COMP 8650 Longfei Zhao u5976992.

Conjugate function

a) $f^*(y) = \sup_x\{x^Ty - f(x)\} \geq x^Ty - f(x)$ for all $x \in dom(f)$

$\therefore f(x) + f^*(y) \geq f(x) + x^Ty - f(x) = x^Ty$

b) $f^*(x) = \sup_x\{x^Ty - f(x)\}$

$\Rightarrow f^*(0) = \sup_x\{-f(x)\}$

$\Rightarrow -f^*(0) = -\sup_x\{-f(x)\}$.

$\because \sup f(x) = -\inf_x(-f(x))$

$\therefore -f^*(0) = \inf_x f(x)$

c). Firstly, $dom(f) = \{x \mid x \succ 0\}$

$f^*(y) = \sup_{dom(f)}\{x^Ty - \sum_{i=1}^n a_i \log x_i\}$.

$= \sup_{dom(f)}\{\sum_{i=1}^n (x_iy_i - a_i \log x_i)\}$.

Now, consider $g = \sum_{i=1}^n (x_iy_i - a_i \log x_i)$

and $g_i = x_iy_i - a_i \log x_i$

$\therefore g = \sum_{i=1}^n g_i$ and $\frac{\partial g}{\partial x_i} = y_i - \frac{a_i}{x_i}$

① if $y_i > 0$,

when $x_i \to \infty$, $g_i \to \infty$ and $g \to \infty$.

② if $y_i = 0$,

if $a_i = 0$, $g_i = 0$

otherwise, $g_i = a_i \log x_i$, it will be ∞ when $x_i \to 0$ or $x_i \to \infty$. (depend on if $a > 0$).

③ if $y_i < 0$.

if $a_i > 0$, when $x_i \to 0$, $g_i \to \infty$.

if $a_i = 0$, $g_i = y_i x$, when $x_i \to \infty$, $g_i \to 0$.

if $a_i < 0$, let $\frac{\partial g_i}{\partial x_i} = y_i - \frac{a_i}{x_i} = 0 \Rightarrow x = \frac{a_i}{y_i}$

Therefore. $f^*(y) = \begin{cases} \sum_{a_i \neq 0}(a_i - a_i \log\frac{a_i}{y_i}), & a \preceq 0, y \preceq 0 \\ \infty & , \text{otherwise}. \end{cases}$

1.2.

~~解~~ assume $b_1 \le b_2 \le \cdots \le b_n$

1) l_1-norm:

$f(x) = \sum_{i=1}^{n} |x - b_i|$

assume $b_k \le x \le b_{k+1}$

$\therefore f(x) = \sum_{i=1}^{k} (x - b_i) + \sum_{i=k+1}^{n} (b_i - x)$

$\therefore \frac{\partial f(x)}{\partial x} = k - (n-k) = 2k - n$

let $\frac{\partial f(x)}{\partial x} = 0$

$\Rightarrow k = \frac{1}{2} n$

$\therefore x^*$ is the median of $\{b_1, b_2, \dots, b_n\}$

2) l_2-norm:

$f(x) = (x1 - b)^T (x1 - b)$

$= nx^2 - 2b^T 1 x + b^T b$

~~let~~

$\therefore \frac{\partial f(x)}{\partial x} = 2nx - 2\sum_{i=1}^{n} b_i$

let $\frac{\partial f(x)}{\partial x} = 0$

$\Rightarrow 2nx^* - 2\sum_{i=1}^{n} b_i = 0$

$\Rightarrow x^* = \frac{\sum_{i=1}^{n} b_i}{n}$, it's the mean of $\{b_1, b_2, \dots, b_n\}$

3) l_∞-norm:

~~if $x < \frac{b_1+b_n}{2}$~~

$f(x) = \max\{|x - b_1|, |x - b_2|, \dots |x - b_n|\}$.

$\therefore$ if $x < \frac{b_1+b_n}{2}$: $f(x) = b_n - x \Rightarrow \frac{\partial f(x)}{\partial x} < 0$

if $x > \frac{b_1+b_n}{2}$: $f(x) = x - b_1 \Rightarrow \frac{\partial f(x)}{\partial x} > 0$

$\therefore$ if $x = \frac{b_1+b_n}{2}$: $f(x) = \frac{b_n - b_1}{2}$ and $\frac{\partial f(x)}{\partial x} = 0$

$\therefore x^* = \frac{b_1+b_n}{2}$

4) l_2-norm for minimize$_x$ $\|x a - b\|_2$

$f(x) = (ax-b)^T(ax-b) = a^Ta x^2 - 2b^Ta x + b^Tb$

$\therefore \frac{\partial f(x)}{\partial x} = 2a^Tax - 2a^Tb$

let $\frac{\partial f(x)}{\partial x} = 0$

$\Rightarrow \quad a^Tax^* = a^Tb$

if a^Ta is invertable

$x^* = (a^Ta)^{-1}a^Tb$

otherwise,

there are infinite solutions for $a^Tax^* = a^Tb$.

6.6.

a) $L(x, r, v) = \sum_{i=1}^m \phi(r_i) + v^T(Ax - b - r)$

The Lagrange dual function is

$$g(v) = \inf_{x,r}\left(\sum_{i=1}^m \phi(r_i) + v^TAx - v^Tb - v^Tr\right)$$

$$\therefore g(v) = \begin{cases} -v^Tb + \inf_r\left(\sum_{i=1}^m \phi(r_i) - v^Tr\right), & A^Tv = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

$$\because \inf_r\left(\sum_{i=1}^m \phi(r_i) - v^Tr\right) = -\sup_r\left(v^Tr - \sum_{i=1}^m \phi(r_i)\right) = -\phi^*(r_i)$$

$$\because \phi(u) = \begin{cases} 0, & |u| \le 1 \\ |u| - 1, & |u| > 1 \end{cases}$$

$$\therefore \phi^*(y) = \sup(yu - \phi(u)) = \begin{cases} |y|, & |y| \le 1 \\ \infty, & |y| > 1 \end{cases}$$

$$\therefore g(v) = \begin{cases} -v^Tb - \sum_{i=1}^m \phi^*(r_i), & A^Tv = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

$\therefore$ the dual is.

maximize $\quad -v^Tb - \|v\|_1$

s.t. $\quad A^Tv=0, \ \|v\|_\infty \le 1$

b) same as (a),

$$g(v) = \begin{cases} -v^Tb - \sum_{i=1}^m \phi^*(r_i), & A^Tv=0 \\ -\infty, & \end{cases}$$

$\because \phi(u)=\begin{cases} u^2, & |u|\le 1\\ 2|u|-1, & |u|>1\end{cases}$

$$\therefore \phi^*(y) = \sup(yu-\phi(u)) = \begin{cases}\sup(yu-u^2), & |u|\le 1 \\ \sup(yu-2|u|-1), & |u|>1\end{cases}$$

$$\therefore \phi^*(y)=\begin{cases} \frac{y^2}{4}, & |y|\le 2\\ \infty, & |y|>0.\end{cases}$$

$\therefore$ the dual is

maximaize $\quad -v^Tb - \frac{1}{4}\|v\|_2^2$

s.t. $\quad A^Tb=0, \ \|v\|_\infty\le 2$

7.6.

~~Bs |y| b/a~~

$\because$ ~~$F_Y(y) = P\{y$~~

$$y = \frac{x+b}{a}$$

$\therefore F_Y(y) = P\{yY\le y\} = P\{X\le ay-b\} = F_X(ay-b)$

$\therefore p_Y(y) = \frac{dF(y)}{dy} = a p_x(ay-b)$

$\therefore$ ~~$\log p_Y$~~ The log-likelihood function is

$\therefore \log p_Y(y) = \cancel{\log a}\ \log a + \log p(ay-b)$

$\because$ p is log-concave

$\therefore$ $\log p(x)$ is concave

$\therefore$ $\log p(ay-b)$ is concave

$\therefore$ The log-likelihood function is a concave function for a and b

Therefore, finding the ML estimate of a and b, given samples $y_1, y_2, \cdots, y_n$ is convex problem.

$$\text{maximize} \quad n\log n + \sum_{i=1}^{n} \log p(ay_i - b)$$

For Laplace distribution,

$$p(x) = e^{-2|x|}$$

$\therefore$ ~~the~~ the problem is

$$\text{maximize} \quad \cancel{nl}\ n\log n - 2\sum_{i=1}^{n} |ay_i - b|$$

~~$\Leftrightarrow$ maximize $n\log n$~~

assume $a \neq 0$.

$\Leftrightarrow$ maximize $n\log n - 2a\sum_{i=1}^{n} |y_i - \frac{b}{a}|$

according to 6.2 (1),

$(\frac{b}{a})^*$ is the median of $\{y_1, \cdots, y_n\}$.

$\therefore$ ~~the~~

$\therefore$ the solution for the ML estimate of a and b should satisfy $\frac{b}{a}$ equals the median of $\{y_1, y_2, \ldots, y_n\}$.

8.24.

consider ~~for~~ $f(u) = (a+u)^T x$, $\|u\|_2 \leq \rho$.

the minimum is $a^T x - \rho\|x\|_2$, where

$\|u\|_2 = \rho$ and it's direction is in opposite of x.

the maximum is $a^T x + \rho\|x\|_2$, where

$\|u\|_2 = \rho$ and its direction is same as x.

由 扫描全能王 扫描创建

~~∴ (a+u)~~

∴ $(a+u)^T x \ge b$ hold for all u with $\|u\| \le \rho$

$\Leftrightarrow$ ~~(a+~~ $a^T x - \rho \|x\|_2 \ge b$ ~~for~~

∴ $(a+u)^T y \le b$ hold for all u with $\|u\| \le \rho$

$\Leftrightarrow$ $a^T y + \rho \|y\|_2 \le b$

∴ the problem equals to.

$$
\begin{aligned}
\text{maximize} \quad & \rho \\
\text{s.t} \quad & a^T x_i - b \ge \rho \|x_i\|_2 , \quad i = 1, 2, \dots, N \\
& b - a^T y_i \ge \rho \|y_i\|_2 , \quad i = 1, 2, \dots, M \\
& \|a\| \le 1
\end{aligned}
$$