MATH 680: Computation Intensive Statistics **Winter 2018**

Lecture 9: February 5

Lecturer: Yi Yang *Scribes: Jiahui Wu*

9.1 Convex solution sets

We can cite LASSO regression as an example:

$$f(\beta) = \|y - X\beta\|_2^2$$
$$f(\beta) = \beta^T X^T X\beta - 2y^T X\beta + C$$
$$\nabla^2 f(\beta) = X^T X \succeq 0$$

Therefore, LASSO problem is not strictly convex, and has infinite solutions. For example, when $p > n$ LASSO regression may not have a unique minimizer.

9.1.1 Huber loss

$$\sum_{i=1}^{n} \rho(y_i - x_i^T\beta), \quad \rho(z) = \begin{cases} z^2/2 & -z > 0 \\ \delta\,|z| = \delta^2/2 & 1 - z \leq 0 \end{cases}$$

When we use huber loss instead of quadratic loss, the effect of outliers will be diminished.

9.2 Hinge form of SVMs

Hinge loss can be written like this:

$$f(z) = (1-z)_+ = \begin{cases} 1 - z & 1 - z > 0 \\ 0 & 1 - z \leq 0 \end{cases}$$

where $z = y_i(X_i^T\beta + \beta_0)$

If we graph this function, it will be similar with the graph of logistical loss.

9.3 Rewriting constraints

We have an optimization problem:

$$\min_x f(x) \qquad \textit{subject to } g_i(x) \leq 0, i = 1, ...m \quad Ax = b$$

There are two methods to rewrite it:

1. $\min_x f(x) \qquad \textit{subject to } x \in C$ where $C = \{x : g_i(x) \le 0, i = 1, ...m, Ax = b\}$;
2. $\min_x f(x) + I_C(x) \quad \textit{where } I_C = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$

The first method can be used in all problems. However, the second method can be used only for convex problems.

9.4 First-order optimality condition

Sufficient and necessary condition of the statement ”Differentiable function f is convex” are:

1. dom(f) is convex;
2. $f(y) \ge f(x) + \bigtriangledown f(x)(y - x)$

First-order optimality condition: Sufficient and necessary condition of the statement ”Feasible point x is optimal” is:

$$\bigtriangledown f(x)(y - x) \ge 0 \quad \textit{for all } y \in C$$

9.4.1 Quadratic minimization

Next, we use quadratic minimization as an example:

$$f(x) = \frac{1}{2}x^TQx + b^Tx + c \qquad \textit{where } Q \succeq 0$$

First ordercondition:

$$\bigtriangledown f(x) = Qx + b = 0$$

if Q is singular and $b \in col(Q)$, we have

$$Qx = -b = -QQ^+b + QZ = Q(-Q^+b + z) \qquad \textit{where } Q^+Q = I \textit{ and } Qz = 0$$

$$x = -Q^+b + z \qquad \textit{where } z \in ker(Q)$$