Optimization Methods **Fall 2021**

Homework 3

Instructor: Lijun Zhang *Name:* 麻超, *StudentId:* 201300066

Notice

- The submission email is: **zhangzhenyao@lamda.nju.edu.cn**.
- 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: One inequality constraint

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 $f^\star$.

Solution:

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

当 $\lambda > 0$ 时, 有 $g(\lambda) = \inf(c^T x + \lambda f(x)) = \lambda \inf((\frac{c}{\lambda})^T x + \lambda f(x)) = -\lambda f_1^*(-\frac{c}{\lambda})$

其 dual problem 为:

$$\begin{aligned} &\text{minimize } -\lambda f_1^*(-\frac{c}{\lambda}) \\ &\text{subject to } \lambda \geq 0 \end{aligned}$$

□

Problem 2: KKT conditions

Consider the problem

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

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

(1) Write the Lagrangian for this problem.

(2) Does strong duality hold in this problem?

(3) Write the KKT conditions for this optimization problem.

Solution:

a) 可以得到其 Lagrangian 为 $L(x,\lambda)=x_1^2+x_2^2+\lambda_1((x_1-1)^2+(x_2-1)^2-2)+\lambda_2(x_1-1)^2+(x_2+1)^2-2)$, 其中 $\lambda\in R^2$

b) 首先, 由于 $\nabla^2 x_1^2+x_2^2=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$, 该矩阵为正定矩阵, 则表明目标函数为凸函数.

其次,$\nabla^2(x_1-1)^2+(x_2-1)^2-2)=\nabla^2(x_1-1)^2+(x_2+1)^2-2)=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$, 为正定矩阵, 表明约束函数为凸函数, 则原问题为凸优化问题.

由于存在 $x=\begin{bmatrix}1 & 0\end{bmatrix}^T\in relint\ D=R^2$, 使得约束不等式 $(x_1-1)^2+(x_2-1)^2-2)=(x_1-1)^2+(x_2+1)^2-2)=-1<0$,

所以满足 Slater 条件, 表明该不等式约束严格成立, 即强对偶性成立.

c) 其 KKT conditions 是

$$
\begin{aligned}
&(x_1-1)^2+(x_2-1)^2-2)\le 0\\
&(x_1-1)^2+(x_2+1)^2-2)\le 0\\
&\lambda_1\ge 0\\
&\lambda_2\ge 0\\
&\lambda_1((x_1-1)^2+(x_2-1)^2-2))=0\\
&\lambda_2((x_1-1)^2+(x_2+1)^2-2))=0\\
&2x_1+2\lambda_1(x_1-1)+2\lambda_2(x_1-1)=0\\
&2x_2+2\lambda_1(x_2-1)+2\lambda_2(x_2+1)=0
\end{aligned}
$$

$\square$

Problem 3: Equality Constrained Least-squares

Consider the equality constrained least-squares problem

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}\|Ax-b\|_2^2\\
\text{subject to} & Gx=h
\end{array}
$$

where $A\in\mathbf{R}^{m\times n}$ with $\mathbf{rank}\ A=n$, and $G\in\mathbf{R}^{p\times n}$ with $\mathbf{rank}\ G=p$.

(1) Derive the Lagrange dual problem with Lagrange multiplier vector v.

(2) Derive expressions for the primal solution $x^\star$ and the dual solution $v^\star$.

Solution:

a)$L(x,v)=\frac{1}{2}\|Ax-b\|_2^2+V^T(Gx-h)=\frac{1}{2}x^TA^TAx+(G^Tv-A^Tb)^Tx-v^Th$

所以当 $x=-(A^TA)^{-1}(G^Tv-A^Tb)$ 取得最小值.

也即用拉格朗日乘数向量 v 导出的拉格朗日对偶问题为

$$g(v)=-2(G^Tv-A^Tb)^T(A^TA)^{-1}(G^Tv-A^Tb)-v^Th$$

b) 其最优条件为 $\nabla xL(x,v)=A^T(Ax^*-b)+G^Tv^*=0, Gx^*=h, x^*=-(A^TA)^{-1}(G^Tv-A^Tb)$

代入 $Gx^*=h$ 可得,$G(A^TA)^{-1}(A^Tb-G^Tv^*)=h$

$\therefore v^*=G((A^TA)^{-1}(G^T)^{-1}(h-G(A^TA)^{-1}(A^Tb))$

□

Problem 4: Negative-entropy Regularization

Please show how to compute

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

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:

原问题可以转化为求如下优化问题的最优解:

$$\text{minimize } b^Tx+c\cdot\sum_{i=1}^n x_i\ln x_i, x_i\geq 0$$
$$\text{subject to } I^Tx-1=0$$

由于凸函数进行保凸运算之后仍为凸函数, 则不等式约束函数为凸函数, 等式约束函数仿射. 所以这是一个凸优化问题, 且当 $x=[1/n,1/n,...,1/n]^T$ 时, 明显有 Slater 条件成立, 即 KKT 条件是保证其最优性的充要条件.

于是引入一个 $\lambda\in R^n, v\in R$, 得到:

(1)$-x\leq 0$

(2)$I^Tx-1=0$

(3)$\lambda\geq 0$

(4)$\lambda_i(-x_i)=0, i=1,2,...,n$

(5)$b_i+c(\ln x_i+1)-\lambda+v=0, i=1,2,...,n$

可以得到 λ 是一个松弛变量, 可以消去, 于是得到如下条件:

(1)$x\geq 0$

(2)$I^Tx-1=0$

(3)$b_ix_i+cx_i(\ln x_i+1)+vx_i=0, i=1,2,...,n$

(4)$b_i+c(\ln x_i+1)+v\geq 0, i=1,2,...,n$

综上可得目标函数 $=\sum_{i=1}^n -cx_i-vx_i$

所以可以得到

$$c+v=\begin{cases}>0 & ,x_i=e^{(-b_i-v)/c-1}\\ \leq 0 & ,/\end{cases}$$

$\therefore x_i = e^{(b_i - v)/c - 1}$, 根据 (2) 可以得到 $\sum_{i=1}^n x_i = e^{(b_i - v)/c-1} = 1$, 所以 v 有唯一解, 满足条件, 得到 x_i 的值.

$\square$

Problem 5: Support Vector Machines

Consider the following optimization problem

$$\text{minimize} \quad \textstyle\sum_{i=1}^n \max\left(0, 1 - y_i(w^T x_i + b)\right) + \frac{\lambda}{2}\|w\|_2^2$$

where $x_i \in \mathbf{R}^d, y_i \in \mathbf{R}, i = 1, \cdots, n$ are given, and $w \in \mathbf{R}^d, b \in \mathbf{R}$ are the variables.

(1) Derive an equivalent problem by introducing new variables $u_i, i = 1, \cdots, n$ and equality constraints

$$u_i = y_i(w^T x_i + b), i = 1, \cdots, n.$$

(2) Derive the Lagrange dual problem of the above equivalent problem.

(3) Give the Karush-Kuhn-Tucker conditions.

Hint: Let $\ell(x) = \max(0, 1-x)$*. Its conjugate function* $\ell^*(y) = \sup\limits_x (yx - \ell(x)) = \begin{cases} y, & -1 \le y \le 0 \\ \infty, & \textit{otherwise} \end{cases}$