MSBD5007 Optimization and Matrix Computation Homework 5

RONG Shuo

May 4, 2025

1. Find explicit formulas for the projection of $y \in \mathbb{R}^n$ onto the following non-empty, closed, and convex sets $S \subset \mathbb{R}^n$, respectively.

   (a) The unit ∞-norm ball
   $$S = \{x \in \mathbb{R}^n \mid \|x\|_\infty \leq 1\}.$$

   (b) The closed halfspace
   $$S = \{x \in \mathbb{R}^n \mid a^T x \leq b\},$$
   where $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$ are given.

2. Consider the optimization problem in $x = (x_1, x_2, x_3) \in \mathbb{R}^3$:
$$\begin{aligned} \min_{x\in\mathbb{R}^3} & \frac{1}{2}(x_1^2 + 4x_2^2 + 9x_3^2) - (4x_1 + 2x_2), \\ \text{s.t.} \quad & x_1 + x_2 + x_3 = 3, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{aligned}$$

   (a) Write down the KKT conditions (stationarity, feasibility, complementary slackness).

   (b) Solve the KKT system to find the optimal solution x^*, the Lagrange multiplier λ^* for the equality constraint, and the multipliers $\mu^* = (\mu_1, \mu_2, \mu_3)$ for the inequality constraints.

3. We wish to compute the projection of $y \in \mathbb{R}^n$ onto the unit $\ell_1$ball, i.e. solve
$$\begin{aligned} \min_{x\in\mathbb{R}^n} & \frac{1}{2}\|x - y\|_2^2 \\ \text{s.t.} \quad & \|x\|_1 \leq 1. \end{aligned}$$

   (a) Derive the Lagrange dual problem and show it can be written as
$$\max_{\lambda\geq 0} d(\lambda), \quad d(\lambda) = \sum_{i=1}^n h_\lambda(y_i) - \lambda,$$

where $h_\lambda : \mathbb{R} \to \mathbb{R}$ is the so-called Huber's function (which is a smooth function consisting of a quadratic and two linear pieces) defined by

$$h_\lambda(t) = \begin{cases} \frac{1}{2}t^2, & |t| \leq \lambda, \\ \lambda|t| - \frac{1}{2}\lambda^2, & |t| \geq \lambda. \end{cases}$$

(b) Prove that strong duality holds.

(c) Find the optimal dual multiplier λ^*.

(d) Give an expression of the projection in terms of λ^*.

4. Let $S \subseteq \mathbb{R}^n$ be a nonempty, closed, convex set, and let $\|\cdot\|_2$ denote 2-norm. The projection of any point $y \in \mathbb{R}^n$ onto S is defined by

$$\mathcal{P}_S(y) = \arg\min_{x \in S} \|x - y\|_2.$$

Prove that the projection $\mathcal{P}_S : \mathbb{R}^n \to S$ is nonexpansive: for all $x, y \in \mathbb{R}^n$,

$$\|\mathcal{P}_S(x) - \mathcal{P}_S(y)\|_2 \leq \|x - y\|_2.$$

1 Answer

1.1 (1)

1.1.1 (a)

We know the projection:

$$\mathcal{P}_S(y) = \arg\min_{\boldsymbol{x} \in S} \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

We consider

$$\min_{\boldsymbol{x} \in \mathbb{S}} \|\boldsymbol{x} - \boldsymbol{y}\|_2$$

This means:

$$\min_{\boldsymbol{x} \in \mathbb{S}} \sum_{i=1}^{n} (x_i - y_i)^2$$

We can solve for each component x_i independently:

$$\min_{x_i \in [-1,1]} (x_i - y_i)^2$$

This is just the Euclidean projection of a scalar onto the interval $[-1, 1]$, which gives:

$$x_i = \min(\max(y_i, -1), 1)$$

and

$$\boldsymbol{x} = [x_i]_1^n$$

1.1.2 (b)

We know the projection:

$$\mathcal{P}_S(y) = \arg\min_{\boldsymbol{x}\in S} \|\boldsymbol{x}-\boldsymbol{y}\|_2$$

$$\mathcal{L}(\boldsymbol{x},\lambda) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 + \lambda(\boldsymbol{a}^T\boldsymbol{x}-b), \quad \lambda \geq 0$$

We know the KKT conditions:

$$\begin{aligned} \nabla_x\mathcal{L} &= \boldsymbol{x}-\boldsymbol{y}+\lambda\boldsymbol{a} = 0 \\ \boldsymbol{x}^* &= \boldsymbol{y}-\lambda\boldsymbol{a} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{a}^T\boldsymbol{x}^* &\leq b \\ \boldsymbol{a}^T(\boldsymbol{y}-\lambda\boldsymbol{a}) &\leq b \\ \boldsymbol{a}^T\boldsymbol{y}-\lambda\|\boldsymbol{a}\|^2 &\leq b \\ \lambda &\geq \frac{\boldsymbol{a}^T\boldsymbol{y}-b}{\|\boldsymbol{a}\|^2} \end{aligned}$$

and we know:

$$\begin{aligned} \lambda(\boldsymbol{a}^T\boldsymbol{x}-b) &= 0 \\ \lambda(\boldsymbol{a}^T\boldsymbol{y}-\lambda\|\boldsymbol{a}\|^2-b) &= 0 \end{aligned}$$

So we get $\lambda = \max(0, \frac{\boldsymbol{a}^T\boldsymbol{y}-b}{\|\boldsymbol{a}\|^2})$
So we can get

$$\mathcal{P}_S(\boldsymbol{y}) = \boldsymbol{y} - \max(0, \frac{\boldsymbol{a}^T\boldsymbol{y}-b}{\|\boldsymbol{a}\|^2})\boldsymbol{a}$$

1.2 (2)

1.2.1 (a)

We can write the Lagrangian Function:

$$\mathcal{L}(x,\lambda,\mu) = \frac{1}{2}(x_1^2+4x_2^2+9x_3^2) - (4x_1+2x_2) - \lambda(x_1+x_2+x_3-3) - \sum_{i=1}^{3}\mu_i x_i$$

$$
\begin{aligned}
\nabla_x \mathcal{L} &= 0 \\
x_1 - 4 - \lambda - \mu_1 &= 0 \\
4x_2 - 2 - \lambda - \mu_2 &= 0 \\
9x_3 - \lambda - \mu_3 &= 0
\end{aligned}
$$

$$
\begin{aligned}
x_1 + x_2 + x_3 &= 3 \\
x_i &\geq 0
\end{aligned}
$$

$$
\begin{aligned}
\mu_i &\geq 0 \\
\mu_i x_i &= 0
\end{aligned}
$$

1.2.2 (b)

We need to do analysis of the solution first.

$$
\begin{aligned}
x_1 > 0, x_2 > 0, x_3 > 0 &\quad \text{invalid} \\
x_1 = 0, x_2 > 0, x_3 > 0 &\quad \text{invalid} \\
x_1 > 0, x_2 = 0, x_3 > 0 &\quad \text{invalid} \\
x_1 > 0, x_2 > 0, x_3 = 0 &\quad \text{valid}
\end{aligned}
$$

if $x_3 = 0$, we know $\mu_1 = 0, \mu_2 = 0$. Therefore, we get:

$$
\begin{aligned}
x_1 &= \lambda + 4 \\
x_2 &= \frac{\lambda + 2}{4} \\
\mu_3 &= -\lambda
\end{aligned}
$$

We know $x_1 + x_2 = 3$.

$$
\begin{aligned}
\lambda + 4 + \frac{\lambda + 2}{4} &= 3 \\
\lambda &= -\frac{6}{5}
\end{aligned}
$$

s.t.

$$x_1 = \frac{14}{5}$$
$$x_2 = \frac{1}{5}$$
$$x_1 = 0$$
$$\mu_3 = \frac{6}{5}$$

Therefore, we can conclude:

$$x_1 = \frac{14}{5}$$
$$x_2 = \frac{1}{5}$$
$$x_1 = 0$$
$$\mu_1 = 0$$
$$\mu_2 = 0$$
$$\mu_3 = \frac{6}{5}$$
$$\lambda = -\frac{6}{5}$$

1.3 (3)

1.3.1 (a)