

Name:

AMATH 515

Homework Set 2, Due Feb 8, 11:59 pm.

- (1) Let $x, y \in \mathbb{R}^n$, and consider a function $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$. We make the following definitions:

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

$$f_t(y) := \min_x \frac{1}{2t} \|x - y\|^2 + f(x).$$

Notice that $\text{prox}_{tf}(y)$ is the minimizer of an optimization problem; in particular it is a vector in \mathbb{R}^n . On the other hand $f_t(y)$ is a function from $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, just as f .

Suppose f is convex.

- (a) Show that f_t is convex.
- (b) Show that $\text{prox}_{tf}(y)$ is uniquely defined for any input y .
- (c) Compute prox_{tf} and f_t , where $f(x) = \|x\|_1$.
- (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 $g_s(x) = f(x) + \frac{1}{2s} \|x - x_0\|^2$. Find the formula for prox_{tg} in terms of prox_{tf} .
- (b) Let $f(x) = \|x\|_2$. Write $\text{prox}_{tf}(y)$ in closed form.
- (c) Let $f(x) = \frac{1}{2} \|x\|_2^2$. Write $\text{prox}_{tf}(y)$ in closed form.
- (d) Let $f(x) = \frac{1}{2} \|Cx\|^2$. Write $\text{prox}_{tf}(y)$ in closed form.

Coding Assignment

Please download `515Hw2.Coding.ipynb`, `solvers.py` and `mnist01.npy` to complete the coding problem (3), (4) and (5).

- (3) Complete three generic solvers we learned from the class in `solvers.py`, including,
- proximal gradient descent,
 - accelerated gradient descent.
 - accelerated proximal gradient descent.

- (4) Compressive sensing, consider the sparse regression problem,

$$\min_x \frac{1}{2} \|Ax - b\|^2 + \lambda \|x\|_1$$

where $A \in \mathbb{R}^{m \times n}$ and $m < n$. When x is sparse, it is possible to recover using the ℓ_1 regularizer. We choose $\lambda = \|A^\top b\|_\infty / 10$.

- (a) By treating $f(x) = \frac{1}{2} \|Ax - b\|^2$ and $g(x) = \lambda \|x\|_1$, complete the function w.r.t. to f and g .
- (b) Apply the proximal gradient algorithm. Do you recover the signal?
- (c) Apply accelerated proximal gradient, is it faster than method of (b)?

- (5) Logistic regression on MNIST data, recall the logistic regression problem,

$$\min_x \sum_{i=1}^m \{ \ln(1 + \exp(\langle a_i, x \rangle)) - b_i \langle a_i, x \rangle \} + \frac{\lambda}{2} \|x\|^2.$$

We will use logistic regression to classify the “0” and “1” images from MNIST. In this example, a_i is our vectorized image, and b_i is the corresponding label. We want to obtain an classifier, so that for a new image a_{new} , we can predict

$$\begin{cases} a_{\text{new}} \text{ is a 0,} & \text{if } \langle a_{\text{new}}, x \rangle \leq 0 \\ a_{\text{new}} \text{ is a 1,} & \text{if } \langle a_{\text{new}}, x \rangle > 0 \end{cases}.$$

- (a) Complete the function, gradient and Hessian for the logistic regression.
- (b) Apply gradient, accelerate gradient and Newton’s method to solve the problem. Which one is the fastest and which one is the slowest?
- (c) What is your accuracy of the classification for the test data.