Name: Philip Pham

AMATH 515 Homework Set 3

Due: Monday, March 9th by midnight..

(1) Compute the conjugates of the following functions.

(a) $f(x) = \delta_{\mathbb{B}_\infty}(x)$.

Proof. We have that

$$\boxed{f^\star(x) = \delta^\star_{\mathbb{B}_\infty}(x) = \max_{y \in \mathbb{B}_\infty} \langle x, y \rangle .}$$

By definition,

$$f^\star(x) = \sup_y \left\{ \langle x, y \rangle - f(y) \right\}.$$

If $y \notin \mathbb{B}_\infty$, then $\langle x, y \rangle - f(y) = -\infty$ for all x. For $y \in \mathbb{B}_\infty$, $\langle x, y \rangle - f(y) = \langle x, y \rangle$, so to maximize $\langle x, y \rangle - f(y)$, we should choose y such that $\langle x, y \rangle$ is maximized. □

(b) $f(x) = \delta_{\mathbb{B}_2}(x)$.

Proof. We have that

$$\boxed{f^\star(x) = \delta^\star_{\mathbb{B}_2}(x) = \max_{y \in \mathbb{B}_2} \langle x, y \rangle .}$$

The proof is identical to the previous one with $\mathbb{B}_\infty$ replaced by $\mathbb{B}_2$. □

(c) $f(x) = \exp(x)$.

Proof. We have that

$$\boxed{f^\star(x) = \begin{cases} x \log x - x = x\left(\log x - 1\right), & x > 0; \\ 0, & x = 0; \\ \infty, & x < 0. \end{cases}}$$

To see this, we can maximize $\langle x, y \rangle - \exp(y)$ with respect to y by taking the derivative, setting it to 0, and solving for y. In doing so, we find that $y = \log x$, which is only defined when $x > 0$. When $x \leq 0$, we see that we can maximize $xy - \exp(y)$ by sending y to $-\infty$. □

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

Proof. We have that

$$\boxed{f^\star(x) = \begin{cases} \infty, & x > 1; \\ 0, & x = 1; \\ x\left(\log x - \log\left(1-x\right)\right) + \log\left(1-x\right), & 0 < x < 1; \\ 0, & x = 0; \\ \infty, & x < 0. \end{cases}}$$

Consider $xy - \log(1+\exp(y))$. The derivative with respect to y is $x - \frac{\exp(y)}{1+\exp(y)}$. When $x \in (0,1)$, we can solve for $y = \log\left(\frac{x}{1-x}\right)$. When $x \geq 1$, the derivative is always positive, so we have that the max is obtained when $y \to \infty$. We note that

$$xy - \log(1+\exp(y)) = xy - \left[y + \exp(-y) - \frac{\exp(-2y)}{2} + \frac{\exp(-3y)}{3} - \frac{\exp(-4y)}{4} + \cdots\right]$$

by a Taylor series expansion, so as $y \to \infty$, we have ∞ if $x > 1$ and 0 if $x = 1$.

When $x \leq 0$, the derivative is always negative, so we maximize the expression with $y \to -\infty$. only the second term remains.

$\square$

(e) $f(x) = x\log(x)$

(2) Let g be any convex function; f is formed using g. Compute f^* in terms of g^*.

(a) $f(x) = \lambda g(x)$.

(b) $f(x) = g(x-a) + \langle x, b\rangle$.

(c) $f(x) = \inf_z \{g(x,z)\}$.

(d) $f(x) = \inf_z \left\{\frac{1}{2}\|x-z\|^2 + g(z)\right\}$

(3) Moreau Identities.

(a) Derive the Moreau Identity:

$$\mathrm{prox}_f(z) + \mathrm{prox}_{f^*}(z) = z.$$

You may find the 'Fenchel flip' useful.

(b) Use either of the Moreau identities and 1a, 1b to check your formulas for

$$\mathrm{prox}_{\|\cdot\|_1}, \quad \mathrm{prox}_{\|\cdot\|_2}$$

from last week's homework.

(4) Duals of regularized GLM. Consider the Generalized Linear Model family:

$$\min_x \sum_{i=1}^n g(\langle a_i, x\rangle) - b^T Ax + R(x),$$

Where g is convex and R is any regularizer.

(a) Write down the general dual obtained from the perturbation

$$p(u) = \min_x \sum_{i=1}^n g(\langle a_i, x\rangle + u_i) - b^T Ax + R(x).$$

(b) Specify your formula to Ridge-regularized logistic regression:

$$\min_x \sum_{i=1}^n \log(1 + \exp(\langle a_i, x\rangle)) - b^T Ax + \frac{\lambda}{2}\|x\|^2.$$

(c) Specify your formula to 1-norm regularized Poisson regression:

$$\min_x \sum_{i=1}^n \exp(\langle a_i, x\rangle) - b^T Ax + \lambda\|x\|_1.$$

Coding Assignment

Please download `515Hw3_Coding.ipynb` and `proxes.py` to complete problem (5).

(5) In this problem you will write a routine to project onto the capped simplex.

The Capped Simplex Δ_k is defined as follows:

$$\Delta_k := \left\{x : 1^T x = k, \quad 0 \le x_i \le 1 \quad \forall i.\right\}$$

This is the intersection of the k-simplex with the unit box.

The projection problem is given by

$$\mathrm{proj}_{\Delta_k}(z) = \arg\min_{x\in\Delta_k} \frac{1}{2}\|x - z\|^2.$$

(a) Derive the (1-dimensional) dual problem by focusing on the $1^T x = k$ constraint.

(b) Implement a routine to solve this dual. It's a scalar root finding problem, so you can use the root-finding algorithm provided in the code.

(c) Using the dual solution, write down a closed form formula for the projection. Use this formula, along with your dual solver, to implement the projection. You can use the unit test provided to check if your code id working correctly.