Math 584 HW 11 Spring 2021, due Thursday, April 22.
Final: Tuesday, May 4, 2:45-4:45.

Refer to Exam 3 review 164)-190). Problems 1) and 2) consider testing 175)-189). More information about multivariate linear regression is in ch.8 of the online notes. Parentheses are the problem from the Olive text.

1) (8.1): Let

$$T(\boldsymbol{W}) = n \ \ [vec(\boldsymbol{L}\hat{\boldsymbol{B}})]^T[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1} \otimes (\boldsymbol{L}\boldsymbol{W}\boldsymbol{L}^T)^{-1}][vec(\boldsymbol{L}\hat{\boldsymbol{B}})].$$

Let

$$\frac{\boldsymbol{X}^T\boldsymbol{X}}{n} = \hat{\boldsymbol{W}}^{-1}.$$

Show $T(\hat{\boldsymbol{W}}) = [vec(\boldsymbol{L}\hat{\boldsymbol{B}})]^T[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1} \otimes (\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}^T)^{-1}][vec(\boldsymbol{L}\hat{\boldsymbol{B}})]$.

2) (8.2): Let $T = [vec(\boldsymbol{L}\hat{\boldsymbol{B}})]^T[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1} \otimes (\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}^T)^{-1}][vec(\boldsymbol{L}\hat{\boldsymbol{B}})]$. Let $\boldsymbol{L} = \boldsymbol{L}_j = [0, ..., 0, 1, 0, ..., 0]$ have a 1 in the jth position. Let $\hat{\boldsymbol{b}}_j^T = \boldsymbol{L}\hat{\boldsymbol{B}}$ be the jth row of $\hat{\boldsymbol{B}}$. Let $d_j = \boldsymbol{L}_j(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}_j^T = (\boldsymbol{X}^T\boldsymbol{X})^{-1}_{jj}$, the jth diagonal entry of $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$. Then $T_j = \frac{1}{d_j}\hat{\boldsymbol{b}}_j^T\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1}\hat{\boldsymbol{b}}_j$. The Hotelling Lawley statistic $U = tr([(n-p)\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}]^{-1}\hat{\boldsymbol{B}}^T\boldsymbol{L}^T[\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}^T]^{-1}\boldsymbol{L}\hat{\boldsymbol{B}}])$. Hence if $\boldsymbol{L} = \boldsymbol{L}_j$, then $U_j = \frac{1}{d_j(n-p)}tr(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1}\hat{\boldsymbol{b}}_j\hat{\boldsymbol{b}}_j^T)$.

Using $tr(\boldsymbol{ABC}) = tr(\boldsymbol{CAB})$ and $tr(a) = a$ for scalar a, show the $(n-p)U_j = T_j$.

3) (8.3): Using the Searle (1982, p. 333) identity
$tr(\boldsymbol{A}\boldsymbol{G}^T\boldsymbol{D}\boldsymbol{G}\boldsymbol{C}) = [vec(\boldsymbol{G})]^T[\boldsymbol{C}\boldsymbol{A} \otimes \boldsymbol{D}^T][vec(\boldsymbol{G})]$, show
$(n-p)U(\boldsymbol{L}) = tr[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1}\hat{\boldsymbol{B}}^T\boldsymbol{L}^T[\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}^T]^{-1}\boldsymbol{L}\hat{\boldsymbol{B}}]$
$= [vec(\boldsymbol{L}\hat{\boldsymbol{B}})]^T[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^{-1} \otimes (\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{L}^T)^{-1}][vec(\boldsymbol{L}\hat{\boldsymbol{B}})]$ by identifying $\boldsymbol{A}, \boldsymbol{G}, \boldsymbol{D}$, and $\boldsymbol{C}$.

Problem 4) is on the back.

4) (8.8): Let $\boldsymbol{y} = \boldsymbol{B}^T\boldsymbol{x} + \boldsymbol{\epsilon}$. Suppose $\boldsymbol{x} = (1, x_2, ..., x_p)^T = (1 \ \boldsymbol{w}^T)^T$ where $\boldsymbol{w} = (x_2, ..., x_p)^T$. Let

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{\alpha}^T \\ \boldsymbol{B}_S \end{pmatrix}.$$

Suppose

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{w} \end{pmatrix} \sim N_{m+p-1}\left[\begin{pmatrix} \boldsymbol{\mu_y} \\ \boldsymbol{\mu_w} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma_{yy}} & \boldsymbol{\Sigma_{yw}} \\ \boldsymbol{\Sigma_{wy}} & \boldsymbol{\Sigma_{ww}} \end{pmatrix}\right].$$

Then $\boldsymbol{y}|\boldsymbol{w} \sim N_m(\boldsymbol{\mu_y} + \boldsymbol{\Sigma_{yw}}\boldsymbol{\Sigma_{ww}^{-1}}(\boldsymbol{w} - \boldsymbol{\mu_w}), \boldsymbol{\Sigma_{yy}} - \boldsymbol{\Sigma_{yw}}\boldsymbol{\Sigma_{ww}^{-1}}\boldsymbol{\Sigma_{ww}})$, and $\boldsymbol{\epsilon} \sim N_m(\boldsymbol{0}, \boldsymbol{\Sigma_{yy}} - \boldsymbol{\Sigma_{yw}}\boldsymbol{\Sigma_{ww}^{-1}}\boldsymbol{\Sigma_{ww}}) = N_m(\boldsymbol{0}, \boldsymbol{\Sigma_\epsilon})$.

Now

$$\boldsymbol{y}|\boldsymbol{x} = \boldsymbol{y}|\begin{pmatrix} 1 \\ \boldsymbol{w} \end{pmatrix} = \boldsymbol{B}^T\boldsymbol{x} + \boldsymbol{\epsilon},$$

and

$$\boldsymbol{y}|\boldsymbol{w} = \boldsymbol{B}^T\boldsymbol{x} + \boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\alpha}^T \\ \boldsymbol{B}_S \end{pmatrix}^T \begin{pmatrix} 1 \\ \boldsymbol{w} \end{pmatrix} + \boldsymbol{\epsilon} = (\boldsymbol{\alpha} \ \ \boldsymbol{B}_S^T)\begin{pmatrix} 1 \\ \boldsymbol{w} \end{pmatrix} + \boldsymbol{\epsilon} = \boldsymbol{\alpha} + \boldsymbol{B}_S^T\boldsymbol{w} + \boldsymbol{\epsilon}.$$

Hence $E(\boldsymbol{y}|\boldsymbol{w}) = \boldsymbol{\mu_y} + \boldsymbol{\Sigma_{yw}}\boldsymbol{\Sigma_{ww}^{-1}}(\boldsymbol{w} - \boldsymbol{\mu_w}) = \boldsymbol{\alpha} + \boldsymbol{B}_S^T\boldsymbol{w}$.

a) Show $\boldsymbol{\alpha} = \boldsymbol{\mu_y} - \boldsymbol{B}_S^T\boldsymbol{\mu_w}$.

b) Show $\boldsymbol{B}_S = \boldsymbol{\Sigma_w^{-1}}\boldsymbol{\Sigma_{wy}}$ where $\boldsymbol{\Sigma_w} = \boldsymbol{\Sigma_{ww}}$.

(Hence $\boldsymbol{B}_S^T = \boldsymbol{\Sigma_{yw}}\boldsymbol{\Sigma_w^{-1}}$.)