BIOS:7600 Homework 7

Chuan Lu

April 24, 2019

1. Problem 5.1, Gaussian tail bound.

   Proof. First, for any $t > 0$,

$$\begin{aligned}\mathbb{P}(Z \geq \lambda) \leq \exp(-t\lambda)\mathbb{E}(e^{tZ}) &= \frac{e^{-t\lambda}}{\sqrt{2\pi(\sigma^2/n)}}\int_{-\infty}^{\infty} e^{tz}e^{-\frac{z^2}{2\sigma^2/n}}dz \\ &= \frac{e^{-t\lambda}}{\sqrt{2\pi(\sigma^2/n)}}\int_{-\infty}^{\infty} \exp(-\frac{(z-\frac{\sigma^2}{n}t)^2}{2\sigma^2/n})\exp(\frac{\sigma^2t^2}{2n})dz \\ &= \exp(-t\lambda + \frac{\sigma^2t^2}{2n}).\end{aligned} \tag{1}$$

   Let

$$t = \frac{n\lambda}{\sigma} > 0, \tag{2}$$

   then the inequality becomes

$$\mathbb{P}(Z \geq \lambda) \leq \exp(-\frac{n\lambda^2}{2\sigma^2}). \tag{3}$$

   By symmetry,

$$\mathbb{P}(Z \leq -\lambda) \leq \exp(-\frac{n\lambda^2}{2\sigma^2}). \tag{4}$$

   Hence we get the result. □

2. Problem 5.3, Prediction bound under RE condition.

   (a) *Proof.* First, as proved in class, we have

$$\frac{1}{n}\|X\delta\|_2^2 \leq 3\lambda\sqrt{|\mathcal{S}|}\|\delta_S\|_2. \tag{5}$$

   By the restricted eigenvalue condition, we have

$$\frac{1}{n}\delta^\top X^\top X\delta \geq \tau\|\delta\|_2^2 \geq \tau\|\delta_S\|_2^2, \tag{6}$$

   so

$$\frac{1}{n}\|X\delta\|_2^2 \leq 3\lambda\sqrt{|\mathcal{S}|}\frac{1}{\sqrt{n\tau}}\|X\delta\|_2. \tag{7}$$

   Hence,

$$\frac{1}{n}\|X\delta\|_2^2 \leq \frac{9}{\tau}\lambda^2|\mathcal{S}|. \tag{8}$$

   □

   (b) *Proof.* Actually, by using the same equality in the last problem, we can get this result quickly. □

3. Problem 6.1, Conditional distribution for random knockoffs.

   Proof. By the conditional distributions of multivariate normal distribution,

$$(\tilde{x_i} \mid x_i = x) \sim N(((\Sigma - S)\Sigma^{-1}x)_i, (\Sigma - (\Sigma - S)\Sigma^{-1}(\Sigma - S))_{ii}) \sim N((x - S\sigma^{-1}x)_i, (2S - S\Sigma^{-1}S)_{ii}). \tag{9}$$

   □

4. Problem 6.2, Selective inference in the p=2 case.

Proof. The condition becomes

$$|\frac{1}{n}x_2^\top(1-P_1)y + \lambda x_2^\top x_1(x_1^\top x_1)^{-1}| \le \lambda. \tag{10}$$

By simplification,

$$-\lambda n(x_1^\top x_1 - x_2^\top x_1) \le (x_1^\top x_1 x_2^\top - x_2^\top x_1 x_1^\top)y \le \lambda n(x_1^\top x_1 - x_2^\top x_1). \tag{11}$$

Then

$$A = \begin{bmatrix} x_1^\top x_1 x_2^\top - x_2^\top x_1 x_1^\top & 0 \\ 0 & x_2^\top x_1 x_1^\top - x_1^\top x_1 x_2^\top, \end{bmatrix}, \quad b = \lambda n \begin{bmatrix} x_1^\top x_1 - x_2^\top x_1 \\ x_1^\top x_1 - x_2^\top x_1 \end{bmatrix}, \tag{12}$$

where $X = [x_1, x_2]$.

$\square$

5. Problem 6.3, HIV drug resistance study.

   (a)