STAT 8260 — Theory of Linear Models
Homework 1 – Due Tuesday, Jan. 29

Homework Guidelines:

- Homework is due by 4:30 on the due date specified above. You may turn it in at the beginning of class or place it in my mailbox in the Statistics Building. **No late homeworks will be accepted without permission granted prior to the due date.**
- Use only standard (8.5 × 11 inch) paper and use only one side of each sheet.
- Homework should show enough detail so that the reader can clearly understand the procedures of the solutions. This is **absolutely essential** for you to receive full credit for your answer since the answers to most of the problems in Rencher appear in the back of the book.
- Problems should appear in the order that they were assigned.

Assignment:

1. Consider the following vectors in $\mathcal{R}^5$: $\mathbf{x} = (2, 1, 1, 1, 4)^T$, and $\mathbf{y} = (3, 6, 1, 5, 7)^T$.

   a. Find $\langle \mathbf{x}, \mathbf{y} \rangle$, $||\mathbf{x}||^2$, $||\mathbf{y}||^2$, $\hat{\mathbf{y}} = p(\mathbf{y}|\mathbf{x})$, and $\mathbf{y} - \hat{\mathbf{y}}$. Show that $\mathbf{x} \perp (\mathbf{y} - \hat{\mathbf{y}})$, and $||\mathbf{y}||^2 = ||\hat{\mathbf{y}}||^2 + ||\mathbf{y} - \hat{\mathbf{y}}||^2$.
   b. Let $\mathbf{w} = (-1, 2, 4, -4, 0)^T$ and $\mathbf{z} = 3\mathbf{x} + 2\mathbf{w}$. Show that $\langle \mathbf{w}, \mathbf{x} \rangle = 0$ and that $||\mathbf{z}||^2 = 9||\mathbf{x}||^2 + 4||\mathbf{w}||^2$. (Why must this be true?)
   c. Let $A_1 = \{1, 2\}, A_2 = \{3, 4\}, A_3 = \{5\}$. Find $p(\mathbf{y}|\mathbf{i}_{A_i})$, $i = 1, 2, 3$.

2. Is projection a linear transformation in the sense that $p(c\mathbf{y}|\mathbf{x}) = cp(\mathbf{y}|\mathbf{x})$ for any real number c? Prove or disprove. What is the relationship between $p(\mathbf{y}|\mathbf{x})$ and $p(\mathbf{y}|c\mathbf{x})$ for $c \neq 0$?

3. Suppose $||\mathbf{x}||^2 > 0$. Use calculus to prove that $||\mathbf{y} - b\mathbf{x}||^2$ is minimum for $b = \mathbf{y}^T\mathbf{x}/||\mathbf{x}||^2$. (That is, provide a calculus-based proof of the Theorem on p. 22 of the notes as an alternative to the one provided there.)

4. Here we will consider subspaces of $\mathcal{R}^4$. Let $\mathbf{x}_1 = (1, 1, 1, 0)^T$, $\mathbf{x}_2 = (1, 1, 0, 0)^T$, $\mathbf{x}_3 = (1, 0, 0, 1)^T$ and $\mathbf{x}_4 = (10, 3, 0, 7)^T$. Let $V_2 = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2)$, $V_3 = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ and $V_4 = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$.

   a. Find matrices $\mathbf{A}_2$, $\mathbf{A}_3$, $\mathbf{A}_4$ whose column spaces are equal to V_2, V_3, V_4, respectively.
   b. Find the dimensions of V_2 and V_3 (i.e., the ranks of $\mathbf{A}_2$ and $\mathbf{A}_3$).
   c. Find bases for V_2 and V_3 that contain vectors with as many zeros as possible.
   d. Give a vector $\mathbf{z} \neq \mathbf{0}$ that is orthogonal to all vectors in V_3.

e. $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and $\mathbf{z}$ are linearly independent. Therefore, $\mathbf{x}_4$ is expressible in the form $\sum_{i=1}^{3} b_i\mathbf{x}_i + c\mathbf{z}$ (this is true because one can form at most d linearly independent vectors in a vector space of dimension d). Show that $c = 0$ and hence that $\mathbf{x}_4 \in V_3$, by determining $\mathbf{x}_4^T\mathbf{z}$. What is $\dim(V_4)$ (the dimension of V_4)?

f. Give a simple description of V_3 in words.

5. Consider $\mathcal{R}^6$, Euclidean 6-space, but let's index a vector $\mathbf{y}$ in this space as $\mathbf{y} = (y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{31})^T$. Think of this indexing scheme as representing data y_{ij} from 3 treatment groups, with 3, 2, and 1 replicates (indexed by j) in groups 1, 2, and 3, respectively (indexed by i). That is, y_{ij} represents the j^{th} response in the i^{th} group where $i = 1, 2, 3$ and $j = 1, \ldots, n_i$, $n_1 = 3$, $n_2 = 2$, $n_3 = 1$. Let $\mathbf{a}_i$ be the indicator for the i^{th} treatment group (e.g., $a_1 = (1, 1, 1, 0, 0, 0)^T$, $a_2 = (0, 0, 0, 1, 1, 0)^T$, etc.). Let $V = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$.

   a. For $\mathbf{y} = (2, 1, 0, 7, 9, 3)^T$ find $\hat{\mathbf{y}} = p(\mathbf{y}|V)$, $\mathbf{y} - \hat{\mathbf{y}}$, $||\mathbf{y}||^2$, $||\hat{\mathbf{y}}||^2$, $||\mathbf{y} - \hat{\mathbf{y}}||^2$.
   b. Give a general non-matrix formula for $\hat{\mathbf{y}} = p(\mathbf{y}|V)$ for any $\mathbf{y}$.

6. Let $\mathbf{x}_1 = \mathbf{j}_4$, $\mathbf{x}_2 = (4, 1, 3, 4)^T$, $\mathbf{y} = (1, 9, 5, 5)^T$. Let $V = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2)$.

   a. Find $\hat{\mathbf{y}} = p(\mathbf{y}|V)$ and $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$.
   b. Find $\hat{\mathbf{y}}_1 = p(\mathbf{y}|\mathbf{x}_1)$ and $\hat{\mathbf{y}}_2 = p(\mathbf{y}|\mathbf{x}_2)$ and show that $\hat{\mathbf{y}} \neq \hat{\mathbf{y}}_1 + \hat{\mathbf{y}}_2$.
   c. Verify that $\mathbf{e} \perp V$.
   d. Find $||\mathbf{y}||^2$, $||\hat{\mathbf{y}}||^2$, $||\mathbf{e}||^2$, and verify that the Pythagorean Theorem holds. Compute $||\hat{\mathbf{y}}||^2$ directly from $\hat{\mathbf{y}}$ and also by using the formula $||\hat{\mathbf{y}}||^2 = \mathbf{y}^T\mathbf{P}\mathbf{y}$ where $\mathbf{P}$ is the projection matrix onto V.
   e. Use Gram-Schmidt orthogonalization to find four mutually orthogonal vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and $\mathbf{v}_4$ such that $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$. *Hint:* You can choose $\mathbf{x}_3$ and $\mathbf{x}_4$ arbitrarily, as long as $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are LIN.

7. (Simple linear regression.) Let $\mathbf{y} = (y_1, \ldots, y_n)^T$, $\mathbf{x} = (x_1, \ldots, x_n)^T$, and $V = \mathcal{L}(\mathbf{j}_n, \mathbf{x})$.

   a. Use Gram-Schmidt orthogonalization on the vectors $\mathbf{j}_n$, $\mathbf{x}$ (in this order) to find orthogonal vectors $\mathbf{j}_n, \mathbf{x}^*$ spanning V. Express $\mathbf{x}^*$ in terms of $\mathbf{j}_n$ and $\mathbf{x}$, then find b_0, b_1 such that $\hat{\mathbf{y}} = b_0\mathbf{j}_n + b_1\mathbf{x}$. To simplify the notation, let $\mathbf{y}^* = \mathbf{y} - p(\mathbf{y}|\mathbf{j}_n) = \mathbf{y} - \bar{y}\mathbf{j}_n$,

$$S_{xy} = \langle \mathbf{x}^*, \mathbf{y}^* \rangle = \langle \mathbf{x}^*, \mathbf{y} \rangle = \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i (x_i - \bar{x})y_i = \sum_i x_i y_i - n\bar{x}\bar{y},$$

$$S_{xx} = \langle \mathbf{x}^*, \mathbf{x}^* \rangle = \sum_i (x_i - \bar{x})^2 = \sum_i (x_i - \bar{x})x_i = \sum_i x_i^2 - n\bar{x}^2,$$

$$S_{yy} = \langle \mathbf{y}^*, \mathbf{y}^* \rangle = \sum_i (y_i - \bar{y})^2.$$

   b. Suppose $\hat{\mathbf{y}} = p(\mathbf{y}|V) = a_0\mathbf{j}_n + a_1\mathbf{x}^*$. Find formulas for a_1 and A_0 in terms of $\bar{y}$, S_{xy}, and S_{xx}.

c. Express $\mathbf{x}^*$ in terms of $\mathbf{j}_n$ and $\mathbf{x}$, and use this to determine formulas for b_0 and b_1 so that $\hat{\mathbf{y}} = b_0\mathbf{j}_n + b_1\mathbf{x}$.
d. Express $||\hat{\mathbf{y}}||^2$ and $||\mathbf{y} - \hat{\mathbf{y}}||^2$ in terms of S_{xy}, S_{xx} and S_{yy}.
e. Use the formula $\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ for $\mathbf{b} = (b_0, b_1)^T$, and verify that this gives the same answer as in (c).
f. for $\mathbf{y} = (2, 6, 8, 8)^T$, $\mathbf{x} = (0, 1, 2, 3)^T$ find $a_0, a_1, \hat{\mathbf{y}}, b_0, b_1, ||\mathbf{y}||^2, ||\hat{\mathbf{y}}||^2, ||\mathbf{y} - \hat{\mathbf{y}}||^2$. Verify that $||\hat{\mathbf{y}}||^2 = b_0\langle \mathbf{y}, \mathbf{j}_4\rangle + b_1\langle \mathbf{y}, \mathbf{x}\rangle$ and that $(\mathbf{y} - \hat{\mathbf{y}}) \perp V$.

8. Let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be a basis of a subspace $V \subset \mathcal{R}^n$. Suppose that $p(\mathbf{y}|V) = \sum_{j=1}^k p(\mathbf{y}|\mathbf{x}_j)$ for every vector $\mathbf{y} \in \mathcal{R}^n$. Prove that $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are mutually orthogonal. *Hint:* Consider the vector $\mathbf{y} = \mathbf{x}_i$ for each i.

9. Show that for $\mathbf{W}_{n\times k} = \mathbf{X}_{n\times k}\mathbf{B}_{k\times k}$ with $\mathbf{B}$ nonsingular and $\mathbf{X}$ of full rank, $\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ remains unchanged if $\mathbf{X}$ is replaced by $\mathbf{W}$. Thus $\mathbf{P}$ is a function of the subspace spanned by the columns of $\mathbf{X}$, not of the particular basis chosen for this subspace (we can change $\mathbf{X}$ without affecting $\mathbf{P}$ as long as we haven't changed $C(\mathbf{X})$).

10. For each subspace V of $\mathcal{R}^3$ give the corresponding projection matrix $\mathbf{P}$. In each case verify that $\mathbf{P}$ is symmetric and idempotent.

    a. $V = \mathcal{L}(\mathbf{x})$ where $\mathbf{x} = (2, -1, -1)^T$.
    b. $V = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2)$ where $\mathbf{x}_1 = (1, 1, 1)^T$, and $\mathbf{x}_2 = (1, -1, 0)^T$.

11. For the subspace $V = \mathcal{L}(\mathbf{j}_n, \mathbf{x})$ of problem 7, what is $\mathbf{P}_V$? (Note that $V = \mathcal{L}(\mathbf{j}_n, \mathbf{x}^*)$, also.) What is $\mathbf{P}_{V^\perp}$?

12. From the text do problem 2.23.

13. From the text do problem 2.24.

Note that problems 1–11 are taken from Stapleton's book. They correspond to problems 1.2.1, 1.2.2, 1.2.3, 1.3.1, 1.3.3, 1.3.4, 1.3.5, 1.3.7, 1.6.1, 1.6.4, and 1.6.5, respectively.