‌گراف و مجموعه تراز

تعریف 2.3.1: برای $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ مجموعه ترازهای پایینی f چنین تعریف

$$S_f[\alpha] := \{x \mid f(x)\leq\alpha\} = f^{-1}(-\infty,\alpha] \quad \text{بسته}$$

$$S_f(\alpha) = \{x \mid f(x)<\alpha\} = f^{-1}(-\infty,\alpha) \quad \text{باز}, \qquad \partial S_f(\alpha) = \{x \mid f(x)=\alpha\} = f^{-1}\{\alpha\} \quad \text{بسته}$$

$A_{ij} = P_iAP_j$ برای $i,j = ۱,۲$ داریم $\mathcal{A} = \Sigma_{i,j=۱,۲}\mathcal{A}_{i,j}$.

چون برای هر $T \in \mathcal{A}$ می توان نوشت

$$\begin{aligned} T &= (P_۱ + P_۲)T(P_۱ + P_۲) \\ &= P_۱TP_۱ + P_۱TP_۲ + P_۲TP_۱ + P_۲TP_۲ \\ &= \Sigma_{i,j=۱,۲}P_iTP_j \\ &= \Sigma_{i,j=۱,۲}T_{ij}. \end{aligned}$$

که در آن $T_{ij} \in \mathcal{A}_{ij}$.

تمرین. اگر f محدب باشد آنگاه: $S_f(\alpha)$, $S_f[\alpha]$, $\gamma S_f(\alpha)$ محدب‌اند.

(تمرین)

تعریف:

$$\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in \mathbb{R}^n\}$$

$$\mathrm{Epi}(f) := \{(x,\alpha) \mid f(x) \leq \alpha\}.$$

بنابراین اگر $x \in \mathrm{dom} f \Leftarrow f(x) < \infty \Leftarrow (x,\alpha) \in \mathrm{Epi}(f)$

گزاره: f محدب $\Longleftrightarrow$ $\mathrm{Epi}(f)$ محدب

اثبات: $(\Rightarrow)$ اگر

$(x_1,a_1),(x_2,a_2) \in \mathrm{Epi}(f)$

$t \in [0,1]$

$$\Rightarrow f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \leq ta_1 + (1-t)a_2$$

$$\Rightarrow (tx_1 + (1-t)x_2, ta_1 + (1-t)a_2) \in \mathrm{Epi}(f)$$

$$t(x_1,a_1) + (1-t)(x_2,a_2)$$

$(\Leftarrow)$ اگر $x_1, x_2 \in \mathrm{dom} f$, $\alpha \in [0,1]$

آنگاه $(x_1, f(x_1)), (x_2, f(x_2)) \in \mathrm{Epi}(f)$

$$\Rightarrow t(x_1, f(x_1)) + (1-t)(x_2, f(x_2)) = (tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in \mathrm{Epi}$$

16

$$\Rightarrow \quad f(tx_1+(1-t)x_2) \le t f(x_1) + (1-t) f(x_2)$$

گزاره: اگر $g_i : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ خانواده‌ای از توابع محدب و آنگاه $g(x) = \sup_{i\in I} g_i(x)$

الف) $\operatorname{dom} g = \bigcap_{i\in I} \operatorname{dom} g_i$

ب) $S_g[\alpha] = \bigcap_i S_{g_i}[\alpha]$ یا $Epi(g) = \bigcap Epi(g_i)$

پ) g محدب است.

اثبات: همه موارد ابتدائی‌اند برای قسمت پ می‌دانیم

g محدب $\iff Epi(g)$ محدب

$$= \bigcap_{i\in I} Epi(g_i)$$

و هر کدام محدب‌اند پس اشتراک محدب.

2.3.1 پشتیبان مجموعه در نقطه مرزی

فرض کنید x_0 یک نقطه مرزی S، $0 \neq a \in \mathbb{R}^n$ است. ابرصفحه با بردار نرمال a گذرنده از نقطه x_0:

$$H := \{x\in\mathbb{R}^n \mid \langle x-x_0, a\rangle = 0\}.$$

این ابرصفحه را ابرصفحه پشتیبان S در x_0 نامند هرگاه:

$$S \subseteq H^+ = \{x\in\mathbb{R}^n \mid \langle x-x_0, a\rangle \ge 0\}$$

یا

$S \subseteq H^-$ ———— ≤ 0

طبیعی است که اگر S در هر دو قرار گیرد پس در H قرار دارد.

پس اگر $H \not\supseteq S$ و پشتیبان S باشد آنگاه آن را پشتیبان سره نامند.

2.5 نتیجه مراجعه کنید (32 pdf)

17

مسئله

$$(P)\quad \min_{C} f$$

که $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ ، $C\subseteq\mathbb{R}^n$ و $C\cap \mathrm{dom} f\neq\emptyset$ را درنظر بگیرید.

هر $\inf_{x\in C} f(x)\in[-\infty,+\infty)$ مقدار بهینه (P) نامیم.

مجموعه جوابهای (P) را با $\mathrm{Argmin}_C f$ نشان می‌دهیم، هر عضو آن را می‌نامیم یا جواب بهینه (P) نامیم.

گزاره 2.5.6. اگر f اکیداً محدب باشد و $C\cap \mathrm{dom} f\neq\emptyset$ آنگاه $\mathrm{Argmin}_C f$ حداکثر یک عضو دارد (تک عضوی).

اثبات:

$\mathrm{Argmin}_C f=\emptyset \longrightarrow$ هیچ عضوی ندارد

$\neq\emptyset \xrightarrow{?}$ باید نشان دهیم فقط یک عضو دارد.

خلف: اگر جواب دو عضو متمایز x, y داشته باشد و $f^*:=\inf_{x\in C} f(x)$

$\rightarrow C\cap \mathrm{dom} f\neq\emptyset \Rightarrow \exists x_0\in C\cap \mathrm{dom} f$

$\Rightarrow f(x_0)<+\infty$

$\Rightarrow f^*\le f(x_0)<+\infty \Rightarrow f^*<+\infty$

و اگر $f^*=-\infty \Leftarrow$ مسئله جواب ندارد ✗.

$\Rightarrow f^*\in\mathbb{R}$ ، $f(x)=f(y)=f^*$.

اکیداً محدب $f \Rightarrow f\Big(\underbrace{\tfrac{x+y}{2}}_{\in C}\Big)<\tfrac12 f(x)+\tfrac12 f(y)=f^*$ ✗.

$f^*\le f\left(\frac{x+y}{2}\right)$

قضیه 2-5-5 (قضیه تصویر)

فرض کنید $C\subseteq\mathbb{R}^n$ بسته ناتهی و محدب است و $z\notin C$ آنگاه:

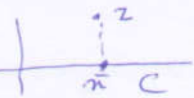
یک) نقطه یکتای $\bar{x}\in C$ وجود دارد که نزدیکترین نقطه به z در C است یعنی جواب یکتای

$$(P)\quad \min_{x\in C}\|x-z\|^2$$

دو) نقطه $\bar{x}$ از نامساوی زیر نیز تشخیص داده می‌شود

$$\langle x-\bar{x},\, z-\bar{x}\rangle\le 0\quad \forall x\in C$$

در این صورت (خود) $\bar{x}$ را تصویر z روی C نامند و آن را با $P_C(z)$ نشان می‌دهند.



اثبات: فرض کن $\tilde{x} \in C$ دلخواه

$$\tilde{C} := \{x \in C \mid \|x-z\| \le \|\tilde{x}-z\|\}$$

$\tilde{C}$ بسته است (اگر $r=\|\tilde{x}-z\|$ آنگاه $\tilde{C} = B_r[z]$) و کراندار (واقع درون دایره‌ای به شعاع r)

بنابراین فشرده است.

از طرفی $f(x) := \|x-z\|^2$ پیوسته است بر مجموعه فشرده، پس مینیمم خود را اختیار می‌کند روی $\tilde{C}$ یعنی

$$(P_1) \quad \min_{x\in\tilde{C}} \|x-z\|^2$$

مسئله جواب دارد. اما f اکیداً محدب، پس جواب یکتا دارد که $\bar{x}$ می‌نامیم.

$$\tilde{C} \subseteq C \Rightarrow \operatorname{Argmin}_{C} f \subseteq \operatorname{Argmin}_{\tilde{C}} f$$

~~(اگر $w \in \operatorname{Argmin}_C f \Rightarrow f(w) = \inf_C f \le \inf_{\tilde{C}} f = \inf_{x\in\tilde{C}} \|x-z\|^2 \le \|\tilde{x}-z\|^2 = f(\tilde{x})$~~

علت: زیرا جوابی که به دنبال آن هستیم f در روی C کمتر شود باید در $\tilde{C}$ نیز باشد چرا که در غیر این صورت از $\tilde{C}$ خارج می‌شود و ...

$\Rightarrow$ ✓

(دوم) هر جواب P_1 هم جواب (P) است زیرا اگر $\bar{x}$ جواب یکتای (P_1) باشد کافی است نشان دهیم

$$(1) \quad f(\bar{x}) \le f(x) \quad \forall x\in C$$

$$f(\bar{x}) = \inf_{C} f \Rightarrow \bar{x} \text{ جواب } (P)$$

علت: (1) برای $x\in\tilde{C}$ که برقرار است

و برای $x\notin\tilde{C}$

$$f(x) = \|x-z\|^2 > \|\tilde{x}-z\|^2 = f(\tilde{x}) \underset{\tilde{x}\in\tilde{C} \text{ چون}}{\ge} \|\bar{x}-z\|^2 = f(\bar{x})$$

نتیجه: چه $x\in\tilde{C}$ و چه $x\notin\tilde{C}$ از پایین $f(\bar{x})$ کراندار است. $\Rightarrow f(\bar{x}) = \inf_C f$

✓

برای قسمت (دو) کافیست نشان دهیم برای نقطه $\hat{x}$

$\hat{x}$ جواب نامساوی مسئله $\Leftrightarrow$ $\hat{x}$ جواب (P)

$(\Leftarrow)$:

$\forall x \in C$

$$\|x-z\|^2 = \|(x-\hat{x})+(\hat{x}-z)\|^2$$

در حقیقت با تساوی

$$\|a+b\|^2 = \langle a+b, a+b\rangle = \langle a,a\rangle + \langle b,b\rangle + 2\langle a,b\rangle = \|a\|^2+\|b\|^2+2\langle a,b\rangle$$

$$= \|x-\hat{x}\|^2 + \|\hat{x}-z\|^2 + 2\underbrace{\langle x-\hat{x}, \hat{x}-z\rangle}_{\geq 0 \text{ از فرض}}$$

$$= \langle x-\hat{x}, z-\hat{x}\rangle \leq 0 \text{ از فرض}$$

$$\geq \|x-\hat{x}\|^2 + \|\hat{x}-z\|^2$$

$$\geq \|\hat{x}-z\|^2 \Rightarrow$$

$$f(x) \geq f(\hat{x}) \Rightarrow f(\hat{x}) = \inf_C f \Rightarrow$$
$\forall x \in C$

$\hat{x}$ جواب یکتای (P) $\Rightarrow \hat{x} = \bar{x}$

$(\Rightarrow)$: اگر $\bar{x}$ جواب (P) باشد طبق ⊙ می‌دانیم باید نشان دهیم

$\bar{x}$ جواب (P) $\Rightarrow$ $f(\bar{x}) = \inf_C f$

$$\|\bar{x}-z\|^2 \leq \|w-z\|^2 \quad \forall w \in C$$

از طرفی

$$w = \bar{x} + a(x-\bar{x}) \quad a \in (0,1)$$

$$\|z-w\|^2 = \|z-\bar{x}-a(x-\bar{x})\|^2$$

$$= \|z-(\underbrace{(1-a)\bar{x}+ax}_{\in C})\|^2 \geq \|z-\bar{x}\|^2$$

$$a\|\bar{x}-x\|^2 - 2\langle z-\bar{x}, x-\bar{x}\rangle + \cancel{\|z-\bar{x}\|^2} \geq \cancel{\|z-\bar{x}\|^2}$$

$$a\|\bar{x}-x\|^2 - 2\langle z-\bar{x}, x-\bar{x}\rangle \geq 0$$

$$\xrightarrow{a \to 0^+} \langle z-\bar{x}, x-\bar{x}\rangle \leq 0 \quad \forall x \in C$$