2.6

زیر گرادیان و زیر دیفرانسیل

تعریف. تابع $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ محدب را در نظر بگیرید. بردار $w\in\mathbb{R}^n$ یک زیرگرادیان f در $x_0\in dom f$ نامیده می‌شود هرگاه

$$f(y)\geq f(x_0)+\langle w, y-x_0\rangle \quad \forall y\in\mathbb{R}^n$$

تذکر: زمانی که f در x_0 مشتق‌پذیر باشد بعداً خواهیم دید که $\nabla f(x_0)\in\partial f(x_0)$ و اینکه $\nabla f(x_0)$ تنها عضو $\partial f(x_0)$ خواهد بود.

در واقع ثابت خواهیم کرد که تک‌عضوی بودن $\partial f(x_0)$ معادل مشتق‌پذیری f در x_0 است.

تمرین: نشان دهید برای تابع محدب f، $\partial f(x_0)$ بسته و محدب است. اگر تابع f پیوسته باشد (همه جا) آنگاه $\partial f(x_0)$ کراندار خواهد بود.

(برای اثبات از اینکه تعریف کرانداری مجموعه‌ها به کمک کراندار بودن را می‌برد استفاده کنید.)

مثال. تابع محدبی $f:\mathbb{R}\to\mathbb{R}$ مثال بزنید که $\partial f(0)=[0,1]$. تابع محدب $f:\mathbb{R}\to\mathbb{R}$ مثال بزنید که $\partial f(0)=[0,+\infty)$.

قضیه 2.6.10: فرض کنید $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ محدب است و $x_0\in dom f$. گزاره‌های زیر معادلند

یک) $\partial f(x_0)\neq\emptyset$

دو) $Epi(f)$ دارای یک ابرصفحه تکیه‌گاه ناعمودی است.

اثبات:

(او ← دو)

$$\exists w\in\partial f(x_0) \rightsquigarrow \langle w, y-x_0\rangle \le f(y)-f(x_0) \quad \forall y$$

$$\le \alpha - f(x_0)$$

اگر $f(y)\le\alpha$

$$\Rightarrow \langle (w,-1), (y,\alpha)-(x_0,f(x_0))\rangle \le 0$$

اگر $H=\{(x,t)\mid \langle (w,-1),(x,t)-(x_0,f(x_0))\rangle=0\}$

آنگاه H یک ابرصفحه تکیه‌گاه است و غیرعمودی است چون $(w,-1)$

یادآوری می‌کنیم که ضرب داخلی (x,t) و (z,s) در $\mathbb{R}^n\times\mathbb{R}$:

$$\langle (z,s),(x,t)\rangle=\langle z,x\rangle+st$$

۲/

(۲ ← (۱ $\exists\,(u,a)$ $\xrightarrow{\text{بنا به}}$ $a\neq 0$

$(x_0,f(x_0))$ مرزی است، نشان می‌دهیم

$$\mathrm{epi}\,f \subseteq H^+ = \left\{(u,t) \;\middle|\; \langle (u,a),(u,t)-(x_0,f(x_0))\rangle \geq 0\right\}$$

$$\Rightarrow \forall (y,\alpha)\in \mathrm{epi}(f) \Rightarrow \langle (u,a),(y,\alpha)-(x_0,f(x_0))\rangle \geq 0$$

$$\langle u, y-x_0\rangle + a\cdot(\alpha - f(x_0)) \geq 0$$

$$\left\langle \frac{u}{a}, y-x_0\right\rangle + (\alpha - f(x_0)) \geq 0$$

$$\alpha \geq f(x_0) + \left\langle \left(-\frac{u}{a}\right), y-x_0\right\rangle \qquad \forall \alpha \geq f(y)$$

$f(y)$ قرار دهید

$$f(y) \geq f(x_0) + \left\langle \left(-\frac{u}{a}\right), y-x_0\right\rangle$$

$$-\frac{u}{a} \in \partial f(x_0) \qquad \boxtimes$$

قضیه 2.6.11 فرض کنید $f:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ محدب، $x_0\in \mathrm{dom} f$. در اینصورت

دو حکم معادلند.

(الف) $w_0\in \partial f(x)$

(ب) $\forall v\in\mathbb{R}^n$ $f'(x_0,v)$ موجود و $f'(x_0,v)\geq \langle v,w_0\rangle$

اثبات:

(الف) ← (ب) اگر $w_0\in\partial f(x_0)$ ، $v\neq 0$

$$f'(x_0,v) := \lim_{t\to 0^+} \frac{f(x_0+tv)-f(x_0)}{t} \geq \lim_{t\to 0^+} \frac{t\langle w_0,v\rangle}{t} = \langle w_0,v\rangle$$

$w_0\in\partial f(x_0)$

(ب) ← (الف) به برهان خلف

$$\Rightarrow w_0 \notin \partial f(x_0)$$

$$\exists\, \bar{y}\in \mathrm{dom} f,\quad \langle w_0, \bar{y}-x_0\rangle \not\leq f(\bar{y})-f(x_0)$$

$$> \qquad '' \qquad ''$$

$f(\bar{y}) < f(x_0) + \langle w_0, \bar{y} - x_0 \rangle$

$\Downarrow$

$f(\bar{y}) - f(x_0) < \langle w_0, \bar{y} - x_0 \rangle \quad (1)$

از طرفی دیگر

$\langle w_0, z \rangle \leq f'(x_0, z) \quad \forall z \neq 0$

$z = y - x_0$

$\langle w_0, \bar{y} - x_0 \rangle \leq f'(x_0, y - x_0) = \lim_{t \to 0^+} \frac{f(x_0 + t(\bar{y} - x_0)) - f(x_0)}{t}$

f محدب است

$\frac{f(x_0 + \bar{y} - x_0) - f(x_0)}{1} = f(\bar{y}) - f(x_0)$

$t \leq 1$ (2)

از (۱) و (۲) تناقض

2.6.7 … $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ … محدب …

$\partial f(x_0) = \{ \nabla f(x_0) \}$

$f(y) - f(x_0) \geq \langle \nabla f(x_0), y - x_0 \rangle = f'(x_0, y - x)$

$\nabla f(x_0) \in \partial f(x_0)$

~~$w_0 \in \partial f(x_0)$ … $\langle w_0, y - x_0 \rangle \leq f(y) - f(x_0)$~~

~~… $\nabla f(x_0)$ …~~

~~$\Rightarrow \langle w_0, w_0 - \nabla f(x_0) \rangle \leq f(\ldots) - f(x_0)$~~

$\exists w_0 \in \partial f(x_0)$ …, $w_0 \neq \nabla f(x_0)$

$\langle w_0, y - x_0 \rangle \leq f(y) - f(x_0)$

$\langle \nabla f(x_0), y - x_0 \rangle \leq f(y) - f(x_0)$

$$\nabla f(x_0) \neq w_0 \in \partial f(x_0) \Rightarrow \langle \nabla f(x_0), v\rangle = f'(x_0, v)$$

$$\geq \langle w_0, v\rangle$$

$$\cancel{\langle w_0 - \nabla f(x_0), v\rangle}$$

$$= \cancel{\langle w_0, v\rangle - \langle \nabla f(x_0), v\rangle}$$

$$\leq \cancel{f(x_0+v) - f(x_0)}$$

$$\Downarrow$$

$$\langle w_0 - \nabla f(x_0), v\rangle \leq 0 \quad \forall v$$

$$w_0 = \nabla f(x_0) \Leftarrow v = 0 \Leftarrow v = w_0 - \nabla f(x_0)$$ از طرفی

2.7 وجود جواب برای مسائل بهینه

در نظر بگیرید مسئله

$$\min_{x \in C} f(x) \quad (*)$$

که $C \subseteq \mathbb{R}^n$ است، و $f: \mathbb{R}^n \to \mathbb{R}$ پیوسته است،

می‌گوییم $x^* \in C$ جواب است که

$$f(x^*) = \min_{x \in C} f(x)$$

$$\Updownarrow$$

$$f(x^*) \leq f(x) \quad \forall x \in C$$

2.7.12 قضیه وایرشتراس

$\emptyset \neq C \subseteq \mathbb{R}^n$ فشرده، و $f: C \to \mathbb{R}$ پیوسته

$(*)$ دارای جواب است.

اثبات:

f پیوسته

پایداری فشردگی

$f(C)$ فشرده در $\mathbb{R}$

بنابر اصل کمال مینیمم دارد

$$\exists x^* \in C, \quad f(x^*) = \min_{x \in C} f(x)$$

---

(۵) کوئرسیو 2·7.11

$f: \mathbb{R}^n \to \mathbb{R}$ اجباری است $\equiv \lim_{\|x\| \to +\infty} f(x) = +\infty$

$\Updownarrow$

$$\forall M>0,\ \exists R>0;\ \|x\|>R \Rightarrow f(x)>M.$$

قضیه: اگر f اجباری باشد آنگاه مجموعه بی‌کران را به بی‌کران تبدیل می‌کند.

علت: (C بی‌کران $\Rightarrow \exists x_n \in C;\ \|x_n\| \to +\infty \Rightarrow f(x_n) \to +\infty \Rightarrow f(C)$ بی‌کران)

تمرین: کدام توابع زیر اجباری‌اند.

الف) $f(x) = \sum_{1}^{n} x_i^2 = \|x\|^2$

ب) $f(x_1, x_2) = x_1^4 + x_2^4 - 3x_1x_2$

ج) $f(x) = e^{\|x\|^2} - (\|x\|^2)^{100}$

تمرین: کدام توابع زیر اجباری نیستند.

الف) $f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2$

ب) $f(x) = c^t x$

مثال: فرض $A \in \mathbb{R}^{n\times n}$ ماتریس متقارن و $f(x) = x^tAx$.

ثابت کنید A معین مثبت است اگر و فقط اگر f اجباری باشد.

($x^tAx \geq \lambda_{\min} \|x\|^2$)

کلافِ مماس .... Tangent bundle

W

التصاق واگنر .... wagner connection

قضیه 2.7.13 فرض کنید $\emptyset \neq C \subseteq \mathbb{R}^n$ و $f:\mathbb{R}^n \to \mathbb{R}$ پیوسته. و فرض کنید

یکی از دو شرط ذیل برقرار باشد

(یک) C بسته و f اجباری

(دو) $\exists \alpha \in \mathbb{R}$ به طوری که $S_f[\alpha] \cap C = \{z \in C \mid f(z) \le \alpha\}$

ناتهی و کراندار است.

آنگاه، $\bar{z} \in C$ وجود دارد به طوری که

$f(\bar{z}) = \inf_{z \in C} f$

اثبات: فرض کنید (یک) برقرار باشد.

$\alpha := \inf_{C} f \Rightarrow \exists z^k \in C, \quad f(z^k) \longrightarrow \alpha$

$\alpha \neq +\infty$ چون از پایین کراندار است

(z^k) کراندار است چون اگر کراندار نباشد و f اجباری است پس $\alpha = +\infty$

در $\mathbb{R}^n$ قضیه بولزانو-وایرشتراس زیردنباله همگرا است

$z^{k_j} \to \bar{z} \Rightarrow f(z^{k_j}) \to f(\bar{z})$

$\alpha = f(\bar{z})$

اگر (دو) برقرار باشد،

ادعا می‌کنیم $\alpha \geq \inf_C f$. اگر نه $\alpha < \inf_C f$

ناتهی $S_f[\alpha]\cap C$ با توجه به توقف

$\phi = S_f[\alpha]\cap C$ $S_f[$

X

$\alpha \geq \inf_C f$

$\alpha = \inf_C f \Rightarrow \{z\in C \mid f(z)\leq\alpha\} \Rightarrow \{z\in C \mid f(z)=\alpha\}$ = مجموعه جواب

$\alpha > \inf_C f \Rightarrow \exists z^k\in C;\ \inf_C f = \lim f(z^k)$

$f(z^k) \leq \alpha$

(z^k) [illegible] f [illegible]

[illegible] [illegible]

خلاصه‌ای از ۲-۸ بدون اثبات:

مسئله بهینه‌سازی بی‌شرط

$$\min_{\mathbb{R}^n} f$$

تعریف: $f:\mathbb{R}^n\to\mathbb{R}$ مشتق‌پذیر روی $D\subseteq\mathbb{R}^n$ باشد. نقطه $x^*\in D$ را نقطه بحرانی f گوییم هرگاه $\nabla f(x^*)=0$.

قضیه (فرمول تیلور) $f:\mathbb{R}^n\to\mathbb{R}$ دوبار مشتق‌پذیر روی C [illegible] و x^* [illegible] یک مجموعه باز [illegible] [illegible] $[x,x^*]$ [illegible]

$\exists z\in[x,x^*];$

$$f(x)= f(x^*)+\nabla f(x^*)^T\cdot(x-x^*)+\frac{1}{2}(x-x^*)^T\nabla^2 f(x^*)(x-x^*) + o(\|x-x^*\|^2)$$

قضیه (شرایط بهینگی مرتبه اول) اگر $f:\mathbb{R}^n\to\mathbb{R}$ مشتق‌پذیر روی مجموعه باز D [illegible] در $x^*\in D$ مینیمم موضعی f باشد آنگاه $\nabla f(x^*)=0$

[illegible] f دوبار مشتق‌پذیر باشد آنگاه $\nabla^2 f(x^*)$ نیمه معین مثبت است.

قضیه ۲.۸.۱۰ فرض کنید ... f دوبار مشتق پذیر باشد آنگاه

$(x-x^*)^T \nabla^2 f(z)(x-x^*) \geq 0 \longrightarrow$ x^* مینیمم سراسری

$\forall x, \quad \forall z \in [x, x^*]$

$'' > 0 \longrightarrow$ مینیمم اکید سراسری

$\leq 0 \longrightarrow$ ماکزیمم سراسری

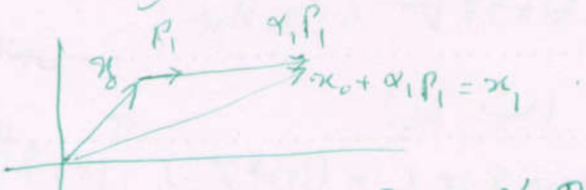
$< 0 \longrightarrow$ ماکزیمم اکید

فصل ۳

۳.۱ روش حل مسائل نامقید

روش monotone (یکنواخت) تکراری یا روش کاهشی تکراری

در این روش معمولاً از یک نقطه شروع می‌کنند برای یافتن نقطه ... در جهت ... حرکت می‌کند



$x_{k+1} := x_k + \alpha_k p^k$

که p^k یک بردار جهت و α_k اندازه گام است به طوری که مسأله ... باشد

$f(x_{k+1}) < f(x_k)$

قضیه ۳.۱.۱۰ فرض کنید $f:\mathbb{R}^n \to \mathbb{R}$ دوبار مشتق پذیر پیوسته باشد و $0 \neq p \in \mathbb{R}^n$ و فرض کنید $\nabla f(x)^T p \neq 0$ آنگاه شرایط زیر معادلند

الف) $0 < \alpha$ وجود دارد به طوری که $f(x+tp) < f(x)$ برای هر $t \in (0,\alpha]$

ب) $\nabla f(x)^T p < 0$

اثبات: ابتدا اگر (لو) درست باشد

$\lambda < 0$ در نظر

z در $[x_0, x_0+\lambda p]$

بسط تیلور

$$f(x+\lambda p) = f(x) + \lambda \nabla f(x)^T p + \frac{\lambda^2}{2} p^T \nabla^2 f(z) p$$

$$= f(x) + \lambda\Big[\underbrace{\nabla f(x)^T p}_{\text{منفی}} + \underbrace{\frac{\lambda}{2} p^T \nabla^2 f(z) p}\Big]$$

چون کراندار است با انتخاب λ کوچک و از $\lambda \in (0, \alpha)$

$$\Rightarrow f(x+\lambda p) < f(x)$$

(۲) اگر (ک) درست باشد

$$f(x+tp) = f(x) + t\Big[\nabla f(x)^T p + \frac{t}{2} p^T \nabla^2 f(z) p\Big]$$

$$[\quad] < 0 \qquad t \in (0, \alpha]$$

حال از $\Leftarrow t \downarrow 0$

$$\Leftarrow \frac{t}{2} p^T \nabla^2 f(z) p \to 0$$

$$\nabla f(x)^T p < 0$$

تعریف 3.1.17: $p \in \mathbb{R}^n$ را که $\nabla f(x)^T p < 0$ جهت نزول باشد

مثال: $f: \mathbb{R}^n \to \mathbb{R}$ که $\nabla f(x) \neq 0$ در این صورت

$p_1 := -\nabla f(x)$ یک جهت نزول است زیرا

$$\nabla f(x)^T (-\nabla f(x)) = -|\nabla f(x)|^2 < 0$$

همچنین اگر $\nabla^2 f(x)$ معین مثبت باشد و $\nabla f(x) \neq 0$ آنگاه

$p_2 := -[\nabla^2 f(x)]^{-1} \nabla f(x)$ یک جهت نزول است چرا که

$$\nabla f(x)^T p_2 = -\underbrace{\nabla f(x)^T [\nabla^2 f(x)]^{-1} \nabla f(x)}_{\text{مثبت}} < 0$$

روش نیوتن 3.1.1

فرض کنید می‌خواهیم $f: \mathbb{R}^n \to \mathbb{R}$ را که دوبار مشتق‌پذیر است کمینه کنیم.

از فرمول تیلور

$$f(x_0+p) = f(x_0) + \nabla f(x_0)^T p + \frac{1}{2} p^T \nabla^2 f(z) p$$

$$\exists z \in [x_0, x_0+p]$$

ایده نیوتن تقریب زدن

$$m_f(p) = f(x_0) + \nabla f(x_0)^T p + \frac{1}{2} p^T \nabla^2 f(x_0) p$$

$$\simeq \quad " \quad\quad " \quad\quad " \; \frac{1}{2} p^T \nabla^2 f(z) p = f(x_0+p)$$

یعنی باید از نقطه x_0 شروع و p را طوری بیابیم که $m_f(p)$ (یعنی تقریب f) به کمترین مقدار برسد، بجای f ... می‌دهیم

ایده:

چون $\nabla m_f(p^*) = 0 \Rightarrow$

$$\nabla f(x_0) + \nabla^2 f(x_0) p^*$$

$$\Rightarrow p^* = -[\nabla^2 f(x_0)]^{-1} \nabla f(x_0)$$

به شرط اینکه $\nabla^2 f(x_0)$ وارون‌پذیر باشد.

گام بعد

$$x_1 := x_0 + p^*$$

$$\vdots$$

$$x_{k+1} := x_k \underbrace{- [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)}_{p^k \text{ می‌نامیم}}$$

دنباله $(x_n)_n$ حاصل که به روش نیوتن مشهور است و p^k را جهت نیوتن گویند.

همانطور که در مثال قبل دیدیم به شرطی که $\nabla^2 f(x_k)$ معین مثبت نباشد ممکن است جهت نزولی نباشد (کم کردن مقدار تابع را نداشته باشد).

قضیه 3.1.29: فرض کنید $A \in \mathbb{R}^{n\times n}$ معین مثبت، $b \in \mathbb{R}^n$، $a \in \mathbb{R}$ و

$$f(x) := a + b^T x + \frac{1}{2} x^T A x$$

آنگاه f محدب و دارای یک مینیمم سراسری یکتا $x^* := -A^{-1}b$ است.

علاوه براین برای هر نقطه اولیه $x_0 \in \mathbb{R}^n$، روش نیوتن به ... در اولین گام همگرا می‌شود، $x_1 = x^*$.

اثبات: قسمت اول تمرین و برای قسمت دوم

$$x_1 = x_0 - [\nabla^2 f(x_0)]^{-1} \nabla f(x_0) = x_0 - A^{-1}(Ax_0 + b) = x_0 - x_0 - A^{-1}b = -A^{-1}b = x^*$$

تمرین: دنباله نیوتن را برای موارد ذیل تحقیق کنید

یک) $f(x_1, x_2)=x_1^4+2x_1^2x_2^2+x_2^4$ با نقطه اولیه $x_0=(1,1)$

دو) $f(t)=t^4-32t+2$ با نقطه اولیه $t_0=3$

در قضیه 3.1.33 (بدون اثبات)

اگر $f:\mathbb{R}^n\to\mathbb{R}$ سه بار مشتق پذیر پیوسته و x^* مینیمم موضعی f و
ماتریس $\nabla^2 f(x^*)$ معین مثبت باشد آنگاه: اگر نقطه شروع x_0 به اندازه کافی نزدیک x^* باشد دنباله تولید شده
از روش نیوتن به x^* همگرا می‌گردد. مرتبه همگرایی حداقل ۲ است.

روش گرادیان (روش سریعترین نزول)

اگر f مشتق پذیر پیوسته و $x_0\in\mathbb{R}^n$. آنگاه الگوریتم سریعترین نزول دنباله (x_k)
با شروع x_0 برای مینیمم سازی f به صورت زیر تعریف می‌شود:

~~جهت کاهش~~ $p=-\frac{\nabla f(x)}{\|\nabla f(x)\|}$ [illegible]

$$x_{k+1}:=x_k-\alpha_k\nabla f(x_k)$$

$$\alpha_k>0$$

که در آن α_k جواب مسئله یک بعدی

$$\min_{\alpha\geq 0} f(x_k-\alpha\nabla f(x^k))$$

است.

این دنباله نیز ... چه اینکه تابع ... بگردد؟ جواب مسئله مینیمم
سازی همگرا است.