Homework 3

Due: March 3, 2021

Let f be a closed proper convex function. The convex conjugate of f, called f^*, is defined by

$$f^*(z) = \sup_x \left\{ z^T x - f(x) \right\}.$$

1. Compute the conjugates of the following functions.

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

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

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

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

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

   Solution.

   (a) By definition

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x \rangle - \delta_{\mathbb{B}_\infty}(x) \\ &= \sup_{x \in \mathbb{B}_\infty} \langle z, x \rangle \end{aligned}$$

   The sup is achieved if $x_i = z_i$, then

$$f^*(z) = \|z\|_1$$

   (b) By definition

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x \rangle - \delta_{\mathbb{B}_2}(x) \\ &= \sup_{x \in \mathbb{B}_2} \langle z, x \rangle \end{aligned}$$

   The sup is achieved if $x_i = \frac{z_i}{\|z\|_2}$, then

$$f^*(z) = \|z\|_2$$

   (c) The condition for sup is

$$\begin{aligned} \nabla(\langle z, x \rangle - \exp(x)) = z - \exp(x) = 0 \\ x = \log z \end{aligned}$$

Then

$$f^*(z) = z \log z - z$$

(d) The condition for sup is

$$\begin{aligned}
\nabla(\langle z, x\rangle - \log(1+\exp(x))) &= z - \frac{\exp(x)}{1+\exp(x)} = 0 \\
\frac{z}{1-z} &= \exp(x) \\
x &= \log(\frac{z}{1-z})
\end{aligned}$$

Then

$$f^*(z) = z\log(\frac{z}{1-z}) - \log(\frac{1}{1-z})$$

(e) The condition for sup is

$$\begin{aligned}
\nabla(\langle z, x\rangle - x\log(x)) &= z - (\log(x)+1) = 0 \\
x &= \exp(z-1)
\end{aligned}$$

Then

$$f^*(z) = \exp(z-1)$$

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

   Solution.

   (a) By definition

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x\rangle - \lambda g(x) \\ &= \lambda(\sup_x \langle \frac{z}{\lambda}, x\rangle - g(x)) \\ &= \lambda g^*(\frac{z}{\lambda}) \end{aligned}$$

   (b) By definition

$$\begin{aligned} f^*(z) &= \sup_x \langle z, x\rangle - g(x-a) - \langle x, b\rangle \\ &= \sup_x \langle z-b, x\rangle - g(x-a) \\ &= \sup_x \langle z-b, x+a\rangle - g(x) \\ &= \langle z-b, a\rangle + \sup_x \langle z-b, x\rangle - g(x) \\ &= \langle z-b, a\rangle + g^*(z-b) \end{aligned}$$

   (c) By definition

$$\begin{aligned} f^*(y) &= \sup_x \left(\langle y, x\rangle - \inf_z \{g(x,z)\}\right) \\ &= \sup_x \left(\sup_z \langle y, x\rangle - g(x,z)\right) \\ &= \sup_{x,z} \langle y, x\rangle - g(x,z) \\ &= \sup_{x,z} \langle y, x\rangle + \langle 0, z\rangle - g(x,z) \\ &= \sup_{x,z} \langle [y,0], [x,z]\rangle - g(x,z) \\ &= g^*(y,0) \end{aligned}$$

   (d) Let

$$h(x,z) = \frac{1}{2}\|x-z\|^2 + g(z)$$

Then

$$
\begin{aligned}
f^*(y) &= h^*(y, 0) \\
&= \sup_{x,z} \langle [y, 0], [x, z] \rangle - h(x, z) \\
&= \sup_{x,z} \langle y, x \rangle - \frac{1}{2}\|x - z\|^2 - g(z) \\
&= \sup_{x,z} \langle y, x - z \rangle + \langle y, z \rangle - \frac{1}{2}\|x - z\|^2 - g(z)
\end{aligned}
$$

Let $x' = x - z$

$$
\begin{aligned}
f^*(y) &= \sup_{x',z} \langle y, x' \rangle + \langle y, z \rangle - \frac{1}{2}\|x'\|^2 - g(z) \\
&= \sup_{x',z} \langle y, x' \rangle - \frac{1}{2}\|x'\|^2 + \sup_{z} \langle y, z \rangle - g(z) \\
&= \frac{1}{2}\|y\|^2 + g^*(y)
\end{aligned}
$$

3. Moreau Identities.

   (a) Derive the Moreau Identity:

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

   (b) Use the Moreau identity and 1a, 1b to check your formulas for

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

   from last week's homework.

Solution.

(a) By definition

$$\begin{aligned}
\text{prox}_f(z) &= \arg\min_{x_1} \frac{1}{2}\|x_1 - z\|^2 + f(x_1) \\
\text{prox}_{f^*}(z) &= \arg\min_{x_2} \frac{1}{2}\|x_2 - z\|^2 + f^*(x_2)
\end{aligned}$$

The optimality condition is

$$\begin{aligned}
0 &\in x_1 - z + \partial f(x_1) \\
0 &\in x_2 - z + \partial f^*(x_2)
\end{aligned}$$

Rearranging

$$\begin{aligned}
z - x_1 &\in \partial f(x_1) \\
z - x_2 &\in \partial f^*(x_2)
\end{aligned}$$

Since f is closed, proper and convex

$$x_1 \in \partial f^*(z - x_1)$$

Let $z - x_1 = y$, then

$$z - y \in \partial f^*(y)$$

This is identical to $z - x_2 \in \partial f^*(x_2)$, then $y = x_2$ and

$$z = x_1 + x_2 = \text{prox}_f(z) + \text{prox}_{f^*}(z)$$

(b) From last week's homework we have

$$\text{prox}_f(z) = \begin{cases} z_i + 1 & z_i < -1 \\ z_i - 1 & z_i > 1 \\ 0 & |z_i| \leq 1 \end{cases}$$

Since $f = \|\cdot\|_1$, $f^* = \delta_{\mathbb{B}_\infty}$, then

$$\begin{aligned}
\operatorname{prox}_{f^*}(z) &= \arg\min_x \frac{1}{2}\|x - z\|^2 + \delta_{\mathbb{B}_\infty}(x) \\
&= \arg\min_{x \in \mathbb{B}_\infty} \frac{1}{2}\|x - z\|^2 \\
&= \operatorname{proj}_{\mathbb{B}_\infty}(z) \\
&= \begin{cases} -1 & z_i < -1 \\ 1 & z_i > 1 \\ z_i & |z_i| \leq 1 \end{cases}
\end{aligned}$$

Then the Moreau identity holds. From last week's homework we have

$$\operatorname{prox}_f(z) = \begin{cases} (1 - \frac{1}{\|z\|})z & \|z\| > 1 \\ 0 & \|z\| \leq 1 \end{cases}$$

Since $f = \|\cdot\|_2$, $f^* = \delta_{\mathbb{B}_2}$, then

$$\begin{aligned}
\operatorname{prox}_{f^*}(z) &= \operatorname{proj}_{\mathbb{B}_2}(z) \\
&= \begin{cases} \frac{z}{\|z\|} & \|z\| > 1 \\ z & \|z\| \leq 1 \end{cases}
\end{aligned}$$

Then the Moreau identity holds.

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 dual obtained by dualizing g.

(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:

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

Solution.

(a) We worked out in the lecture that, if we have a general primal problem

$$\min_x f(x) + h(Ex-d) + \langle c, x\rangle$$

Then the dual problem is

$$\sup_z -\langle z, d\rangle - h^*(z) - f^*(-A^T z - c)$$

For the GLM family, we have

$$f(x) = R(x), \quad h(x) = \sum_{i=1}^n g(x_i), \quad E = A, \quad d = 0, \quad c = -A^T b$$

Then the resulting dual problem is

$$\sup_z -\langle z, 0\rangle - \sum_{i=1}^n g^*(z_i) - R^*(-A^T z + A^T b) = \sup_z -\sum_{i=1}^n g^*(z_i) - R^*(A^T(b-z))$$

(b) In this case, $g(x) = \log(1+\exp(x))$ and $R(x) = \frac{\lambda}{2}\|x\|^2$. We worked out in 1d that

$$g^*(z) = z\log(\frac{z}{1-z}) - \log(\frac{1}{1-z})$$

For $R(x)$ we have

$$R^*(z) = \sup_x \langle z, x\rangle - \frac{\lambda}{2}\|x\|^2$$

The condition for sup is

$$\begin{aligned} z - \lambda x &= 0 \\ x &= \frac{z}{\lambda} \end{aligned}$$

Then we have

$$R^*(z) = \frac{\|z\|^2}{2\lambda}$$

Putting everything together, the dual problem is

$$\sup_z -\langle z, \log(\frac{z}{1-z})\rangle + \sum_{i=1}^{n} \log(\frac{1}{1-z_i}) - \frac{\|A^T(b-z)\|^2}{2\lambda}$$

(c) In this case, $g(x) = \exp(x)$ and $R(x) = \lambda\|x\|_1$. We worked out in 1c that

$$g^*(z) = z\log z - z$$

From 2b we know that

$$R^*(z) = \delta_{\lambda\mathbb{B}_\infty}(z)$$

Putting everything together, the dual problem is

$$\sup_z -\langle z, \log z\rangle + \sum_{i=1}^{n} z_i - \delta_{\lambda\mathbb{B}_\infty}(A^T(b-z))$$

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

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

(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 is working correctly.

Solution.

(a) The dual problem can be formulated as

$$\min_{x \in [0,1]^n} \sup_{\tau} \frac{1}{2}\|x - z\|^2 + \tau(1^T x - k) = \sup_{\tau} \min_{x \in [0,1]^n} \frac{1}{2}\|x - z\|^2 + \tau 1^T x - \tau k$$

The optimality condition in x is

$$x - z + \tau = 0$$
$$x = z - \tau$$

Since x is constrained in the unit box, this minimizer is true when $0 \le z_i - \tau \le 1$. Outside the unit box we will take the endpoints. Then

$$x^* = \begin{cases} z_i - \tau & \tau \le z_i \le \tau + 1 \\ 1 & z_i > \tau + 1 \\ 0 & z_i < \tau \end{cases} \tag{1}$$

Define the object in the dual problem as $f_\tau(z)$ and substitute in x^*

$$f(\tau) = \begin{cases} \frac{1}{2}\tau^2 + \tau(z_i - \tau) - \tau k & \tau \le z_i \le \tau + 1 \\ \frac{1}{2}|1 - z_i|^2 + \tau - \tau k & z_i > \tau + 1 \\ \frac{1}{2}|z_i|^2 - \tau k & z_i < \tau \end{cases}$$

Then the dual problem is

$$\sup_{\tau} f(\tau)$$

(b) We wish to solve

$$0 = f'(\tau) = -k + \sum_i \begin{cases} z_i - \tau & \tau \leq z_i \leq \tau + 1 \\ 1 & z_i > \tau + 1 \\ 0 & z_i < \tau \end{cases}$$

(c) After we compute τ, we can get x using the formula in (1).