Optimization Theory and Algorithm II	**Homework 1 - 08/10/2021**

Homework 1

Lecturer: Xiangyu Chang *Scribe:* 吴博成

Edited by: Junbo Hao

1 HW 1

1.1 问题重述

思考以下凸优化问题:

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t. } \mathbf{x} \leq 0$$

然后，根据一般的最优性条件证明最优点 x^* 满足

$$\nabla f\left(\mathrm{x}^*\right) \leq 0$$

$$x_i^*\left(\nabla f\left(\mathrm{x}^*\right)\right)_i=0, i=1, \ldots, n$$

1.2 问题求解

原题有约束凸优化问题可以转化为:

$$\min_{\mathbf{x}} f(\mathbf{x})+\lambda x$$
$$\text{其中}\lambda \geq 0$$

对该等价问题使用一班的最优性条件可知，当目标函数梯度为 0 时有最优解，即:

$$\nabla f\left(\mathrm{x}^*\right)+\lambda=0$$

所以:

$$\nabla f\left(\mathrm{x}^*\right) \leq 0$$

接下来证明:

$$x_i^* \nabla f(x^*)_i=0$$

由条件和上述已经证明的结论，当且仅当 $x_i^*<0$ 以及 $\nabla f(x^*)_i<0$ 时,$x_i^* f(x^*)_i \neq 0$
因为 $\lambda x \geq 0$ 且 $f(x)=-\lambda$,
所以: $x^* \nabla f(x^*) \leq 0$, 当且仅当 $x_i^*<0$ 以及 $\nabla f(x^*)_i<0$ 时,$x_i^* f(x^*)_i \neq 0$

假设 $x_i^*<0$ 以及 $\nabla f(x^*)_i<0$, $\exists \epsilon \rightarrow 0^+$, $d=\begin{pmatrix} 0 \\ f(x^*)_i \\ \ldots\ldots \\ 0 \end{pmatrix}$, 使得 $y=x^*+\epsilon d\ (y<0)$

代入泰勒公式可得：

$$f(y) = f(x^*) + \epsilon d \nabla f(x^*) + O(\epsilon^2 d \nabla f(x^*)) < f(x^*)$$

与 $f(x^*)$ 是最小值矛盾，故原假设不成立，即 $x_i^* \nabla f(x^*)_i = 0$

2 HW 2

2.1 问题重述

写出下列问题的 Lagrange 对偶问题:

$$
\begin{aligned}
&\min \mathbf{c}^\top \mathbf{x} \\
&\text{s.t. } A\mathbf{x} \leq \mathbf{b}
\end{aligned}
$$

2.2 问题求解

拉格朗日函数:

$$L(x,\lambda) = c^T x + \lambda^T(Ax-b) = (c^T + \lambda^T A)x - \lambda^T b$$

拉格朗日对偶函数:

$$
\begin{aligned}
g(\lambda) &= \inf_{x\in D}\{L(x,\lambda)\} \\
&= \inf_{x\in D}\left\{(c^T+\lambda^T A)x - \lambda^T b\right\} \\
&= \begin{cases} -\lambda^T b & , c^T + \lambda^T A = 0 \\ -\infty & , otherwise \end{cases}
\end{aligned}
$$

拉格朗日对偶问题:

$$
\begin{aligned}
\max \quad & -\lambda^T b \\
s.t. \quad & c^T + \lambda^T A = 0 \\
& \lambda \geq 0
\end{aligned}
$$

3 HW 3

3.1 问题重述

计算共轭函数:

(i) $f(\mathrm{x}) = \delta_{B_\infty}(\mathrm{x})$

(ii) $f(\mathrm{x}) = \delta_{\mathbb{R}^n_-}(\mathrm{x})$

(iii) $f(x) = \log(1 + \exp(x))$

(iv) $f(\mathrm{x}) = g(\mathrm{x} - \mathbf{a}) + < \mathrm{x}, \mathbf{b} >$

(vi) $f(\mathrm{x}) = \inf_{\mathbf{z}} \left\{ \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|^2 + g(\mathbf{z}) \right\}$

3.2 问题求解

(i)

$$\begin{aligned} f^*(y) &= \sup_x \left\{ y^T x - f(x) \right\} \\ &= \sup_{x \in B_{\|\cdot\|_\infty \le 1}} \left\{ y^T x - f(x) \right\} \\ &= \sigma_{B_0}(x) \end{aligned}$$

(ii)

$$\begin{aligned} f^*(y) &= \sup_x \left\{ y^T x - f(x) \right\} \\ &= \sup_{x \in R^n_-} \left\{ y^T x \right\} \\ &= \begin{cases} 0 & , y \in R^n_+ \\ \infty & , other \end{cases} \\ &= \delta_{R^n_+}(y) \end{aligned}$$

(iii)

$$f^*(y) = \sup_x \left\{ y^T x - log(1 + e^x) \right\}$$
$$l(x, y) = y^T x - log(1 + e^x)$$

其对 x 求导可得: $y - \frac{e^x}{1+e^x} = 0$，解得 $x = -\log(1 - y)$, $y \in (0, 1)$

当 $y = 0$ 时，$\sup_x \left\{ y^T x - log(1 + e^x) \right\} = \sup_x \left\{ -log(1 + e^x) \right\} = \infty$

当 $y = 1$ 时，$\sup_x \left\{ y^T x - log(1 + e^x) \right\} = \sup_x \left\{ x - log(1 + e^x) \right\} = 0$

综上所述:

$$f^*(y) = \begin{cases} y \log(\frac{y}{1-y}) + \log(1 - y) & , y \in (0, 1) \\ 0 & , y = 1 \\ \infty & , othe \end{cases}$$

(iv) 令 $z = x - a$

$$
\begin{aligned}
f^*(y) &= \sup_x \left\{ y^T(z+a) - f(x) \right\} \\
&= \sup_Z \left\{ y^T(z+a) - g(z) - b^T(z+a) \right\} \\
&= \sup_Z \left\{ (y^T - b^T)(z+a) - g(z) \right\} \\
&= \sup_Z \left\{ (y^T - b^T)z - g(z) \right\} + (y^T - b^T)a \\
&= g^*(y-b)^T + (y-b)^T a
\end{aligned}
$$

(vi)

$$
\begin{aligned}
f^*(y) &= \sup_x \left\{ y^T x - \inf_z \left\{ \frac{1}{2}\|x-z\|^2 + g(z) \right\} \right\} \\
&= \sup_x \left\{ \sup_z \left\{ -\frac{1}{2}\|x-z\|^2 - g(z) \right\} + y^T x \right\} \\
&= \sup_z \sup_x \left\{ -\frac{1}{2}\|x-z\|^2 - g(z) + y^T x \right\} \\
&= \sup_z \sup_x \left\{ -\frac{1}{2}x^T x + x^T z - \frac{1}{2}z^T z - g(z) + y^T x \right\} \\
&= \sup_z \left\{ \sup_x \left\{ -\frac{1}{2}x^T x + x^T z + y^T x \right\} - \frac{1}{2}z^T z - g(z) \right\} \\
&= \sup_z \left\{ \frac{(y+z)^T(y+z)}{2} - \frac{1}{2}z^T z - g(z) \right\} \\
&= \sup_z \left\{ \frac{y^T y + y^T z + z^T y + z^T z}{2} - \frac{1}{2}z^T z - g(z) \right\} \\
&= \sup_z \left\{ y^T z - g(z) \right\} + \frac{y^T y}{2} \\
&= g^*(y) + \frac{y^T y}{2}
\end{aligned}
$$

4 HW 4

4.1 问题重述

(1) 定义负熵函数 $f(x) = x\log x$ and $x \geq 0, 0\log 0 = 0$, 写出其共轭函数.

(ii) 写出熵最大化问题的 Lagrange 对偶问题:

$$\begin{aligned}
&\min_{\mathbf{x}} \sum_i x_i \log(x_i), \\
&\text{s.t. } A\mathbf{x} \succeq \mathbf{b}, \\
&\sum_i x_i = 1
\end{aligned}$$

(iii) 假设熵最大化问题具有强对偶性，我们得到了最优对偶变量 λ^* 和 ν^*, 利用 λ^* 和 ν^* 计算最优原始变量 x*.

4.2 问题求解

(i)

$$\begin{aligned}
f^*(y) &= \sup_x \left\{ y^T x - f(x) \right\} \\
&= \sup_x \left\{ y^T x - x\log(x) \right\}
\end{aligned}$$

$$\begin{aligned}
&\text{令}\nabla(y^T x - x\log(x)) = y - 1 - log(x) = 0 \\
&\text{解得}x = e^{y-1} \\
&\Rightarrow f^*(y) = y^T e^{y-1} - e^{y-1}(y-1) = e^{y-1}
\end{aligned}$$

(ii) 拉格朗日函数:

$$L(x, \lambda, v) = \sum_i x_i \log(x_i) - \lambda^T(Ax - b) + \sum_i v_i x_i - V^T$$

拉格朗日对偶函数:

$$\begin{aligned}
g(\lambda, v) = \inf_x L(x, \lambda, v) &= \inf_x \left\{ \sum_i x_i \log(x_i) - \sum_i \lambda_i a_i x_i + \sum_i v_i x_i \right\} + \lambda^T b - V^T \\
&= \inf_x \left\{ \sum_i (v_i - \lambda_i a_i) x_i + \sum_i x_i \log(x_i) \right\} + \lambda^T b - V^T \\
&= \sum_i - \sup_x \left\{ (\lambda_i a_i - v_i) x_i - \sum_i x_i \log(x_i) \right\} + \lambda^T b - V^T \\
&= -\sum_i f^*(\lambda_i a_i - v_i) + \lambda^T b - V^T
\end{aligned}$$

由（1）得:

$$\sum_i f^*(y) = \sum_i e^(y_i - 1) \Rightarrow -\sum_i f^*(\lambda_i a_i - v_i) = -\sum_i e^{\lambda_i a_i - v_i}$$

故有:

$$g(\lambda, v) = -\sum_i e^{\lambda_i a_i - v_i} + \lambda^T b - V^T$$

拉格朗日对偶问题:

$$\begin{aligned} \max_{\lambda, v} \quad & -\sum_i e^{\lambda_i a_i - v_i - 1} + \lambda^T b - V^T \\ s.t. \quad & \lambda \geq 0 \end{aligned}$$

其中，a_i 是 A 的第 i 列向量

(iii)

$$\begin{aligned} \frac{\partial L(x, \lambda^*, v^*)}{\partial x_i} &= 1 + \log(x_i) - A^T \lambda^* + v^T = 0 \\ \Rightarrow x_i &= e^{v^{*T} + A^T \lambda^* - 1} \end{aligned}$$

即 x^* 的每一个分量 x_i 满足 $x_i = e^{v^{*T} + A^T \lambda^* - 1}$

5 HW 5

5.1 问题重述

写出拉格朗日对偶问题:

(i)

$$\min_x f(x) + g(Ax)$$

(ii) 岭回归:

$$\min_x \frac{1}{2}\|Ax - b\|^2 + \lambda\|x\|_2^2$$

5.2 问题求解

(1)

令 $z = Ax$

则:

$$\begin{aligned} g(v) &= \inf_x \left\{ f(x) + g(z) + v^T(Ax - z) \right\} \\ &= -\sup_x \left\{ -f(x) - g(z) - v^T Ax + v^T z) \right\} \\ &= -\sup_x \left\{ -f(x) - v^T Ax \right\} + \sup_z \left\{ -g(z) + v^T z) \right\} \\ &= -f^*(-A^T v) - g^*(v) \end{aligned}$$

拉格朗日对偶问题:

$$\max_v -f^*(-A^T v) - g^*(v)$$

(2) 将原问题转化为令 $z = Ax - b$

则:

$$\begin{aligned} g(v) &= \inf_x \left\{ \frac{1}{2}\|z\|_2^2 + \lambda\|x\|_2^2 + v^T(Ax - b - z) \right\} \\ &= \inf_x (\left\{ \frac{1}{2}\|z\|_2^2 - v^T z \right\} + \lambda(\|x\|_2^2 + \frac{v^T Ax}{\lambda}) - v^T b) \\ &= \inf_x \left\{ \frac{1}{2}\|z\|_2^2 - v^T z \right\} + \inf_x \left\{ \lambda(\|x\|_2^2 + \frac{v^T Ax}{\lambda}) \right\} - v^T b \\ &= -\frac{1}{2}\|v\|_2^2 + \frac{\lambda}{4}\|\frac{v^T A}{\lambda}\|_2^2 - v^T b \end{aligned}$$

拉格朗日对偶问题:

$$\max_v -\frac{1}{2}\|v\|_2^2 + \frac{\lambda}{4}\|\frac{v^T A}{\lambda}\|_2^2 - v^T b$$