Additional Problems

Santhosh Nadig

February 1, 2018

A2.30

We need to show that the infimal convolution of two functions f and g on $\mathbf{R}^n$ with $f(x) = \|x\|_1$ and $g(x) = (1/2)\|x\|_2^2$ given by

$$h(x) = \inf_y (\|y\|_1 + \frac{1}{2}\|x - y\|_2^2)$$

is the Huber Penalty

$$h(x) = \sum_{i=1}^{n} \phi(x_i), \qquad \phi(u) = \begin{cases} u^2 & |u| \leq 1 \\ |u| - \frac{1}{2} & |u| > 1. \end{cases}$$

For $y_i < 0$, the argument that minimizes $h(x_i)$ is given by setting $h_1'(x_i) = -1 - (x_i - y_i)$ to 0, yielding $y_i^* = x_i + 1$ with the condition $x_i < -1$. Similarly, for $y_i > 0$, we get $y_i^* = x_i - 1$ with the condition $x_i > 1$. When $x_i \in [-1, 1]$, the function is non-differentiable meaning $y_i^* = 0$. Thus, we have,

$$y_i^* = \begin{cases} x_i - 1 & x_i > 1 \\ 0, & -1 \leq x_i \leq 1 \\ x_i + 1 & x_i < -1. \end{cases} \tag{1}$$

Substituting $y_i = y_i^*$ for each element in $h(x)$, we get

$$\phi(x_i) = \begin{cases} x_i^2 & |x_i| \leq 1 \\ |x_i| - \frac{1}{2} & |x_i| > 1. \end{cases} \tag{2}$$

which is the Huber loss.

A2.31 (b,c)

The function $h : \mathbf{R} \to \mathbf{R}$ is convex, non-decreasing, with dom $h = \mathbf{R}$ and $h(t) = h(0), t \leq 0$. We need to show that the conjugate of a function $f(x) = h(\|x\|_2)$ with $x \in \mathbf{R}^n$ is $f^*(y) = h^*(\|y\|_2)$.

$$f^*(y) = \sup_{x \in \mathbf{R}^n} (y^T x - f(x))$$

From the property of dual-norm, we have $y^T x \leq \|y\|_* \|x\|$, where the strict equality holds for some optimal y. Using the fact that the Euclidean norm is self-dual, we may write

$$\begin{aligned} f^*(y) &= \sup_{x \in \mathbf{R}^n} (\|y\|_2 \|x\|_2 - f(x)) \\ &= \sup_{x \in \mathbf{R}^n} (\|y\|_2 \|x\|_2 - h(\|x\|_2)) \\ &= \sup_{\|x\|_2 \in \mathbf{R}} (\|y\|_2 \|x\|_2 - h(\|x\|_2)) = h^*(\|y\|_2) \end{aligned} \tag{3}$$

Using this result, we need to derive the conjugate of $f(x) = 1/p \|x\|_2^p$ with $h(t)$ defined as

$$h(t) = \frac{1}{p} \max\{0, t\}^p = \begin{cases} \frac{1}{p} t^p, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

The conjugate of $h(x)$ is given by

$$h^*(z) = \sup_{x \in \mathbf{R}} (zx - 1/p x^p)$$

By setting the derivative w.r.t x of the above to 0, we find that the x that achieves the maximum is $x = z^{1/(p-1)}$. Thus,

$$\begin{aligned} h^*(z) &= z \cdot z^{1/(p-1)} - 1/p z^{p/(p-1)} \\ &= \frac{p-1}{p} z^{p/(p-1))} \\ &= \frac{1}{p/(p-1)} z^{p/(p-1)} \end{aligned}$$

Using 3, we have

$$f^*(y) = \frac{1}{p/(p-1)} \|y\|_2^{p/(p-1)} \tag{4}$$

We can easily verify that $f^{**} = f$ (since f is closed, convex etc.)

$$\begin{aligned} f^{**}(x) &= \frac{\frac{1}{p/(p-1)}}{(p/(p-1)-1)} \|x\|_2^{\frac{p/(p-1)}{(p/(p-1)-1)}} \\ &= \frac{\frac{1}{p/(p-1)}}{\frac{1}{(p-1)}} \|x\|_2^{\frac{p/(p-1)}{1/(p-1)}} \\ &= \frac{1}{p} \|x\|_2^p = f(x). \end{aligned}$$