Optimization Methods **Fall 2020**

Homework 3

Instructor: Lijun Zhang *Name:* 黄彦骁, *StudentId:* 191300020

Notice

- The submission email is: **opt4grad@163.com**.
- Please use the provided LaTeX file as a template. If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: Negative-entropy Regularization

Please show how to compute

$$\underset{x\in\Delta^n}{\operatorname{argmin}}\, b^\top x + c\cdot\sum_{i=1}^n x_i \ln x_i$$

where $\Delta^n = \{x|\sum_{i=1}^n x_i = 1, x_i \geq 0, i = 1,\cdots,n\}$, $b\in\mathbb{R}^n$ and $c\in\mathbb{R}$.

Solution :

问题可以被转化为求

$$min\ b^T x + c\sum_{i=1}^n x_i log x_i$$
$$1^T x = 1$$

由于约束为线性约束且能找到可行解，问题为凸问题，所以满足 $Slater$ 条件，问题有强对偶性。

函数 $f(x) = b^T x + c\sum_{i=1}^n x_i log x_i$ 的对偶函数为 $g(v) = -v - f^*(-v)$, 其中 f^* 函数为 f 的共轭函数，为 $f^*(y) = \sum_{i=1}^n e^{\frac{y_i - b_i - c}{c}}$。所以其对偶问题为:

$$max\ \ -v - \sum_{i=1}^n e^{\frac{-v-b_i-c}{c}}$$

该问题的最优解为 $v^* = c(1+log\frac{\sum_{i=1}^n e^{-\frac{b_i}{c}}}{c})$。而这个解也就对应于原问题的最优解。$(v^*)$ 处的 $lagrange$ 函数为:

$$L(x,v^*) = b^T x + c\sum_{i=1}^n x_i log x_i + v^*(1^T x - 1)$$

所以最优值对应的点有 $x_i^* = \frac{1}{e^{\frac{b_i+c+v^*}{c}}}$

Problem 2: One inequality constraint

(1) With $c \neq 0$, express the dual problem of

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & f(x) \leq 0, \end{aligned}$$

in terms of the conjugate $f^\star$.

(2) Explain why the problem you give is convex. We do not assume f is convex.

Solution :

（1） $\lambda = 0$ 时， $g(\lambda) = inf\ c^T x = -\infty$

对 $\lambda > 0$, $g(\lambda) = inf(c^T x + \lambda f(x)) = \lambda inf(\frac{c}{\lambda}^T x + f(x)) = -\lambda f_1^*(-\frac{c}{\lambda})$, 所以对偶问题为：

$$maximize\ -\lambda f_1^*(-\frac{c}{\lambda})$$

$$\lambda \geq 0$$

（2）因为我们极大化的目标函数 $g(\lambda)$ 为凹函数。且可行集为凸集，所以给出的对偶问题为凸优化问题。

Problem 3: KKT conditions

Consider the problem

$$\begin{aligned} &\min_{x \in \mathbb{R}^2} x_1^2 + x_2^2 \\ \text{s.t.}\quad &(x_1 - 1)^2 + (x_2 - 1)^2 \leq 2 \\ &(x_1 - 1)^2 + (x_2 + 1)^2 \leq 2 \end{aligned}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.

a) Write the Lagrangian for this problem.

b) Does strong duality hold in this problem?

c) Write the KKT conditions for this optimization problem.

Solution :

（1） $lagrange$ 对偶函数为：

$$g(\lambda_1, \lambda_2) = inf_{x_1, x_2} L(x_1, x_2, \lambda_1, \lambda_2)$$

其中:

$$\begin{aligned} &L(x_1, x_2, \lambda_1, \lambda_2) \\ &\quad = x_1^2 + x_2^2 + \lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 2) + \lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 2) \\ &\quad = (1 + \lambda_1 + \lambda_2)x_1^2 + (1 + \lambda_1 + \lambda_2)x_2^2 - 2(\lambda_1 + \lambda_2)x_1 - 2(\lambda_1 - \lambda_2)x_2 + 2(\lambda_1 + \lambda_2) \end{aligned}$$

函数 L 取到最小值时有：

$$x_1 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2}, x_2 = \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2}$$

则有：

$$g(\lambda_1, \lambda_2) = \begin{cases} -\frac{(\lambda_1+\lambda_2)^2 + (\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + 2(\lambda_1 + \lambda_2) & 1 + \lambda_1 + \lambda_2 \geq 0 \\ -\infty & otherwise \end{cases}$$

所以对偶问题为:

$$\begin{aligned} &maximize\ \frac{2(\lambda_1 + \lambda_2) + (\lambda_1 + \lambda_2)^2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \\ &subject\ to\ \lambda_1, \lambda_2 \geq 0 \end{aligned}$$

（2）显然有 x_1, x_2 存在使得 $(x_1-1)^2+(x_2-1)^2 < 2, (x_1-1)^2+(x_2+1)^2 < 2$ 成立, 例如 $x_1 = 0, x_2 = 1$, 且原问题为凸问题，所以强对偶性成立。

（3）KKT 条件为：

$$(x_1-1)^2+(x_2-1)^2 \le 2, (x_1-1)^2+(x_2+1)^2 \le 2$$

$$\lambda_1 \ge 0, \lambda_2 \ge 0$$

$$2x_1+2\lambda_1(x_1-1)+2\lambda_2(x_1-1)=2x_2+2\lambda_1(x_2-1)+2\lambda_2(x_2+1)=0$$

$$\lambda_1((x_1-1)^2+(x_2-1)^2-2)=\lambda_2((x_1-1)^2+(x_2+1)^2-2)=0$$

Problem 4: Matrix eigenvalues

We denote by $f(A)$ the sum of the largest r eigenvalues of a symmetric matrix $A \in \mathbb{S}^n$ (with $1 \le r \le n$), i.e.,

$$f(A)=\sum_{k=1}^{r}\lambda_k(A),$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A sorted in decreasing order. Show that the optimal value of the optimization problem

$$\begin{aligned} \max \quad & \boldsymbol{tr}(AX) \\ \text{s.t.} \quad & \boldsymbol{tr}X = r \\ & 0 \preceq X \preceq I, \end{aligned}$$

with variable $X \in \mathbb{S}^n$, is equal to $f(A)$.

Solution :

问题可以转化为：

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} a_i x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i = r \\ & 0 \le x_i \le i, \quad i=1,2,3\ldots,n \end{aligned}$$

则显然最大特征值值 a_1 的最大权重为 1，接下来的次大值显然一样如此，故最优值 $p^*=\sum_{i=1}^{r} a_i = f(A)$.

Problem 5: Determinant optimization

Derive the dual problem of the following problem

$$\begin{aligned} \min \quad & \log\det X^{-1} \\ \text{s.t.} \quad & A_i^T X A_i \preceq B_i,\ i=1,\ldots,m \end{aligned}$$

where $X \in \mathbb{S}_{++}^n$, $A_i \in \mathbb{R}^{n\times k_i}$, $B_i \in \mathbb{S}_{++}^{k_i}$, $k_i \in \mathbb{N}_+$, $i=1,\ldots,m$.

Solution :

引入乘子 $Z_1 \dots, Z_m \in S^k$，则 $Lagrange$ 函数为:

$$
\begin{aligned}
L(X, Z_1, \dots, Z_m) = & \\
&= log\ det\ X^{-1} + \sum_{i=1}^{m} tr((A_i^T X A_i - B_i) Z_i) \\
&= -\sum_{i=1}^{m} tr(B_i Z_i) + log\ det\ X^{-1} + tr((\sum_{i=1}^{m}(A_i A_i^T Z_i))X)
\end{aligned}
$$

所以有:

$$
\begin{aligned}
g(Z_1, \dots, Z_m) &= inf_{x \in S_{++}^n}(L(X, Z_1, \dots, Z_m)) \\
&= -\sum_{i=1}^{m} tr(B_i Z_i) + f_0^*(-\sum_{i=1}^{m} A_i A_i^T Z_i) \\
&= \begin{cases} log\ det\ (\sum_{i=1}^m A_i A_i^T Z_i) - \sum_{i=1}^m tr(B_i Z_i) + n & \sum_{i=1}^m A_i A_i^T Z_i \succ 0 \\ -\infty & otherwise \end{cases}
\end{aligned}
$$

所以对偶问题为:

$$
\begin{aligned}
& maxmize\ log\ det\ (\sum_{i=1}^{m} A_i A_i^T Z_i) - \sum_{i=1}^{m} tr(B_i Z_i) + n \\
& subject\ to\ \sum_{i=1}^{m} A_i A_i^T Z_i \succ 0 \\
& \qquad\qquad Z \succeq 0
\end{aligned}
$$