CS680, Spring 2020, Assignment 10

Zhijie Wang, [redacted]

July 21, 2020

Exercise 1

$$
\begin{aligned}
&\max_{\forall j, \|\mathbf{z}_j\|_2 \leqslant \lambda} \|(X+Z)\mathbf{w} - \mathbf{y}\|_2 \\
&\max_{\forall j, \|\mathbf{z}_j\|_2 \leqslant \lambda} \|X\mathbf{w} - y + (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_d)\mathbf{w}\|_2 \\
\leqslant &\|X\mathbf{w} - y\|_2 + \max_{\forall j, \|\mathbf{z}_j\|_2 \leqslant \lambda} \|\sum_{i=1}^{d} \mathbf{z}_i w_i\|_2 \\
\leqslant &\|X\mathbf{w} - y\|_2 + \sum_{i=1}^{d} |\mathbf{w}_i|\lambda \\
&\|X\mathbf{w} - y\|_2 + \lambda\|\mathbf{w}\|_1
\end{aligned}
\tag{1}
$$

Therefore, $\min_{\mathbf{w}\in\mathbb{R}^d} \max_{\forall j, \|\mathbf{z}_j\|_2 \leqslant \lambda} \|(X+Z)\mathbf{w} - \mathbf{y}\|_2$ is equivalent to $\min_{\mathbf{w}\in\mathbb{R}^d} \|X\mathbf{w} - y\|_2 + \lambda\|\mathbf{w}\|_1$

Exercise 2

1. Re-write the summation as vector and matrix multiplication.

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathbb{E}\left[\left(y_i - \mathbf{w}^T\tilde{\mathbf{x}}_i\right)^2\right] \\
&\mathbb{E}\left[\left(\mathbf{y} - \tilde{X}\mathbf{w}\right)^2\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T\tilde{X}^T\mathbf{y} + \mathbf{w}^T\tilde{X}^T\tilde{X}\mathbf{w}\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T\left(X+\epsilon\right)^T\mathbf{y} + \mathbf{w}^T\left(X+\epsilon\right)^T\left(X+\epsilon\right)\mathbf{w}\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^TX^T\mathbf{y} - 2\mathbf{w}^T\epsilon^T\mathbf{y} + \mathbf{w}^TX^TX\mathbf{w} + \mathbf{w}^T\epsilon^TX\mathbf{w} + \mathbf{w}^TX^T\epsilon\mathbf{w} + \mathbf{w}^T\epsilon^T\epsilon\mathbf{w}\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^TX^T\mathbf{y} + \mathbf{w}^TX^TX\mathbf{w}\right] + \mathbb{E}\left[-2\mathbf{w}^T\epsilon^T\mathbf{y} + \mathbf{w}^T\epsilon^TX\mathbf{w} + \mathbf{w}^TX^T\epsilon\mathbf{w}\right] + \mathbb{E}\left[\mathbf{w}^T\epsilon^T\epsilon\mathbf{w}\right] \\
&\left(\mathbf{y} - X\mathbf{w}\right)^2 + 0 + \mathbf{w}^T\mathbb{E}\left[\epsilon^T\epsilon\right]\mathbf{w} \\
&\sum_{i=1}^{n}\left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \mathbf{w}^T\sum_{i=1}^{n}\mathbb{E}\left[\epsilon_i^2\right]\mathbf{w} \\
&\sum_{i=1}^{n}\left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \mathbf{w}^T\sum_{i=1}^{n}\left\{\mathbb{E}\left[\epsilon_i\right]^2 + Var\left[\epsilon_i\right]\right\}\mathbf{w} \\
&\sum_{i=1}^{n}\left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \mathbf{w}^T\sum_{i=1}^{n}\left\{0_{d\times d} + \lambda\mathbb{I}_{d\times d}\right\}\mathbf{w} \\
&\sum_{i=1}^{n}\left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + n\lambda\|\mathbf{w}\|_2^2
\end{aligned}
\tag{2}
$$

Therefore, $\min \sum_{i=1}^{n} \mathbb{E}\left[\left(y_i - \mathbf{w}^T\tilde{\mathbf{x}}_i\right)^2\right]$ is equivalent to $\min \sum_{i=1}^{n}\left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \lambda\|\mathbf{w}\|_2^2$

2. Re-write the summation like Ex2.1.

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathbb{E}\left[\left(y_i - \mathbf{w}^T \tilde{\mathbf{x}}_i\right)^2\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T\tilde{X}^T\mathbf{y} + \mathbf{w}^T\tilde{X}^T\tilde{X}\mathbf{w}\right] \\
&\mathbb{E}\left[\mathbf{y}^T\mathbf{y}\right] - 2\mathbf{w}^T\mathbb{E}\left[\tilde{X}\right]^T\mathbf{y} + \mathbf{w}^T\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]\mathbf{w} \\
&\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T\mathbb{E}\left[\tilde{X}\right]^T\mathbf{y} + \mathbf{w}^T\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]\mathbf{w}
\end{aligned}
\tag{3}
$$

For $\mathbb{E}\left[\tilde{X}\right]$,

$$
\begin{aligned}
\mathbb{E}\left[\tilde{X}\right]_{ji} &\quad p \cdot \frac{x_{ji}}{p} \quad x_{ji} \\
\mathbb{E}\left[\tilde{X}\right] &\quad X
\end{aligned}
\tag{4}
$$

For $\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]$

$$
\begin{aligned}
\left(\tilde{X}^T\tilde{X}\right)_{ji} &\quad \sum_{k=1}^{n} \tilde{x}_{kj}\tilde{x}_{ki} \\
\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]_{ji} &\quad \begin{cases} \sum p \cdot \frac{x_{kj}}{p} \cdot \frac{x_{ki}}{p} & \frac{1}{p}(X^TX)_{ji} \qquad j \quad i \\ \sum p \cdot p \cdot \frac{x_{kj}}{p} \cdot \frac{x_{ki}}{p} & (X^TX)_{ji} \qquad j \neq i \end{cases} \\
\mathbb{E}\left[\tilde{X}^T\tilde{X}\right] &\quad X^TX + \left(\frac{1}{p} - 1\right)\operatorname{diag}(X^TX)
\end{aligned}
\tag{5}
$$

Back to Eq.(3)

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathbb{E}\left[\left(y_i - \mathbf{w}^T \tilde{\mathbf{x}}_i\right)^2\right] \\
&\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T\mathbb{E}\left[\tilde{X}\right]^T\mathbf{y} + \mathbf{w}^T\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]\mathbf{w} \\
&\mathbf{y}^T\mathbf{y} - 2\mathbf{w}^TX^T\mathbf{y} + \mathbf{w}^TX^TX\mathbf{w} - \mathbf{w}^TX^TX\mathbf{w} + \mathbf{w}^T\mathbb{E}\left[\tilde{X}^T\tilde{X}\right]\mathbf{w} \\
&\sum_{i=1}^{n} \left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \left(\frac{1}{p} - 1\right)\mathbf{w}^T\operatorname{diag}\left(X^TX\right)\mathbf{w} \\
&\sum_{i=1}^{n} \left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \left(\frac{1}{p} - 1\right)\left\|\operatorname{diag}\left(X^TX\right)^{\frac{1}{2}}\mathbf{w}\right\|_2^2
\end{aligned}
\tag{6}
$$

Therefore, min $\sum_{i=1}^{n} \mathbb{E}\left[\left(y_i - \mathbf{w}^T \tilde{\mathbf{x}}_i\right)^2\right]$ is equivalent to min $\sum_{i=1}^{n} \left(y_i - \mathbf{w}^T\mathbf{x}_i\right)^2 + \left(\frac{1}{p} - 1\right)\left\|\operatorname{diag}\left(X^TX\right)^{\frac{1}{2}}\mathbf{w}\right\|_2^2$