به نام خدا

تمرین سری **دوم** درس **بهینه‌سازی**

(نیمسال دوم ۱۴۰۰)

۱- مساله‌ی بهینه‌سازی زیر را با متغیر $x \in \mathbb{R}^n$ درنظر بگیرید.

$$\begin{array}{ll} minimize & c^T x \\ subject\ to & Ax \preccurlyeq b \quad for\ all\ A \in \mathcal{A} \end{array}$$

در این مساله $\mathcal{A} \subseteq \mathbb{R}^{m\times n}$ به‌صورت زیر تعریف می‌شود:

$$\mathcal{A} = \{A \in \mathbb{R}^{m\times n} \mid \bar{A}_{ij} - V_{ij} \le A_{ij} \le \bar{A}_{ij} + V_{ij}, i = 1,2,\dots,m, j = 1,2,\dots,n\}$$

فرض می‌شود که ماتریس‌های $\bar{A}$ و V ثابت‌هایی هستند که در مساله داده شده‌اند. این مساله را می‌توانیم به‌عنوان یک مساله‌ی LP درنظر بگیریم که در آن هرکدام از ضرایب ماتریس A در یک بازه قرار گرفته‌اند و هدفی که دنبال می‌شود آن است که x به گونه‌ای باشد تا تمام قیدهای مربوط به این ضرایب برآورده گردد. این مساله را به فرم یک مساله‌ی LP بنویسید. (توجه: درنظر داشته باشید که نباید در پاسختان ابعادی وجود داشته باشد که با افزایش n یا m به صورت نمایی رشد کند.)

۲- مساله‌ی بهینه‌سازی زیر را درنظر بگیرید.

$$\begin{array}{ll} minimize & c^T x \\ subject\ to & Ax \preccurlyeq b \\ & x_i \in \{0,1\}, \qquad i = 1,\dots,n \end{array} \qquad (۱)$$

در این مساله x یک بردار n بعدی است که تنها می‌تواند مقادیر صفر یا یک داشته باشد. به‌طور کلی حل کردن مسائلی از این دست علیرغم آن‌که مجموعه‌ی جواب‌های ممکن متناهی است (حداکثر 2^n حالت) بسیاردشوار است. یک روش کلی برای حل این مسائله استفاده از این ساده‌سازی است که قید صفر و یک بودن x_i ها با قید نامساوی $0 \le x_i \le 1$ جایگزین شود:

$$\begin{array}{ll} minimize & c^T x \\ subject\ to & Ax \preccurlyeq b \\ & 0 \le x_i \le 1, \qquad i = 1,\dots,n \end{array} \qquad (۲)$$

استفاده از این ساده‌سازی باعث می‌شود که حل مساله به‌دست آمده بسیار ساده‌تر از مساله اولیه باشد.

a) نشان دهید که پاسخ بهینه‌ی مساله‌ی (۲) یک کران پایین از پاسخ بهینه‌ی مساله‌ی (۱) است. اگر مساله‌ی (۲) یک مساله infeasible باشد، در مورد مساله‌ی (۱) چه می‌توان گفت؟

b) اگر پاسخ مساله‌ی (۲) به‌گونه‌ای باشد که در آن $x_i \in \{0,1\}$ است؛ در مورد پاسخ مساله‌ی (۱) چه نظری می‌توان داد؟

۳- مساله پیدا کردن مقدار کمینه برای تابع $f(x,y) = (1-x)^2 + 100(y-x^2)^2$ را درنظر بگیرید. با فرض آنکه در نقطه شروع $x_0 = [5,5]^T$ قرار داشته باشیم و همچنین در اولین تکرار $\Delta_0 = 1$ باشد و مقدار $\widehat{\Delta} = 100$, $\eta = 0.15$ باشد. مراحل اجرای الگوریتم trust-region (dogleg) را برای یک گام بنویسید.

۴- نشان دهید که اگر $f(x)$ یک تابع درجه دو اکیدا محدب باشد آنگاه تابع $h(\sigma)$ که به‌صورت $h(\sigma) = f(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1})$ تعریف می‌شود نیز بر روی متغیر $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_{k-1})^T$ یک تابع درجه دو اکیدا محدب است. در اینجا بردارهای $p_0, p_1, \dots, p_{k-1}$ به‌صورت خطی از یکدیگر مستقل هستند. (راهنمایی: می‌توانید عبارت $\sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}$ را به‌صورت حاصلضرب یک ماتریس در بردار به فرم $P\sigma$ بنویسید که در آن P یک ماتریس است که ستون‌های آن بردارهای p_i می‌باشند.)

۵- با توجه به مفاهیم نرخ همگرایی به سوالات زیر پاسخ دهید:

الف) همگرایی سری $x_k = \frac{1}{k!}$ را از نظر Q-superlinearly و Q-quadratically بررسی کنید.

ب) سری $\{x_k\}$ به صورت زیر تعریف شده است:

$$x_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k}, & k \text{ even} \\ \frac{(x_{k-1})}{k}, & k \text{ odd} \end{cases}$$

همگرایی این سری را از نظر Q-superlinearly و Q-quadratically بررسی کنید.

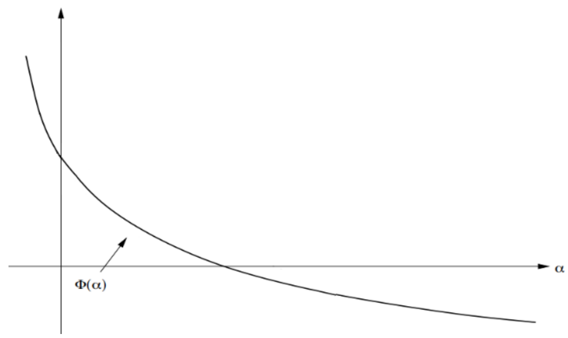
۶- روش Steepest Descent به تابعی به شکل زیر اعمال شده است:

$$Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \cdot A \cdot \mathbf{x} - \mathbf{b}^T \cdot \mathbf{x} + c$$

به طوری که A، b و c به ترتیب ماتریس، بردار و ثابت هستند. درصورت برقرار بودن چه شرایطی برای ماتریس A، روش Steepest Descent در یک گام به نقطه مینیمم مساله همگرا می‌شود؟ نقطه‌ی شروع می‌تواند هر نقطه‌ی دلخواهی مانند x_0 باشد (فرض کنید که نقطه شروع پاسخ مساله نباشد)؟ (راهنمایی: درصورتی روش

مذکور در یک گام به کمینه‌ی تابع Q همگرا می‌شود که خط جستجو (search line) $x_0 + \alpha d_0$ شامل کمینه‌ی تابع Q باشد)

۷- نشان دهید که اگر $0 < c_2 < c_1 < 1$ باشد ممکن است هیچ اندازه قدمی وجود نداشته باشد که بتواند شروط ولف را برآورده سازد. (راهنمایی: برای این منظور می‌توانید این امر را برای تابعی مانند شکل زیر نشان دهید.)



۸- پاسخی تحلیلی برای مساله‌ی LP زیر ارائه دهید.

$$\begin{aligned} &minimize\ c^T x \\ &subject\ to\!: l \preccurlyeq x \preccurlyeq u \end{aligned}$$

لازم به‌ذکر است در مساله‌ی فوق همواره $l \preccurlyeq u$ برقرار است.

تمرین‌های پیاده‌سازی

۹- در سوال دوم این تمرین مساله‌ی بهینه‌سازی (۱) را درنظر بگیرید که به‌صورت زیر تعریف شده بود.

$$\begin{array}{ll} minimize & c^T x \\ subject\ to & Ax \preccurlyeq b \\ & x_i \in \{0,1\}, \qquad i = 1, \dots, n \end{array} \qquad (۱)$$

فرض کنید مقدار بهینه برای این مساله p^* باشد. حال x^{rlx} را به عنوان پاسخ مساله‌ی (۲) که به صورت زیر تعریف شده بود درنظر بگیرید.

$$\begin{array}{ll} minimize & c^T x \\ subject\ to & Ax \preccurlyeq b \\ & \mathbf{0} \preccurlyeq x \preccurlyeq \mathbf{1} \end{array} \qquad (۲)$$

بنابراین $L = c^T x^{rlx}$ یک کران پایین برای p^* خواهد بود. از پاسخ x^{rlx} همچنین می‌توان به‌عنوان یک حدس اولیه $\hat{x}$ برای مساله‌ی (۱) استفاده کرد به این صورت که درایه‌های $\hat{x}$ را براساس حدآستانه $t \in [0,1]$ و به‌صورت زیر به صفر و یک تغییر دهیم که در آن $i = 1, \dots, n$ است:

$$\hat{x}_i = \begin{cases} 1 & x_i^{rlx} \geq t \\ 0 & otherwise \end{cases}$$

مشخصا $\hat{x}$ یک بردار بولین است که تمام درایه‌های آن مقادیر صفر یا یک دارند. اگر این پاسخ یک پاسخ feasible برای مساله‌ی (۱) باشد آنگاه می‌توانیم مقدار تابع هدف به ازای آن ($U = c^T \hat{x}$) را به‌عنوان کران بالای p^* درنظر بگیریم. هرچقدر U و L نزدیکتر باشند، $\hat{x}$ به پاسخ بهینه نزدیکتر است. لازم به‌ذکر است که این روش حل، ممکن است همیشه موثر نباشد چرا که ممکن است حالتی رخ دهد که در آن به ازای تمامی مقادیر ممکن برای t، پاسخ $\hat{x}$ یک پاسخ Infeasible شود. با این حال در برخی موارد می‌توان پاسخ خوبی به دست آورد.

مشابه کد زیر داده‌های مساله فوق را تولید کنید.

```
n=100;
m=300;
A=rand(m,n);
b=A*ones(n,1)/2;
c=-rand(n,1);
```

با استفاده از داده‌های تولید شده، پاسخ بهینه را برای مساله‌ی (۲) به‌دست آورید (برای این کار می‌توانید از cvx استفاده کنید). مقدار کران پایین L را محاسبه کنید. به ازای ۱۰۰ مقدار مختلف برای حدآستانه t که با فاصله‌های یکسان در بازه $[0,1]$ قرار گرفته‌اند پاسخ $\hat{x}$ را محاسبه کرده و مقدار تابع هدف ($c^T x$) را به ازای آن به‌دست آورید. برای هرکدام از پاسخ‌های به‌دست آمده‌ی $\hat{x}$ مقدار $\max_i (A\hat{x} - b)_i$ را نیز محاسبه کنید. سپس این مقادیر را در نموداری نسبت به t رسم کنید. در نمودار رسم شده همچنین مشخص کنید که به ازای چه مقادیری از t پاسخ $\hat{x}$ یک پاسخ feasible بوده است و به ازای چه مقادیری این پاسخ Infeasible بوده است. درنهایت مقداری را برای t پیدا کنید که به ازای آن $\hat{x}$ یک پاسخ feasible باشد و کمترین مقدار را برای تابع هدف نتیجه دهد. این پاسخ را به عنوان کران بالای U درنظر بگیرید. فاصله میان کران بالا و کران پایین ($U - L$) متناظر با پاسخ بهینه‌ی p^* را نیز گزارش کنید.

۱۰- مساله‌ی $\min_x \frac{1}{2} x^T A_i x - b^T x$ را که در آن $i \in \{1,2\}$ است را در نظر بگیرید که در آن $b \in \mathbb{R}^n$ برداری است که تمام عناصر آن مقدار ۱ دارد و A_i به‌صورت زیر تعریف می‌شود.

$$A_1 = tridiag(-1,4,-1)_{n\times n} = \begin{bmatrix} 4 & -1 & 0 & 0 & \dots & 0 \\ -1 & 4 & -1 & 0 & \dots & 0 \\ 0 & -1 & 4 & -1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 4 & -1 \\ 0 & 0 & 0 & \dots & -1 & 4 \end{bmatrix}_{n\times n}$$

$$A_2 = hilb(n) = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \ddots & \ddots & \frac{1}{n+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \dots & \frac{1}{2n-1} \end{bmatrix}_{n\times n}$$

a) الگوریتم‌های steepest descent، Newton و BFGS را با بهره‌گیری از متد backtracking line search (الگوریتم ۳.۱ کتاب Nocedal) پیاده‌سازی نمایید و از آن‌ها برای حل مساله‌ی بهینه‌سازی داده شده برای $i = 1,2$ و همچنین $n = 40,400$ پیاده‌سازی کنید و برای هرکدام زمان حل مساله و مقدار $\|A_i x^* - b\|_2$ را محاسبه کنید و در جدولی نمایش دهید (x^* پاسخ بهینه‌ای است که برای مساله به‌دست آمده است).

b) مساله‌ی فوق را برای ماتریس‌های $\hat{A}_i = A_i + \epsilon, i = 1,2$ به ازای $n = 40,400$حل کنید و برای هرکدام $\|\hat{A}_i \hat{x}^* - b\|_2$، $\|A_i - \hat{A}_i\|_2$ و $\|x^* - \hat{x}^*\|_2$ را در جدولی نمایش دهید. در اینجا $\hat{x}^*$ پاسخ بهینه‌ای است که پس از حل مساله به‌دست می‌آید و x^* پاسخ بهینه‌ای است که از حل مساله‌ی نظیر در بخش a محاسبه شده است. (فرض کنید $\epsilon = 1 \times 10^{-6}$ باشد)

c) تحلیل خود را از نتایج به‌دست آمده از قسمت‌های a و b ارائه دهید.

فرمت گزارش:

گزارش بایستی حاوی تمام نتایج بدست آمده از شبیه‌سازی‌های کامپیوتری در قالب فایل PDF باشد.

درصورتی که تمرینات را به‌صورت دستنویس حل می‌کنید. فایل‌های عکس تمرینات را با کیفیت مناسب و به ترتیب سوالات در یک فایل pdf قرار دهید و درنهایت این فایل را آپلود نمایید.

فایل گزارش خود را به شکل « StdNum_HWNum.pdf» نام‌گذاری کنید. (مانند 9272203_HW2.pdf)

فرمت کدها:

برای هر تمرین شبیه‌سازی کامپیوتری بایستی فایل کد جداگانه در محیط MATLAB، Python یا R تهیه شود.

هر فایل کد خود را به شکل « Q_k » نام‌گذاری کنید. که k بیانگر شماره سوال شبیه‌سازی خواهد بود.

نحوه تحویل:

فایلهای کد و گزارش خود را که طبق فرمت‌های فوق تهیه شده‌اند، در قالب یک فایل فشرده بارگذاری نمایید. فایل فشرده را به شکل «StdNum_HW2.zip» نام‌گذاری کنید. (مانند 9272203_HW2.zip)

تذکر:

- در صورتیکه پارامتر خاصی در سوالات مشخص نشده با توجه به اطلاعاتی که در ارتباط با محدوده پارامتر دارید، مقدار دلخواهی انتخاب کنید و آن را در گزارش توضیح دهید.
- ارسال تمرین‌های تئوری تا یک روز تاخیر بلامانع است. پس از آن پاسخ این بخش بارگذاری خواهد شد و لذا مواردی که پس از بارگذاری پاسخ‌ها ارسال شوند قابل قبول نخواهد بود.
- برای تمرین‌های بخش پیاده‌سازی در مجموع ۷ روز تاخیر مجاز است (برای کل تمرینات جمع تاخیرهای شما نباید از ۷ روز بیشتر شود). در صورت تاخیر بیشتر از ۷ روز کسر ۵ درصد نمره از نمره کل تمرینات پیاده‌سازی به ازای هر روز تاخیر مد نظر قرار خواهد گرفت.
- در صورت شبیه بودن تمارین دانشجویان، نمره تمرین بین دانشجویان با تمرین مشابه تقسیم خواهد شد.
- در صورت وجود هرگونه سوال یا ابهام با یکی از ایمیل‌های shervin.halat@gmail.com و یا b.roshanfekr@aut.ac.ir ارتباط برقرار کنید.