به نام خدا

بهینه‌سازی محدّب ۱ (۱-۲۵۷۵۶)

تمرین پنجم

ترم بهار ۰۲-۱۴۰۱

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: جمعه ۱۲ خرداد ۱۴۰۲، ساعت ۲۳:۵۹

$(*)$ مسائلی که با ستاره مشخّص شده‌اند امتیازی هستند و حل‌کردن آن‌ها نمره‌ی امتیازی خواهد داشت!

۱ خواص زیرگرادیان

برای هر تابع اکیدا محدب f و هر دو بردار دلخواه $\mathbf{x}$ و $\mathbf{y}$ ثابت کنید ۴ شرط زیر با هم معادلند.

۱. $\mathbf{y} \in \partial f(\mathbf{x})$

۲. $\mathbf{x} \in \partial f^*(\mathbf{y})$

۳. مقدار بیشینه‌ی تابع $g(\mathbf{z}) = \langle \mathbf{z}, \mathbf{y} \rangle - f(\mathbf{z})$ در $\mathbf{z} = \mathbf{x}$ است.

۴. $f(\mathbf{x}) + f^*(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle$

۲ مشتق و زیرگرادیان

تابع محدّب و کران‌دار f را در نظر بگیرید. در کلاس ثابت کردیم اگر تابع f در نقطه‌ی $\mathbf{x}$ مشتق‌پذیر باشد، بردار ∇f تنها عضو مجموعه‌ی زیرگرادیان تابع f در نقطه‌ی $\mathbf{x}$ است. حال عکس این گزاره را ثابت کنید. یعنی ثابت کنید اگر مجموعه‌ی زیرگرادیان تابع f در نقطه‌ی $\mathbf{x}$ تک عضوی باشد، تابع f در نقطه‌ی $\mathbf{x}$ مشتق‌پذیر و مشتق آن همان تک عضو مجموعه‌ی زیرگرادیان است.

۳ توابع قویّاً محدّب

تابع f یک تابع قویاً محدّب با ضریب α است، اگر و تنها اگر تابع $g(\mathbf{x}) = f(\mathbf{x}) - \frac{\alpha}{2}\|\mathbf{x}\|^2$ محدّب باشد.

در این سوال می‌خواهیم یک شرط لازم و کافی برای قویاً محدّب بودن تابع f روی یک مجموعه‌ی محدّب $\mathcal{C}$ بر حسب بردارهای زیرگرادیان بیان کنیم. ثابت کنید شرط لازم و کافی برای اینکه تابع **محدب** $f : \mathbb{R}^n \to \mathbb{R}$ قویاً محدب با ضریب α روی مجموعه‌ی محدّب $\mathcal{C}$ باشد، این است که به ازای هر دو بردار $\mathbf{x}_1$ و $\mathbf{x}_2$ در $\mathcal{C}$ و به ازای هر $\mathbf{s} \in \partial f(\mathbf{x}_1)$ داشته باشیم:

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \langle \mathbf{s}, \mathbf{x}_2 - \mathbf{x}_1 \rangle + \frac{\alpha}{2}\|\mathbf{x}_2 - \mathbf{x}_1\|^2.$$

یا معادلاً به ازای هر $\mathbf{s}_1 \in \partial f(\mathbf{x}_1)$ و $\mathbf{s}_2 \in \partial f(\mathbf{x}_2)$

$$\langle \mathbf{s}_2 - \mathbf{s}_1, \mathbf{x}_2 - \mathbf{x}_1 \rangle \geq \alpha \|\mathbf{x}_2 - \mathbf{x}_1\|^2.$$

۴ توابع لیپشیتز

می‌گوییم تابع $f : \mathbb{R}^n \to \mathbb{R}$، L–لیپشیتز است اگر به ازای هر $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ داشته باشیم:

$$|f(\mathbf{u}) - f(\mathbf{v})| \leq L\|\mathbf{u} - \mathbf{v}\|.$$

حال فرض کنید $\mathcal{A}$ مجموعه‌ای باز و محدّب باشد و تابع محدّب $f : \mathcal{A} \to \mathbb{R}$ را در نظر بگیرید. ثابت کنید که تابع f، L–لیپشیتز است، اگر و فقط اگر برای هر $\mathbf{w} \in \mathcal{A}$ و $\mathbf{v} \in \partial f(\mathbf{w})$ داشته باشیم:

$$\|\mathbf{v}\| \leq L.$$

۵ فشرده‌سازی

در محبث فشرده‌سازی سیگنال‌ها، تعداد بیت‌های لازم برای فشرده‌سازی m متغیر گوسی با واریانس‌های $\sigma_i^۲, i = ۱, ..., m$ با میزان خطای D_i به طوری که مجموع خطا حداکثر D باشد، توسط مسئله‌ی بهینه‌سازی زیر توصیف می‌شود:

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=۱}^{m} \max\left\{\frac{۱}{۲}\ln(\frac{\sigma_i^۲}{D_i}), ۰\right\} \\ \text{subject to:} \quad & \sum_{i=۱}^{m} D_i = D \end{aligned}$$

۱. نشان دهید مساله‌ی فوق مساله‌ی بهینه‌سازی محدب است.

۲. با استفاده از روش KKT مقادیر بهینه‌ی $D_۱, ..., D_m$ برای این مسئله را به دست آورید. (جواب فرم بسته ندارد اما شکل water-filling خواهد داشت.)

۶ آنتروپی

آنتروپی نسبی بین دو بردار $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{++}$ به صورت زیر تعریف می‌شود:

$$D(\mathbf{x}, \mathbf{y}) = \sum_{k=۱}^{n} x_k \log(\frac{x_k}{y_k}).$$

آنتروپی نسبی، به صورت مشترک نسبت به $\mathbf{x}$ و $\mathbf{y}$ محدّب است. در مساله‌ی زیر در تلاشیم بردار $\mathbf{x}$ را طوری بیابیم که آنتروپی نسبی آن با بردار داده‌شده‌ی $\mathbf{y}$ کمینه شود به طوری که شرط‌های تساوی زیر روی $\mathbf{x}$ برقرار باشند:

$$\begin{aligned} \text{Minimize} \quad & D(\mathbf{x}, \mathbf{y}) = \sum_{k=۱}^{n} x_k \log(\frac{x_k}{y_k}) \\ \text{subject to:} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \sum_{k=۱}^{n} x_k = ۱ \end{aligned}$$

و فرض می‌کنیم که پارامترهای $\mathbf{y} \in \mathbb{R}^n_{++}$، $\mathbf{A} \in \mathbb{R}^{m\times n}$ و $\mathbf{b} \in \mathbb{R}^m$ داده‌شده هستند. دوگان لاگرانژین این مسئله را یافته و آن را ساده کنید تا به فرم زیر برسید:

$$\text{Maximize} \quad \mathbf{b}^\top \mathbf{z} - \log\left(\sum_{k=۱}^{n} y_k e^{\mathbf{a}_k^\top \mathbf{z}}\right)$$

که در آن منظور از $\mathbf{a}_k$، ستون k–ام ماتریس $\mathbf{A}$ است.

۷ محاسبه‌ی زیرگرادیان

۱. تابع $f(x) = \max\{x + ۸, ۹x, x^۲\}$ با دامنه‌ی $[۰, +\infty)$ را در نظر بگیرید. زیرگرادیان تابع f را در نقاط $x = ۱$ و $x = ۳$ را محاسبه کنید.

۲. زیرگرادیان تابع $f(\mathbf{x}) = \|x\| = \sqrt{\sum_{k=۱}^{n} x_k^۲}$ را در مبدأ مختصات بیابید.