STAT 741 Homework Assignment 2 Due Monday, February 27

1. Exercise 11.3 from Rencher.

Solution: The theorems from Rencher are as follows:

Theorem 2.7A. The system of equations $\mathbf{Ax} = \mathbf{c}$ has at least one solution vector if and only if rank($\mathbf{A}$,$\mathbf{c}$) = rank($\mathbf{A}$).

Theorem 11.2A. If $\mathbf{X}$ is $n \times p$ of rank $k < p \leq n$, the system of equations $\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{Y}$ is consistent.

We know that rank($\mathbf{X}$) = rank($\mathbf{X}^T\mathbf{X}$) = k. Therefore the matrix $[\mathbf{X}^T\mathbf{X}, \mathbf{Y}^T\mathbf{X}]$ has rank at least k. But $[\mathbf{X}^T\mathbf{X}, \mathbf{X}^T\mathbf{Y}] = \mathbf{X}^T[\mathbf{X}, \mathbf{Y}]$ has rank at most k. Therefore the rank of $[\mathbf{X}^T\mathbf{X}, \mathbf{Y}^T\mathbf{X}]$ is exactly k. Applying Theorem 2.7A, we conclude that the normal equations are consistent.

2. (a) Prove that if S_1 and S_2 are statements (for example, statements about parameters being covered by confidence sets) such S_i is true with probability $1 - \alpha_i$, then the probability that S_1 and S_2 are both true is at least $1 - \alpha_1 - \alpha_2$.

(b) If $\mathcal{E}_i$ is the event that S_i is true, sharpen the inequality of (a) if $\mathcal{E}_1$ and $\mathcal{E}_2$ are independent.

Solution: (a) Using Bonferroni's inequality,

$$P[\mathcal{E}_1^c \cup \mathcal{E}_2^c] \leq P[\mathcal{E}_1^c] + P[\mathcal{E}_2^c] = \alpha_1 + \alpha_2.$$

Therefore $P[\mathcal{E}_1 \cap \mathcal{E}_2] \geq 1 - \alpha_1 - \alpha_2$.

(b) Using independence,

$$P[\mathcal{E}_1 \cap \mathcal{E}_2] = (1 - \alpha_1)(1 - \alpha_2) = 1 - \alpha_1 - \alpha_2 + \alpha_1\alpha_2.$$

3. Let $\psi_1, \ldots, \psi_q$ be linearly independent estimable functions and consider the problem of obtaining a confidence set for the parameter vector $\boldsymbol{\theta} = (\psi_1, \ldots, \psi_q, \sigma)$.

(a) Using the result of Problem 2(a), find a confidence set for $\boldsymbol{\theta}$ in the form of an elliptical cylinder with axis $\psi_1 = \hat{\psi}_1, \ldots, \psi_q = \hat{\psi}_q$.

(b) Using the result of Problem 2(b), find a confidence set for $\boldsymbol{\theta}$ in the form of a right frustum of an elliptical cone with the same axis. (A right frustum of a cone is the region bounded by the cone and two planes orthogonal to the axis of the cone.) Draw pictures for the case $q = 2$.

Solution: (a) Let S_1 be the statement that $\boldsymbol{\psi}$ lies in the following $1 - \alpha_1$ confidence ellipsoid:

$$C_1 = \{\psi \,|\, (\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})(\mathbf{A}^T\mathbf{A})^{-1}(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}}) \leq qs^2 F_{q,n-r,\alpha_1}\}$$

where $\hat{\boldsymbol{\psi}} = \mathbf{A}^T\mathbf{Y}$ is the least squares estimator of $\boldsymbol{\psi}$. Let S_2 be the statement that σ lies in the confidence interval:

$$s\sqrt{(n-r)/\chi^2_{n-r,\alpha_2/2}} \leq \sigma \leq s\sqrt{(n-r)/\chi^2_{n-r,1-\alpha_2/2}}.$$

Using Problem 2(a), S_1 and S_2 is a $1 - \alpha_1 - \alpha_2$ confidence statement about $\boldsymbol{\theta}$ defined by an elliptic cylinder whose cross section is congruent to C_1 and whose axis is $\hat{\boldsymbol{\psi}}$.

(b) Recall that $\hat{\boldsymbol{\psi}} \sim N(\boldsymbol{\psi}, \sigma^2\mathbf{A}^T\mathbf{A})$. Therefore the exponent of this q-variate normal density

$$(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})(\mathbf{A}^T\mathbf{A})^{-1}(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})/\sigma^2$$

has a χ^2 distribution with q degrees of freedom, and $\hat{\boldsymbol{\psi}}$ is independent of s^2. Therefore the combined confidence statement is S_1' and S_2, where S_1' is the statement that $\boldsymbol{\psi}$ satisfies the inequality

$$(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})(\mathbf{A}^T\mathbf{A})^{-1}(\boldsymbol{\psi} - \hat{\boldsymbol{\psi}})/\sigma^2 \leq \chi^2_{q,\alpha_1}.$$

This statement corresponds to the frustum of an elliptical cone, as required. According to Problem 2(b), the confidence is $(1 - \alpha_1)(1 - \alpha_2)$

4. Suppose that $\mathbf{x}$ is a point in a space of one or more dimensions and that the "true value" of the variable Y given $\mathbf{x}$ is $g(\mathbf{x})$, where g is an unknown function. Points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ (the *design* of the experiment) are selected and data $(\mathbf{x}_i, Y_i)$, $i = 1, \ldots, n$, are observed. An estimate $\hat{g}$ is formed by some method.

(a) Show that the mean squared error $E\{[\hat{g}(\mathbf{x}) - g(\mathbf{x})]^2\}$ can be written

$$E\{[\hat{g}(\mathbf{x}) - g(\mathbf{x})]^2\} = \mathrm{Var}[\hat{g}(\mathbf{x})] + [E\{\hat{g}(\mathbf{x})\} - g(\mathbf{x})]^2 = V(\mathbf{x}) + B(\mathbf{x}).$$

The quantity $B(\mathbf{x})$ measures bias and is often more important than the sampling variability.

(b) Consider functions of the form $\hat{g}(\mathbf{x})) = \sum_{j=1}^{p} \tilde{\beta}_j h_j(\mathbf{x})$, where the h_j are given functions of $\mathbf{x}$ and the $\tilde{\beta}_j$ are linear functions of the Y_i. Among these functions, the $\hat{g}$ computed by minimizing $\sum_{i=1}^{n} \sum_{j=1}^{p} [Y_i - b_j h_j(\mathbf{x}_i)]^2$ also minimizes the quantity $\sum_{i=1}^{n} B(\mathbf{x}_i)$.

Solution: (a) By adding and subtracting $E\{\hat{g}(\mathbf{x})\}$ we obtain

$$\begin{aligned} E\{[\hat{g}(\mathbf{x}) - g(\mathbf{x})]^2\} &= E\{[\hat{g}(\mathbf{x}) - E\{\hat{g}(\mathbf{x})\} + E\{\hat{g}(\mathbf{x})\} - g(\mathbf{x})]^2\} \\ &= E\{[\hat{g}(\mathbf{x}) - E\{\hat{g}(\mathbf{x})\}]^2\} + E\{[\hat{g}(\mathbf{x})\} - g(\mathbf{x})]^2\} \\ &= \mathrm{Var}[\hat{g}(\mathbf{x})] + [E\{\hat{g}(\mathbf{x})\} - g(\mathbf{x})]^2 \end{aligned}$$

because the cross-product has expectation zero.

(b) Let $\mathbf{X}$ be the matrix with (i, j) entry $h_j(\mathbf{x}_i)$. The minimizer of $\sum_{i=1}^{n} \sum_{j=1}^{p} [Y_i - b_j h_j(\mathbf{x}_i)]^2$ satisfies the normal equations $\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{Y}$, and

$$E\{\hat{\boldsymbol{\beta}}\} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E\{\mathbf{Y}\} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \begin{bmatrix} g(\mathbf{x}_1) \\ \vdots \\ g(\mathbf{x}_n) \end{bmatrix}.$$

Now consider the deterministic least squares problem of minimizing

$$\sum_{i=1}^{n} B(\mathbf{x}_i) = \sum_{i=1}^{n} [\sum_{j=1}^{p} b_j h_j(\mathbf{x}_i) - g(\mathbf{x}_i)]^2 = \sum_{i=1}^{n} [\sum_{j=1}^{p} b_j h_j(\mathbf{x}_i) - E\{Y_i\}]^2.$$

By differentiating we see that the solution $\mathbf{b}^* = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E\{\mathbf{Y}\} = E\{\hat{\boldsymbol{\beta}}\}$.