98101588 HW5 محمدامین رامی

-1

ابتدا ثابت می‌کنیم اگر 1 برقرار باشد، 2 نیز برقرار است. فرض کنید $y \in \partial f(x)$ آنگاه:

$$x \in \partial f^*(y) \iff \forall \varepsilon \in D_{f^*}: f^*(\varepsilon) - f^*(y) \ge x^T(\varepsilon - y)$$

$$\iff \sup_z \{z^T\varepsilon - f(z)\} - \sup_z \{z^T y - f(z)\} \ge x^T(\varepsilon - y)$$

$$\iff \sup_z \{\varepsilon^T(z-x) - f(z)\} \ge \sup_z \{y^T(z-x) - f(z)\} \quad \text{(I)}$$

چون $y \in \partial f(x)$ است، مقدار $\sup_z \{y^T(z-x) - f(z)\}$ برابر $-f(x)$ است.
از طرفی تابع $\varepsilon^T(z-x) - f(z)$ در $z = x$ مقدار $-f(x)$ را به خود می‌گیرد.
پس رابطه (I) برقرار است و گزاره 1 گزاره 2 را نتیجه می‌دهد.

همچنین چون $y \in \partial f(x)$ داریم:

$$f(z) - f(x) \ge y^T(z-x) \Rightarrow y^T x - f(x) \ge y^T z - f(z)$$

$$\Rightarrow \arg\max_z \{y^T z - f(z)\} = x \Longrightarrow$$ گزاره 3 از 1 نتیجه می‌شود.

داریم: $f^*(y) = \max_z \{y^T z - f(z)\}$ ثابت کردیم مقدار بیشینه تابع $y^T z - f(z)$ در $z = x$ است، مقدار آن $y^T x - f(x)$ است. پس داریم:

$$f(x) + f^*(y) = +f(x) + \langle x, y\rangle - f(x) = \langle x, y\rangle \Longrightarrow$$ گزاره 1 گزاره 4 را نتیجه می‌دهد

فرض کنید 3 درست باشد. ثابت می‌کنیم 1 نتیجه خواهد شد:

$\forall z:$

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

$$\Rightarrow f(z) - f(x) \geq \langle y, z-x\rangle \Rightarrow y \in \partial f(x)$$

به طریق مشابه از 4 گزاره 1 ثابت می‌شود. حال ثابت می‌کنیم از 2 نیز 1 ثابت می‌شود.

$$g(y) = f^*(y) \Rightarrow y \in g^*(x) \qquad g^*(x) = (f^{**})^*(x) = f(x)$$

$$\Rightarrow y \in \partial f(x) \Rightarrow$$ گزاره 2 گزاره 1 را نتیجه می‌دهد

-2

تابع $h(t)$ را به صورت مقابل تعریف می‌کنیم: $h(t) = f(x+tv)$ که v یک جهت دلخواه است. حال تعریف کنید: $t = \lambda z$ واضح است که $z \geq t$ پس داریم:

$$\frac{h(z) - h(0)}{z} \geq \frac{h(t) - h(0)}{t}$$

همچنین داریم:

$$h(t) - h(0) = f(x+tv) - f(x) \geq t\, v^T g$$

که g بردار تک عضوی گرادیان است. پس داریم:

$$\frac{h(z) - h(0)}{z} \geq \frac{h(t) - h(0)}{t} \geq v^T g$$

پس با توجه به تحدب حد زیر وجود دارد:

$$\lim_{t \to 0} \frac{h(t) - h(0)}{t} = v^T g$$

پس داریم:

$$\lim_{t\to 0} \frac{f(x+tv)-f(x)}{t} = g^T v$$

پس f مشتق پذیر است و گرادیان آن در x برابر g است.

-3

برای اثبات، کافی است سلسله استدلال های زیر را به صورت دوطرفه انجام دهیم:

$$g(x) = f(x) - \frac{\alpha}{2}\|x\|^2$$

محدب است $g(x) \iff g(x_2) - g(x_1) \geq \langle s_g(x_1), x_2 - x_1 \rangle$

$$\iff f(x_2) - \frac{\alpha}{2}\|x_2\|^2 - f(x_1) + \frac{\alpha}{2}\|x_1\|^2 \geq \langle s, x_2 - x_1\rangle - \alpha\langle x_1, x_2 - x_1\rangle$$

$$\iff f(x_2) - f(x_1) \geq \langle s, x_2 - x_1\rangle + \frac{\alpha}{2}x_2^T x_2 - \frac{\alpha}{2}x_1^T x_1 - \alpha x_1^T x_2 + \alpha x_1^T x_1$$

$$\iff f(x_2) - f(x_1) \geq \langle s, x_2 - x_1\rangle + \frac{\alpha}{2}\left(x_1^T x_1 - 2x_1^T x_2 + x_2^T x_2\right)$$

$$\iff f(x_2) - f(x_1) \geq \langle s, x_2 - x_1\rangle + \frac{\alpha}{2}\|x_2 - x_1\|^2 \iff$$ حکم

-4

ابتدا ثابت می کنیم اگر f، L-لیپشیتز باشد، $L \geq \|v\|$ است. داریم:

$$L\|w+v-w\| \geq f(w+v) - f(w) \geq v^T(w+v-w)$$

$$\implies L\|v\| \geq v^T v = \|v\|^2 \implies L \geq \|v\|$$

حال ثابت می‌کنیم اگر $\|v\| \le L$ باشد آنگاه f L-لیپشیتز است.

این موضوع را با برهان خلف ثابت می‌کنیم. فرض کنید w و u دو نقطه باشند که داشته باشیم:

$$f(w) \geq f(u) \quad , \quad f(w) - f(u) > L\|w-u\| \, , \quad v \in \partial f(w)$$

آنگاه داریم:

$$f(u) \geq f(w) + v^T(u-w) \quad \text{I}$$

$$f(w) - L\|u-w\| > f(u) \quad \text{II}$$

$$\text{I}, \text{II} \implies f(w) - L\|u-w\| > f(w) + v^T(u-w)$$

$$\implies L\|u-w\| < v^T(w-u) \leq \|v\| \, \|u-w\|$$

$$\implies L < \|v\| \longrightarrow \text{تناقض}$$

پس حکم اصلی درست است و تابع f باید L-لیپشیتز باشد.

-5

-1 فرض کنید $f_i(D_i) = \frac{1}{2}\ln\left(\frac{\sigma_i^2}{D_i}\right)$ آنگاه داریم:

$$\frac{d^2 f_i(D_i)}{d D_i^2} = \frac{1}{2}\,\frac{1}{D_i^2} > 0 \implies f_i(D_i) \text{ is convex.}$$

تابع ثابت 0 هم کانوکس است. ماکسیمم دو تابع کانوکس هم کانوکس است. همچنین مجموع یک سری تابع کانوکس با ضرایب مثبت نیز کانوکس است. پس objective function نیز کانوکس است.

قید مسئله نیز خطی است. پس مسئله کانوکس است.

-2

با استفاده از قضیه KKT مسئله را حل می‌کنیم. داریم:

$$L(D,v) = \sum_{i=1}^{n} \max\left\{\frac{1}{2}\ln\left(\frac{\sigma_i^2}{D_i}\right), 0\right\} + v(1^T D - D_0)$$

$$\Longrightarrow \partial f_i(D_i) + v = 0 \Longrightarrow v = -\partial f_i(D_i)$$

$$\partial f_i(D_i) = \begin{cases} \frac{-1}{2D_i} & D_i < \sigma_i^2 \\ \left[-\frac{1}{2\sigma_i^2}, 0\right] & D_i = \sigma_i^2 \\ 0 & D_i > \sigma_i^2 \end{cases}$$

$$\Longrightarrow v = \begin{cases} \frac{1}{2D_i} & D_i < \sigma_i^2 \\ \left[0, \frac{1}{2\sigma_i^2}\right] & D_i = \sigma_i^2 \\ 0 & D_i > \sigma_i^2 \end{cases}$$

حال مسئله دو حالت دارد.

حالت اول: $D < \sum_{i=1}^{n} \sigma_i^2$ در این حالت، حالت $D_i > \sigma_i^2$ نمی‌تواند اتفاق بیوفتد.

پس داریم:

$$D_i = \begin{cases} \sigma_i^2 & 0 < v < \frac{1}{2\sigma_i^2} \\ \frac{1}{2v} & \frac{1}{2\sigma_i^2} \leq v \end{cases} = \min\left\{\frac{1}{2v}, \sigma_i^2\right\}$$

$$\Longrightarrow \sum_{i=1}^{m} \min\left\{\frac{1}{2v}, \sigma_i^2\right\} = D \quad , \quad D_i = \min\left\{\frac{1}{2v}, \sigma_i^2\right\}$$

پس مسئله به صورت یک مسئله water-Filling است که از آن v به دست می‌آید و سپس D_i ها به دست می‌آیند

حالت دوم: اگر $D \geq \sum_{i=1}^{m} \sigma_i^2$

در این حالت کافی است قرار دهیم: $D_i = \sigma_i^2$ در این صورت، مقدار تابع هزینه، کمترین مقدار ممکن یعنی 0 خواهد شد و همچنین قیدها نیز ارضا شده‌اند.

-6

لاگرانژین به صورت زیر است:

$$L(x, z, \mu) = \sum_k x_k \log\left(\frac{x_k}{y_k}\right) + b^T z - z^T A x + \mu - \mu 1^T x$$

پس برای بهینه کردن داریم:

$$1 + \log\left(\frac{x_k}{y_k}\right) - a_k^T z - \mu = 0$$

$$\Rightarrow x_k = y_k e^{a_k^T z + \mu - 1}$$

حال برای دوگان لاگرانژین داریم:

$$g(z, \mu) = b^T z + \mu - \sum_k y_k e^{a_k^T z + \mu - 1}$$

$$\Rightarrow \text{minimize } b^T z + \mu - \sum_k y_k e^{a_k^T z + \mu - 1}$$

از طرفی اگر روی μ بهینه کنیم، خواهیم داشت:

$$\mu = 1 - \log \sum_k y_k e^{a_k^T z}$$

ساده سازی

$$\Rightarrow \text{minimize } b^T z - \log\left(\sum_k y_k e^{a_k^T z}\right)$$

-7

.1

اگر $x=1$ باشد، توابع $9x$ و $x+8$ فعال هستند.

$$\partial(9x)\Big|_{x=1}=\{9\} \qquad \partial(x+8)\Big|_{x=1}=\{1\}$$

$$\Longrightarrow \partial(\max\{9x, x+8, x^2\}) = [1,9]$$

اگر $x=3$، فقط تابع $9x$ فعال است. پس:

$$\partial(9x)\Big|_{x=3}=\{9\} \Longrightarrow \partial(\max\{9x, x+8, x^2\}) = \{9\}$$

.2 g ساب‌گرادیان تابع در مبدا است اگر و تنها اگر داشته باشیم:

$$f(y)-f(0)\ge g^Ty \iff \|y\|\ge g^T\|y\|$$

واضح است هر برداری g که $\|g\|\le 1$ باشد عضوی از ساب‌گرادیان است. این موضوع از نامساوی کوشی شوارتز به راحتی نتیجه می‌شود.

ادعا می‌کنیم اگر $\|g\|>1$ باشد، g نمی‌تواند ساب‌گرادیان باشد، زیرا اگر باشد، داریم:

$$f(g)\ge g^Tg \Longrightarrow \|g\|\ge\|g\|^2 \Longrightarrow 1\ge\|g\| \Longrightarrow \times$$

پس ساب‌گرادیان تابع در مبدا، مجموعه زیر است:

$$\partial f(0)=\{g \mid \|g\|\le 1\}$$