Convex Optimization Problems

October 8, 2024

1 Convex optimization problems

Definition 1. Optimization problem

$$\min_{x \in D} f(x)$$
$$\text{subject to } g_i(x) \leq 0, \quad i = 1, ..., m$$
$$h_j(x) = 0, \quad j = 1, ..., p$$

Here $D = \text{dom}(f) \cap \bigcap_{i=1}^{m} \text{dom}(g_i) \cap \bigcap_{i=1}^{p} \text{dom}(h_j)$, common domain of all the functions.

Definition 2. Convex optimization problem: optimization problem set-up above provided that the functions f and g_i, $i = 1, ..., m$ are convex, and h_j, $j = 1, ..., p$ are affine:

$$h_j(x) = a_j^T x + b_j, \quad j = 1, ..., p$$

- Affine function

$$h_j(x) = 0 \Leftrightarrow h_j(x) \leq 0 \; h_j(x) \geq 0$$

Comments: Note we can represent the constraints as follow:

1. $g(x) \geq 0$ and $-g(x) \leq 0$.
2. $h(x) \leq 0$ and $h(x) \geq 0 \iff h(x) = 0$.
3. Domain of convex optimization problem is always convex (intersection of convex sets is also convex set).
4. $\min_x f(x) \iff \max_x -f(x)$

Motivation for convex problems: local minima = global minima!

Proof. Use contradiction. If x is not a global minima, then there must exist some feasible $z \in D$ such that

$$f(z) < f(x)$$

then

$$||z - x||_2 > \rho$$

Now we choose

$$y = tx + (1 - t)z$$

for some $0 \leq t \leq 1$, then □

- $y \in D$
- y satisfies the constraints

$$\begin{aligned} h_j(y) &= a_j^T(tx + (1 - t)z) + b_j \\ &= 0 \end{aligned}$$

$$
\begin{aligned}
g_i(y) &\leq t g_i(x) + (1-t) g_i(z) \\
&\leq 0
\end{aligned}
$$

- Now take a very large value of t such that $||y - x||_2 \leq \rho$. By the convexity of f, we have

$$
\begin{aligned}
f(y) &= f(tx + (1-t)z) \\
&\leq t f(x) + (1-t) f(z) \\
&< t f(x) + (1-t) f(x) \\
&= f(x).
\end{aligned}
$$

  Therefore we have found y in the neighborhood of x and $y < x$. This contradicts with the fact that x is the local minimum.

1.1 Convex solution sets

We can cite LASSO regression as an example:

$$f(\beta) = \left\| y - X\beta \right\|_2^2$$

$$f(\beta) = \beta^T X^T X \beta - 2y^T X \beta + C$$

$$\nabla^2 f(\beta) = X^T X \succeq 0$$

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

1.2 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.

1.3 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.

1.4 Rewriting constraints

We have an optimization problem:

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

There are two methods to rewrite it:

1. $\min_x f(x)$ subject to $x \in C$ where $C = \{x : g_i(x) \leq 0, i = 1, ...m, Ax = b\}$;
2. $\min_x f(x) + I_C(x)$ 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.

1.5 First-order optimality condition

Sufficient and necessary condition of the statement "differentiable function f is convex" are:

1. $\mathrm{dom}(f)$ is convex;
2. $f(y) \geq f(x) + \nabla f(x)(y - x)$

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

$$\nabla f(x)(y - x) \geq 0 \quad \text{for all } y \in C$$

1.6 Quadratic minimization

Next, we use quadratic minimization as an example:

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

First order condition:

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

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

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

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