Homework 6

Requirements:

1. Digital format (can be typeset or photos, ought to write clearly if written by hand), upload to `https://course.pku.edu.cn/`.
2. Submit by next class
3. A problem is not counted if nobody can work it out
4. Each homework 10 points; 1 point deducted for each week's delay

Problems:

1. Show that the following functions are convex.

(a) $f(\mathbf{x}) = -\log\left(-\log\left(\sum_{i=1}^{m} \exp\left(\mathbf{a}_i^T\mathbf{x} + b_i\right)\right)\right)$ on

$$\text{dom } f = \left\{ \mathbf{x} \,\middle|\, \sum_{i=1}^{m} \exp\left(\mathbf{a}_i^T\mathbf{x} + b_i\right) < 1 \right\}.$$

You can use the fact that $\log\left(\sum_{i=1}^{n} e^{y_i}\right)$ is convex.

(b) $f(\mathbf{x}, u, v) = -\sqrt{uv - \mathbf{x}^T\mathbf{x}}$ on $\text{dom } f = \{(\mathbf{x}, u, v) | uv > \mathbf{x}^T\mathbf{x}, u, v > 0\}$. Use the fact that $\mathbf{x}^T\mathbf{x}/u$ is convex in $(\mathbf{x}, u)$ for $u > 0$, and that $-\sqrt{x_1x_2}$ is convex on $\mathbb{R}^2_{++}$.

(c) $f(\mathbf{x}, u, v) = -\log(uv - \mathbf{x}^T\mathbf{x})$ on $\text{dom } f = \{(\mathbf{x}, u, v) | uv > \mathbf{x}^T\mathbf{x}, u, v > 0\}$.

(d) $f(\mathbf{x}, t) = -\left(t^p - \|\mathbf{x}\|_p^p\right)^{1/p}$ where $p > 1$ and $\text{dom } f = \{(\mathbf{x}, t) | t \geq \|\mathbf{x}\|_p\}$. You can use the fact that $\|\mathbf{x}\|_p^p/u^{p-1}$ is convex in $(\mathbf{x}, u)$ for $u > 0$, and that $-x^{1/p}y^{1-1/p}$ is convex on $\mathbb{R}^2_+$.

(e) $f(\mathbf{x}, t) = -\log\left(t^p - \|\mathbf{x}\|_p^p\right)$ where $p > 1$ and $\text{dom } f = \{(\mathbf{x}, t) | t > \|\mathbf{x}\|_p\}$. You can use the fact that $\|\mathbf{x}\|_p^p/u^{p-1}$ is convex in $(\mathbf{x}, u)$ for $u > 0$.

2. Show that the cross entropy $H(\mathbf{w}) = -\sum_{i=1}^{N} [y_i \ln h(\mathbf{x}_i) + (1 - y_i)\ln(1 - h(\mathbf{x}_i))]$ is a convex function on $\mathbb{R}^n$, where $h(\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x})$, $y_i \in \{0, 1\}$, and $\sigma(x) = 1/(1 + e^{-x})$ is the logistic (also called sigmoid) function.

3. Derive the conjugates of the following functions.

(a) Sum of largest elements. $f(\mathbf{x}) = \sum_{i=1}^{r} x_{[i]}$ *on* $\mathbb{R}^n$.

(b) Piecewise-linear function on $\mathbb{R}$. $f(x) = \max_{i=1,\ldots,n}(a_ix + b_i)$ *on* $\mathbb{R}$. *You can assume that the* a_i *are sorted in increasing order, i.e.,* $a_1 \leq \ldots \leq a_m$*, and that none of the functions* $a_ix + b_i$ *is redundant, i.e., for each* k *there is at least one* x *with* $f(x) = a_kx + b_k$.

(c) Power function. $f(x) = x^p$ *on* $\mathbb{R}_{++}$, *where* $p > 1$. *Repeat for* $p < 0$.

(d) Geometric mean. $f(\mathbf{x}) = -(\prod x_i)^{1/n}$ *on* $\mathbb{R}^n_{++}$.

(e) Negative generalized logarithm for second-order cone. $f(\mathbf{x}, t) = -\log(t^2 - \mathbf{x}^T\mathbf{x})$ *on* $\{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|\mathbf{x}\|_2 < t\}$.

Note: Computing a conjugate function needs to specify its domain as well.

4. Properties of conjugate functions.

(a) Conjugate of convex plus affine function. *Define* $g(\mathbf{x}) = f(\mathbf{x}) + \mathbf{c}^T\mathbf{x} + d$, *where* f *is convex. Express* g^* *in terms of* f^* *(and* $\mathbf{c}, d$*).*

(b) Conjugate of perspective. *Express the conjugate of the perspective of a convex function* f *in terms of* f^*.

(c) Conjugate and minimization. *Let* $f(\mathbf{x}, \mathbf{z})$ *be convex in* $(\mathbf{x}, \mathbf{z})$ *and define* $g(\mathbf{x}) = \inf_z f(\mathbf{x}, \mathbf{z})$. *Express the conjugate* g^* *in terms of* f^*. *As an application, express the conjugate of* $g(\mathbf{x}) = \inf\{h(\mathbf{z}) \mid \mathbf{A}\mathbf{z} + \mathbf{b} = \mathbf{x}\}$, *where* h *is convex, in terms of* h^*, $\mathbf{A}$, *and* $\mathbf{b}$.