Name: Hongda Li
Class: AMATH 515

Problem 1

(a)

Objective: Find the conjugate of $\delta_{\mathbb{B}_\infty}(x)$. The function has definition:

$$\delta_{\mathbb{B}_\infty} = \begin{cases} 0 & \|x\|_\infty \leq 1 \\ \infty & \text{else} \end{cases}$$

The definition of conjugate is:

$$\delta_{\mathbb{B}_\infty}^*(z) = \sup_x \left\{ z^T x - \delta_{\mathbb{B}_\infty}(x) \right\} \tag{1a1}$$

$$= \sup_{\|x\|_\infty} \left\{ z^T x \right\}$$

$$= \max_{x_i} \left\{ \sum_{i=1}^n z_i x_i : |x_i| \leq 1 \right\}$$

$$= \sum_{i=1}^n |z_i|$$

$$= \|z\|_1$$

(b)

Objective: Find conjugate for indication function of the 2 norm. From the definition of the function we have:

$$\delta_{\mathbb{B}_2}(x) = \begin{cases} 0 & \|x\|_2 \leq 1 \\ \infty & \text{else} \end{cases}$$

Then by definition of the conjugate we have:

$$\delta_{\mathbb{B}_2}^*(z) = \sup_x \left\{ z^T x - \delta_{\mathbb{B}_2}(x) \right\} \tag{1b1}$$

$$= \sup_{\|x\|_2 \leq 1} \left\{ z^T x \right\}$$

$$= \frac{z^T z}{\|z\|_2}$$

$$= \|z\|_2$$

(c)

Objective: Find conjugate of $\exp(x)$, assuming vectorized exponential. Notice that the function is smooth and concave and it's univariate.

$$\exp^*(z) = \sup_x \left\{ zx - \exp(x) \right\} \tag{1c1}$$

Assume $z > 0$ consider:

$$\partial_x[zx - \exp(x)] = 0 \tag{1c2}$$

$$[z - \exp(x)] = 0$$

$$x = \log(z)$$

$$\implies \exp^*(z) = \sup_x \{zx - \exp(x)\} = z \ln(z) - z$$

1

Assume $z = 0$, we have:
$$\exp^*(0) = \sup_x \{-\exp(x)\} = 0 \tag{1c3}$$

Assume $z < 0$, we have:
$$\exp^*(z) = \sup_x \{-\exp(x)\} = \infty \tag{1c4}$$

This is because $\lim_{x \to -\infty}(zx - \exp(x))$ gives ∞ if $z < 0$ Therefore:

$$\exp^*(z) = \begin{cases} z\ln(z) - z & z > 0 \\ 0 & z = 0 \\ \infty & z < 0 \end{cases}$$

Taking the limit on the first case, we can merge the second case and the first case.

(d)

Objective: Find the conjugate of the function $f(x) = \ln(1 + \exp(x))$. Notice that the function is convex and smooth, therefore, the inverse of it will be concave, hence setting the derivative to zero will provide us with the optimal.
By definition of conjugacy we have:

$$f^*(z) = \sup_x \{zx - \ln(1 + \exp(x))\}$$

consider the partial derivative wrt to x:

$$\partial_x [zx - \ln(1 + \exp(x))] = 0 \tag{1d1}$$

$$z = \frac{\exp(x)}{1 + \exp(x)}$$

$$z(1 + \exp(x)) = \exp(x)$$

$$z + z\exp(x) = \exp(x)$$

$$z = (1 - z)\exp(x)$$

$$\frac{z}{1 - z} = \exp(x)$$

$$x = \ln\left(\frac{z}{1 - z}\right) \quad \text{Substitute this back}$$

$$\implies f^*(z) = z\ln\left(\frac{z}{1 - z}\right) - \ln\left(1 + \frac{z}{1 - z}\right)$$

It's not all sunshine and rainbow, we have some conditions for z:

$$\text{sign}(z) = \text{sign}(1 - z) \implies z \in (0, 1) \tag{1d2}$$

We can try to cover 2 of the edge cases of when $z = 0$ and $z = 1$ first, notice that this means we have:

$$\sup_x \{x - \ln(1 + \exp(x))\} = 0 \tag{1d3}$$

$$\text{Notice: } \ln(1 + \exp(x)) > x \quad \forall x$$

$$\lim_{x \to \infty} (x - \ln(1 + \exp(x))) = \lim_{x \to \infty} (\ln(\exp(x)) - \ln(1 + \exp(x)))$$

$$= \lim_{x \to \infty} \left(\ln\left(\frac{\exp(x)}{1 + \exp(x)}\right)\right) = 0$$

It's bounded above by zero. Similarly, when $z = 0$, we have:

$$\sup_x \{-\ln(1 + \exp(x))\} = 0 \tag{1d4}$$

This is bounded above by zero, and $x \to -\infty$, this approaches zero.

2

1. Consider $z < 0$ then $\sup_x\{zx - \ln(1 + \exp(x))\} = \infty$ because as $x \to -\infty$, the expression is approaching infinity.

2. Consider $z > 1$ then $\sup_x\{zx - \ln(1 + \exp(x))\} = \infty$ because:

$$\lim_{x \to \infty} (zx - \ln(1 + \exp(x))) = \lim_{x \to \infty} (\ln(\exp(zx)) - \ln(1 + \exp(x))) = \lim_{x \to \infty} \left(\ln \left(\frac{\exp(zx)}{1 + \exp(x)} \right) \right) = \infty$$

$$(1d5)$$

Summing all the cases up we have the conjugate function to be:

$$f^*(x) = \begin{cases} z \ln \left(\frac{z}{1-z} \right) - \ln \left(\frac{1}{1-z} \right) & z \in (0, 1) \\ 0 & z = 1 \vee z = 0 \\ \infty & \text{else} \end{cases} \qquad (1d6)$$

Note: The second case can be merged into the first case, if we take the limit.

(e)

Objective: Look for the conjugate of $f(x) := x \ln(x)$. Notice that the function is convex and smooth. The conjugate is defined to be:

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

And noticing that we can take thd derivative, setting it to zero and then solve for x, so we can get the value coming out of sup, we treat this function as a univariate function.

$$\partial_x[zx - x\ln(x)] = 0 \qquad (1e1)$$
$$z - (\ln(x) + 1) = 0$$
$$z - 1 = \ln(x)$$
$$\exp(z - 1) = x$$

No edge cases, this is well-defined for $x > 0$. Substituting it back to the expression inside of sup, we have:

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

Problem 2

(a)

Objective: Express the conjugate of the function $f(x) := \lambda g(x)$ in terms of the conjugate of $g(x)$.

$$f^*(z) = \sup_x \left\{ z^T x - \lambda g(x) \right\} \qquad (2a1)$$

$$= \lambda \sup_x \left\{ \frac{z^T}{\lambda} x - g(x) \right\}$$

$$= \lambda g \left(\frac{z}{\lambda} \right)$$

(b)

Objective: $f(x) := g(x-a) + b^T x$ find conjugate of $f(x)$ in term of conjugate of g.

$$f^*(z) = \sup_x \left\{ z^T x - g(x-a) - b^T x \right\} \tag{2b1}$$

$$= \sup_x \left\{ (z^T - b^T)x - g(x-a) \right\} \quad \text{let: } y = x - a$$

$$= \sup_{y+a} \left\{ (z^T - b^T)(y+a) - g(y) \right\}$$

$$= \sup_{y+a} \left\{ (z^T - b^T)y + g(y) \right\} + (z^T - b^T)a$$

$$= g^*(z) + (z^T - b^T)a$$

(c)

Objective: Figure out conjugate of $f(x) := \inf_z \{g(x,z)\}$ using conjugate of $g(x,y)$.
The format of taking the conjugate of a bi-variable function is hinted in the lecture, and let's define it to be:

$$g^*(w,z) = \sup_{x,y} \left\{ w^T x + z^T y - g(x,y) \right\}$$

So then we can start with the definition of the conjugate of f:

$$f^*(y) = \sup_x \left\{ y^T x - \inf_z \{g(x,z)\} \right\} \tag{2c1}$$

$$= \sup_x \left\{ y^T x + \sup_z \{-g(x,z)\} \right\}$$

$$= \sup_{x,z} \left\{ y^T x + w^T \mathbf{0} - g(x,z) \right\}$$

$$= g^*(y, \mathbf{0})$$

(d)

Objective: Find the conjugate of $f(x) := \inf_z \{\frac{1}{2}\|x-z\|^2 + g(z)\}$
Notice that, the expression fits the pattern described in (c), and therefore, we are interested in looking for the conjugate of the function inside the inf:

$$h(x,y) := \frac{1}{2}\|x-y\|^2 + g(y) \quad f(x) = \inf_z \{h(x,z)\}$$

Let's work it out:

$$h^*(v,w) = \sup_{x,y} \left\{ \underbrace{v^T x + w^T y - \frac{1}{2}\|x-y\|^2 - g(y)}_{\text{smooth wrt x}} \right\} \tag{2d1}$$

$$\nabla_x \left[v^T x + w^T y - \frac{1}{2}\|x-y\|^2 - g(y) \right] = \mathbf{0}$$

$$v - (x-y) = \mathbf{0}$$

$$x = v + y$$

$$h^*(v,w) = \sup_y \left\{ v^T(v+y) + w^t y - \frac{1}{2}\|v\|^2 - g(y) \right\}$$

$$= \sup_y \left\{ (v^T + w^T)y - g(y) \right\} + \frac{1}{2}\|v\|^2$$

$$= g^*(v+w) + \frac{1}{2}\|v\|^2$$

4

Using 2c1, we have:

$$f^*(y) = h^*(y, \mathbf{0}) \tag{2d2}$$

$$f^*(y) = g^*(y) + \frac{1}{2}\|y\|^2$$

Problem 3

(a)

Objective: Derive the equality:

$$\operatorname*{prox}_{f}(z) + \operatorname*{prox}_{f^*}(z) = z$$

With the assumption the f is closed and convex.

Claim 3a1

$$u = \operatorname*{prox}_{f}(z) \iff z - u \in \partial f(u)$$

Claim 3a2

$$z \in \partial f(x) \iff x \in \partial f^*(z)$$

Proof:

$$z - u \in \partial(u) \tag{3a3}$$
$$u \in \partial f^*(z - u) \quad \text{by 3a2}$$
$$z - (z - u) \in \partial f^*(z - u)$$
$$z - u = \operatorname*{prox}_{f^*}(z) \quad \text{by 3a1}$$

Therefore:

$$\operatorname*{prox}_{f}(z) + \operatorname*{prox}_{f^*}(z) = z - u + u = z$$

(b)

Objective: Verify part (a) using $\operatorname{prox}_{\|\cdot\|_1}, \operatorname{prox}_{\|\cdot\|_2}$ From part (1)(a), we conclude that:

$$\delta^*_{\mathbb{B}_\infty}(x) = \delta_{\mathbb{B}_1}(x)$$

From previous HW2 we gathered that:

$$\left(\operatorname*{prox}_{1,\|\cdot\|_1}(y)\right)_i = \begin{cases} 0 & y \in [-1, 1] \\ y_i + 1 & y_i < -1 \\ y_i - 1 & y_i > 1 \end{cases}$$

$$\left(\operatorname*{prox}_{1,\|\cdot\|_\infty}(y)\right)_i = \begin{cases} y_i & y_i \in [-1, 1] \\ -1 & y_i < -1 \\ 1 & y_i > 1 \end{cases}$$

If I add them together, element-wise, each case matches each other, cancelling out, and then we just have the vector y.

For the second part, we need the prox with the $\|\cdot\|_2$, and the conjugate of $\|\cdot\|_2$ to verify.

From 1b1 we have:

$$(\delta_{\mathbb{B}_2}(x))^* = \|x\|_2$$

From Previous HW we know that:

$$\operatorname*{prox}_{\|\cdot\|}(z) = \begin{cases} (\|z\|_2 - t)\widehat{z} & \|z\|_2 > t \\ \mathbf{0} & \|z\|_2 \leq t \end{cases}$$

5

And the prox for the $\delta_{\mathbb{B}_2}$ is like:

$$\operatorname*{prox}_{\delta_{\mathbb{B}_2}}(z) = \operatorname*{proj}_{\mathbb{B}_2}(z) = \begin{cases} \widehat{z} & \|z\|_2 \geq 1 \\ z & \text{else} \end{cases}$$

Adding then case by case we have:

$$(\|z\|_2 - t)\widehat{z} + \widehat{z} = \|z\|_2\widehat{z} = z \quad \mathbf{0} + z = z$$

Under both cases, the identity is verified.

Problem 4

(a)

Claim 4a1: From the textbook, we have the following Dualization process from the Fenchel Legendre Transform:

$$\inf_x \{h(Ax) + f(x)\} = \sup_y \{-h^*(y) - f^*(-A^T y)\}$$

Objective: Find the dual for the generalized MLE models, which is in the format of:

$$\min_x \left\{ \underbrace{\sum_{i=1}^n g(a_i^T x) - b^T Ax}_{h(Ax)} + \underbrace{R(x)}_{f(x)} \right\}$$

And therefore, the remaining part is to for ues to identify the solution for $h^*(z)$, the conjugate in term of $g(x)$. Take note, we made the assumption that $g : \mathbb{R} \mapsto \mathbb{R}$ and the conjugate of this function is easy to find.

$$h(Ax) = \sum_{i=1}^n \left(g(a_i^T x)\right) - b^T Ax \tag{4a2}$$

$$\implies h(x) = \sum_{i=1}^n \left(g(x_i)\right) - b^T x$$

$$\implies h^*(z) = \sup_x \left\{ z^T x - \sum_{i=1}^n \left(g(x_i)\right) + b^T x \right\}$$

$$= \sup_x \left\{ (z^T + b^T)x - \sum_{i=1}^n \left(g(x_i)\right) \right\}$$

$$= \sup_{x_1, x_2, \cdots, x_n} \left\{ \sum_{i=1}^n \left((z_i + b_i)x_i - g(x_i)\right) \right\}$$

$$= \sum_{i=1}^n \left(\sup_{x_i} \left\{(z_i + b_i)x_i - g(x_i)\right\} \right)$$

$$= \sum_{i=1}^n g^*(z_i + b_i)$$

4a1, and use results from 4a2 above, we have the dual problem for the generalized MLE to be:

$$\sup_z \left\{ -\sum_{i=1}^n g^*(z_i + b_i) - R^*(-A^T z) \right\} \tag{4a3}$$

6

(b)

Objective: Find the dual of:

$$\min_x \left\{ \sum_{i=1}^n \left(\underbrace{\ln(1 + \exp(a_i^T x))}_{g(a_i^T x)} \right) - b^T Ax + \underbrace{\frac{\lambda}{2}\|x\|^2}_{R(x)} \right\}$$

What is: $R^*(z)$ exactly?

$$R^*(z) = \sup \left\{ z^T x - \frac{\lambda}{2}\|x\|^2 \right\} \tag{4b1}$$

$$= \lambda \sup \left\{ \frac{z^T x}{\lambda} - \frac{1}{2}\|x\|^2 \right\}$$

$$= \frac{\lambda}{2} \left\| \frac{z}{\lambda} \right\|^2$$

The conjugate of the function $\frac{1}{2}\|x\|^2$ is itself, this is established in the lecture.
From 1d6, we can express the conjugate of the all the sum of the conjugate, while also adding a constraint with the dual objective to take care of the case when the conjugate function reaches infinity.

$$-\sum_{i=1}^n (g^*(z_i + b_i)) = \sum_{i=1}^n \left((z_i + b_i) \ln \left(\frac{z_i + b_i}{1 - z_i - b_i} \right) - \ln \left(\frac{1}{1 - z_i - b_i} \right) \right) \quad \forall z_i + b_i \in [0, 1] \tag{4b2}$$

For the case when any of the constraint $z_i + b_i \in [0, 1]$, the sum of conjugate will be $-\infty$, we can add this as a constraint for the dual maximization problem, stating them together we have:

$$\sup_z \left\{ -\sum_{i=1}^n (g^*(z_i + b_i)) - \frac{\lambda}{2} \left\| \frac{z}{\lambda} \right\|^2 \quad \text{s.t} \quad z + b \in [0, 1]^n \right\} \tag{4b3}$$

Along with 4b2, 4b3, we have the whole expression for the dual.

(c)

Objective: Find Dual for the dual of the 1-norm regularized the poisson objective:

$$\min_x \left\{ \sum_{i=1}^n \left(\underbrace{\exp(a_i^T x)}_{g(a_i^T x)} \right) - b^T Ax + \underbrace{\lambda\|x\|_1}_{R(x)} \right\}$$

2 parts remaining, we need to find the conjugate of the $g(x)$ and the 1-norm regularization term.

$$(\lambda\|x\|_1)^*(z) = \sup_x \left\{ z^T x - \lambda\|x\|_1 \right\} \tag{4c1}$$

$$= \lambda \sup_x \left\{ \frac{z^T x}{\lambda} - \|x\|_1 \right\}$$

$$= \delta_{\mathbb{B}_\infty} \left(\frac{z}{\lambda} \right)$$

$$(\exp(x))^*(z) = \sup_x \left\{ zx - \exp(x) \right\} \tag{4c2}$$

Consider:

$$\partial_x[zx - \exp(x)] = 0$$

$$z - \exp(x) = 0$$

$$x = \ln(z) \quad \text{Assuming: } z > 0$$

$$\implies (\exp(x))^*(z) = z\ln(z) - z$$

Notice that, when $z \leq 0$, it's either that $\exp(x)$ is a constant, or $\exp(x) \to 0$ as $x \to -\infty$. So taking $x \to -\infty$, we will make expression inside sup to be ∞ because of the zx in it. Therefore:

$$(\exp(x))^*(z) = \begin{cases} z\ln(z) - z & z \geq 0 \\ \infty & z < 0 \end{cases} \tag{4c3}$$

Using results from 4a3, and 4c3 and 4c1, we have:

$$\sup_{z \in \mathbb{R}^n_+} \left\{ -\sum_{i=1}^n \left((z_i + b_i)\ln(z_i + b_i) - (z_i + b_i)\right) \quad \text{s.t:} \quad \left\|\frac{-A^T z}{\lambda}\right\|_\infty < 1 \right\} \tag{4c4}$$

There are 2 constraints for the dual variable z, one is under the sup and other one is in the braces.

Problem 5

(a), (b)

The k-simplex and box is given as:

$$\Delta_k := \{x : 1^T x = k, 0 \leq x_i \leq 1 \; \forall i\}$$

Objective: Find:

$$\text{proj}(z) = \underset{x \in \Delta_k}{\text{argmin}} \frac{1}{2}\|x - z\|^2$$

Strategies: Start with the problem, find the Lagrangian, measure the optimality conditions of the dual, phrase a root finding problem from the optimality conditions, and then connect the primal to the dual variable to get the optimal to the primal by checking the derivative for the primal problem

By focusing on the equality constraint, we express the box constraint implicitly, giving us the primal objective function in the form of:

$$\min_{x \in [0,1]^n} \left\{ \frac{1}{2}\|x - z\|^2 + \delta_0(1^T x - k) \right\} \tag{5a1}$$

Where the equality constraints are expressed in the form of the indictor function, and it has the dual expression:

$$\delta_0(1^T x - k) = \sup_z \left\{ (1^T x - k)y - \delta_0^*(y) \right\} \tag{5a2}$$

$$= \sup_y \left\{ (1^T x - k)y \right\}$$

Continuing on 5a1, we will have:

$$\min_{x \in [0,1]^n} \left\{ \frac{1}{2}\|x - z\|^2 + \sup_y \left\{ (1^T x - k)y \right\} \right\} \tag{5a3}$$

$$\min_{x \in [0,1]^n} \sup_y \left\{ \frac{1}{2}\|x - z\|^2 + (1^T x - k)y \right\}$$

$$\min_x \sup_y \left\{ \underset{x \in [0,1]^n}{\text{proj}} \left(\frac{1}{2}\|x - z\|^2 + (1^T x)y \right) - ky \right\}$$

$$\min_x \sup_y \left\{ \sum_{i=1}^n \left(\min_{x_i \in [0,1]} \left(\frac{1}{2}(x_i - z_i)^2 + x_i y \right) \right) - ky \right\}$$

$$\implies \mathcal{L}(x, y) = \sum_{i=1}^n \left(\min_{x_i \in [0,1]} \left(\frac{1}{2}(x_i - z_i)^2 + x_i y \right) \right) - ky$$

8

Where the primal is min and dual is max, with x as a vector with length n for the primal problem and y as a single scalar for the dual problem. The expression from 5a3 can be simplified because there is a closed form for the inner minimization:

$$x_i^+ = \operatorname*{argmin}_{x_i \in [0,1]} \left\{ \frac{1}{2}(x_i - z_i)^2 + x_i y \right\} \tag{5a4}$$

$$x_i^+ = \max(0, \min(1, z_i - y))$$

$$x^+ = \max(0, \min(1, z - y)) \leftarrow \text{Vector Form}$$

Which is simply, if we have $x_i \in [0,1]$ and it set the derivative of the objective inside to be zero, then that is the point, else, we choose the $x_i \in [0,1]$ that is as close to the optimal point as possible. This will give us the new Lagrangian in the form of:

$$\mathcal{L}(x,y) = \sum_{i=1}^n \left(\frac{1}{2}(x_i^+ - z_i)^2 + x_i^+ y \right) - ky \tag{5a5}$$

Which, take note that, this is purely a function wrt to y, so the dual as a max problem, and this problem looks piece-wise quadratic, which means that we can use root finding on the derivative to solve for the optimal dual solution. Firstly, we would love to know the derivative of the function $x_i^+(y)$ wrt y first, which can be think of like this:

$$\partial_y x_i^+(y) = \begin{cases} z_i - y & z_i - y \in [0,1] \\ 0 & z_i - y < 0 \\ 1 & z_i - y > 1 \end{cases} \tag{5a6}$$

And next, we are going to take the derivative expression inside the summation wrt to y, giving us:

$$\partial_y \left(\frac{1}{2}(x_i^+ - z_i)^2 + x_i^+(y)y \right) \tag{5a7}$$

$$= (x_i^+ - z_i)\partial_y[x_i^+] + x_i^+ + \partial_y[x_i^+]y$$

$$= \partial_y[x_i^+](x_i^+ - z_i + y) + x_i^+$$

Observe that:

$$(x_i^+ - z_i + y) = \begin{cases} z_i - y - z_i + y = 0 & z_i - y \in [0,1] \\ y - z_i & z_i - y < 0 \\ 1 - z_i + y & z_i - y > 0 \end{cases} \tag{5a8}$$

use 5a7, multiply cases by cases, get:

$$\partial_y[x_i^+](x_i^+ - z_i + y) = 0$$

Fascinating, this means that we have:

$$\partial_y \mathcal{L}(x,y) = \sum_{i=1}^n \left(\partial_y \left(\frac{1}{2}(x_i^+(y) - z_i)^2 + x_i^+(y)y \right) \right) - k \tag{5a9}$$

$$= \sum_{i=1}^n \left(x_i^+(y) \right) - k$$

$$\implies 0 = \sum_{i=1}^n \left(\max(0, \min(1, z_i - y^+)) \right) - k$$

Solving for y^+ using bisection method will give us the solution for the dual problem in this case, and the last thing to check is that, the optimal variable for the dual y^+ can give us the solution for primal solution. Notice that from 5a5 the Lagrangian of the function is not related to the primal variable anymore, that means, if we have the solution for the dual variable y, then the solution for primal is just x^+. Yeah, because $\partial_x \mathcal{L}(x,y) = 0 \implies x = x^+$.

(b)

It's actually back in part (a).

(c)

See the last paragraph in part problem 5 (a).