Name: Philip Pham

AMATH 515 Homework Set 2

(1) Recall that

$$\operatorname{prox}_{tf}(y) = \arg\min_x \frac{1}{2t}\|x-y\|^2 + f(x)$$
$$f_t(y) = \min_x \frac{1}{2t}\|x-y\|^2 + f(x).$$

Suppose f is convex.

(a) Prove that f_t is convex.

Proof. Let $h(x,y) = \frac{1}{2t}\|x-y\|^2 + f(x)$, so $f_t(y) = \min_x h(x,y)$. h is a convex function of x as it is the sum of a ℓ_2-norm and a convex function. Only the first term depends on y, which is a ℓ_2-norm, which is convex, so h is convex as a function of y.
Now, using the convexity, we have that

$$\begin{aligned} f_t\left(\lambda y_1 + (1-\lambda)y_2\right) &= \min_x h\left(x, \lambda y_1 + (1-\lambda)y_2\right) \\ &\leq h\left(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2\right) \\ &\leq \lambda h\left(x_1, y_1\right) + (1-\lambda)h\left(x_2, y_2\right). \end{aligned}$$

We can choose x_1 and x_2 to be anything. For instance, we could choose $x_1 = \arg\min_x h\left(x, y_1\right)$ and $x_2 = \arg\min_x h\left(x, y_2\right)$. In this case, we'd have that

$$f_t\left(\lambda y_1 + (1-\lambda)y_2\right) \leq \lambda h\left(x_1, y_1\right) + (1-\lambda)h\left(x_2, y_2\right) = \lambda f_t\left(y_1\right) + (1-\lambda) f_t\left(y_2\right),$$

so f_t is convex. □

(b) Prove that prox_{tf} is a single-valued mapping.

Proof. In h, the first term is strongly convex with $\alpha = t^{-1} \gneq 0$. Since h is the sum of convex functions, h will also be strongly convex with $\alpha \gneq 0$. Therefore, h has a unique global minimizer, so prox_{tf} is a single-valued mapping. □

(c) Compute prox_{tf} and f_t, where $f(x) = \|x\|_1$.

Proof. We can rewrite the objective as a sum of positive terms

$$\frac{1}{2t}\|x-y\|^2 + \|x\|_1 = \sum_{i=1}^n \left[\frac{1}{2t}\left(x_i - y_i\right)^2 + \left|x_i\right|\right].$$

Each term is independent of each other, so we can minimize each term separately. The derivate of each term is undefined at $x_i = 0$, but otherwise,

$$\frac{\partial}{\partial x_i}\left(\frac{1}{2t}(x_i - y_i)^2 + |x_i|\right) = \frac{x_i - y_i}{t} + \operatorname{sign}(x_i).$$

Solving for x_i when $|y_i| \geq t$, we find $x_i = y_i - \operatorname{sign}(y_i)\, t$. Otherwise, we note that the derivative is negative when $x_i < 0$ and positive when $x_i > 0$, so the solution must be $x_i = 0$. Thus, we have that

$$\left[\operatorname{prox}_{tf}(y)\right]_i = \begin{cases} y_i - \operatorname{sign}(y_i)\, t, & |y_i| \geq t; \\ 0, & \text{otherwise.} \end{cases}$$

$\square$

(d) Compute prox_{tf} and f_t for $f = \delta_{\mathbb{B}_\infty}(x)$, where $\mathbb{B}_\infty = [-1, 1]^n$.

(2) More prox identities.

(a) Suppose f is convex and let $g(x) = f(x) + \frac{1}{2}\|x - x_0\|^2$. Find formulas for prox_{tg} and g_t in terms of prox_{tf} and f_t.

(b) The elastic net penalty is used to detect groups of correlated predictors:

$$g(x) = \beta\|x\|_1 + (1-\beta)\frac{1}{2}\|x\|^2, \quad \beta \in (0, 1).$$

Write down the formula for prox_{tg} and g_t.

(c) Let $f(x) = \frac{1}{2}\|Cx\|^2$. Write $\operatorname{prox}_{tf}(y)$ in closed form.

(d) Let $f(x) = \|x\|_2$. Write $\operatorname{prox}_{tf}(y)$ in closed form.

Coding Assignment