Statistics 600 Problem Set 1

Due in class on Tuesday, October 6th

1. Prove that the "horizontal residuals" in simple linear regression sum to zero in a least squares fit of Y (the dependent variable) on X (the independent variable). The horizontal residuals are the line segments connecting each data point X_i,Y_i to the fitted line $\hat{\alpha} + \hat{\beta}X$.

   Solution: To get the i^{th} horizontal residual, solve

   $$\hat{\alpha} + \hat{\beta}X = Y_i$$

   to get $\hat{X}_i = (Y_i - \hat{\alpha})/\hat{\beta}$, so the residual becomes $H_i \equiv X_i - (Y_i - \hat{\alpha})/\hat{\beta}$. Now if we sum these values we get

   $$\begin{aligned} \sum_i H_i &= \sum_i X_i - (Y_i - \bar{Y} + \hat{\beta}\bar{X})/\hat{\beta} \\ &= \sum_i (-Y_i + \bar{Y} + \hat{\beta}(X_i - \bar{X}))/\hat{\beta} \\ &= \sum_i (\bar{Y} - Y_i)/\hat{\beta} + \sum_i (X_i - \bar{X}) \\ &= 0. \end{aligned}$$

2. Suppose P is the projection onto a d-dimensional subspace S, and T is some other d-dimensional subspace. Show that there exists an orthogonal matrix Q such that $Q'PQ$ is the projection onto T.

   Solution: Let n be the dimension of the vector space containing S and T, let U_S be an $n \times d$ matrix whose columns are an orthonormal basis for S, and let U_T be a $n \times d$ matrix whose columns are an orthonormal basis for T. Similarly, let $U_{S^\perp}$ and $U_{T^\perp}$ be $n \times n - d$ matrices whose columns are orthonormal bases for $S^\perp$ and $T^\perp$. Then $A = U_S U_T'$ is a square matrix that maps T to S, and is orthogonal when restricted to T. Similarly, $B = U_{S^\perp} U_{T^\perp}'$ is a square matrix that maps $T^\perp$ to $S^\perp$, and is orthogonal when restricted to $T^\perp$. By direct calculation, you can confirm that $Q = A + B$ is orthogonal and maps T to S, and $T^\perp$ to $S^\perp$. Now if V is any vector and we decompose V uniquely as $V_T + V_{T^\perp}$, where $V_T \in T$ and $V_{T^\perp} \in T^\perp$, then by direct calculation, $Q'PQV = V_T$, which proves that $Q'PQ$ is the projection onto T.

3. Explicitly construct the projection matrix onto the span of $(1, 1, 1, 0)'$ and $(0, 1, 1, 1)'$.

Solution: Let $V_1 = (1,1,1,0)'/\sqrt{3}$ be the unit vector in the direction of $(1,1,1,0)'$. Then the space spanned by V_1 and $(0,1,1,1)'$ is equal to the space spanned by V_1 and $(-2,1,1,3)'/\sqrt{15}$. Thus the projection can be written

$$\frac{1}{3}\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{15}\begin{pmatrix} 4 & -2 & -2 & -6 \\ -2 & 1 & 1 & 3 \\ -2 & 1 & 1 & 3 \\ -6 & 3 & 3 & 9 \end{pmatrix} = \frac{1}{15}\begin{pmatrix} 9 & 3 & 3 & -6 \\ 3 & 6 & 6 & 3 \\ 3 & 6 & 6 & 3 \\ -6 & 3 & 3 & 9 \end{pmatrix}.$$

4. Suppose we observe data from a simple linear model $Y = \alpha + \beta X + \epsilon$ where $X, Y \in \mathcal{R}^n$, $E(\epsilon|X) = 0$ and $\mathrm{cov}(\epsilon|X) = \sigma^2 I$. Suppose X and Y are partitioned as

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \qquad X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},$$

where Y_1 and Y_2 each have half the length of Y, and X_1 and X_2 each have half the length of X. Let $\hat\beta_1$ and $\hat\beta_2$ denote the least squares estimates obtained by regressing Y_1 on X_1 and Y_2 on X_2, respectively, and let $\tilde\beta = (\hat\beta_1 + \hat\beta_2)/2$.

(a) If $\bar X_1 = \bar X_2 = \bar X$, state a condition such that $\tilde\beta$ has the same variance as the least squares estimate $\hat\beta$ obtained by regressing Y on X. Then state whether when this condition holds, $\tilde\beta$ is the least squares estimate, or is a different estimate with the same variance.

Solution: Let $T_1 = \sum_{i=1}^{n/2} Y_i(X_i - \bar X_1)$ and $T_2 = \sum_{i=n/2+1}^{n} Y_i(X_i - \bar X_2)$, and let $S_1 = \sum_{i=1}^{n/2}(X_i - \bar X_1)^2$ and $S_2 = \sum_{i=n/2+1}^{n}(X_i - \bar X_2)^2$. Then

$$\hat\beta_1 = \beta + T_1/S_1,$$

$$\hat\beta_2 = \beta + T_2/S_2.$$

and

$$\tilde\beta = \beta + \frac{T_1}{2S_1} + \frac{T_2}{2S_2}.$$

Since $\mathrm{var}(T_j) = \sigma^2 S_j$ for $j = 1, 2$, it follows that

$$\mathrm{var}\tilde\beta = \frac{\sigma^2}{4S_1} + \frac{\sigma^2}{4S_2}.$$

The variance of the least squares estimate using all the data is $\sigma^2/(S_1+S_2)$. The two variances are equal if only if

$$(S_1 + S_2)^2 = 4S_1S_2,$$

which is easily seen to hold if and only if $S_1 = S_2$. This is the condition required for the variance of $\tilde{\beta}$ to equal the variance of $\hat{\beta}$, and it is easy to see that when $S_1 = S_2$, $\tilde{\beta} = \hat{\beta}$.

(b) Now consider the more general case where $\bar{X}_1$ and $\bar{X}_2$ may differ. Show that in this case var$\tilde{\beta}$ is always greater than var $\hat{\beta}$, and derive a concise expression for the difference.

Solution: Picking up from above, the difference in variances is

$$\frac{\sigma^2}{4S_1} + \frac{\sigma^2}{4S_2} - \frac{\sigma^2}{S_1+S_2} = \frac{\sigma^2(S_1-S_2)^2}{2S_1S_2(S_1+S_2)}.$$

5. Suppose we have a bivariate regression ($p = 2$ with an intercept in the model). For simplicity, assume the covariates satisfy $\bar{X}_1 = \bar{X}_2 = 0$, var$(X_1) =$ var$(X_2) = 1$, and cov$(X_1, X_2) = r$. Derive an expression for cor$(\hat{\beta}_1, \hat{\beta}_2)$. Note which of the following values your answer does and does not depend on: β (the true regression coefficients), n (the sample size), σ^2 (the true error variance), r (as defined above).

Solution: Since the covariates are standardized, it follows that

$$X'X/n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & r \\ 0 & r & 1 \end{pmatrix}.$$

Thus,

$$\mathrm{cov}\hat{\beta} = \sigma^2 n^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/(1-r^2) & -r/(1-r^2) \\ 0 & -4/(1-r^2) & 1/(1-r^2) \end{pmatrix},$$

so

$$\mathrm{cov}(\hat{\beta}_1, \hat{\beta}_2) = -\frac{\sigma^2 r}{n(1-r^2)}$$

and cor$(\hat{\beta}_1, \hat{\beta}_2) = -r$, which obviously only depends on r.

6. Derive an expression for the sample covariance between the observed and fitted values, $\widehat{\text{cov}}(\hat{Y}, Y)$. Consider whether this covariance can either be negative, or equal to zero.

   Solution: Let P be the projection matrix onto $\text{col}(X)$. Then,

$$\begin{aligned}\widehat{\text{cov}}(\hat{Y}, Y) &= (PY)'(Y - \bar{Y})/(n-1) \\ &= \frac{(Y - \bar{Y} + \bar{Y})'P(Y - \bar{Y})}{n-1} \\ &= \frac{(Y - \bar{Y})'P(Y - \bar{Y})}{n-1} + \frac{\bar{Y}'P(Y - \bar{Y})}{n-1}.\end{aligned}$$

   Here, $\bar{Y}$ is interpreted as an n-vector in which all values are equal to the sample mean of the Y_i. This can be written $\bar{Y} = n^{-1}\mathbf{1}\mathbf{1}'Y$, where $\mathbf{1}$ is a n-vector of 1's. Since there is an intercept in the model, $P\mathbf{1} = 1$, so the second summand above is equal to

$$\frac{n^{-1}\mathbf{1}\mathbf{1}'(Y - \bar{Y})}{n-1},$$

   which is zero since $\mathbf{1}'(Y - \bar{Y}) = 0$. Thus

$$\text{cov}(\hat{Y}, Y) = \frac{(Y - \bar{Y})'P(Y - \bar{Y})}{n-1} \geq 0.$$

7. (a) Is the product of projection matrices a projection matrix? Prove the result, or give a counterexample.

   Solution: The statement is false. Let V and W be unit vectors, so that VV' and WW' are both projection matrices. Idempotence is a necessary condition for a matrix to be a projection, we can check whether the product $VV'WW'$ is idempotent. First note that

$$VV'WW' = (V'W)VW'.$$

   Then,

$$VV'WW'VV'WW' = (V'W)^3VW'.$$

   Thus the matrix is idempotent if and only if $V'W$ equals 0, -1, or 1. In any other case, it is not idempotent, so cannot be a projection.

(b) Is the product of square orthogonal matrices a square orthogonal matrix? Prove the result, or give a counterexample.

Solution: For a square matrix M to be orthogonal, it is sufficient that $M'M = I$. If A and B are orthogonal and $C = AB$, then $C'C = B'A'AB = I$, so the statement is true.