Homework 4

1. Consider the effects of multicollinearity on predictions from a standard linear model.

   (a) How does multicollinearity affect the global measure of prediction performance, trace($\Sigma_{\hat{Y}\hat{Y}}$)?

   Since trace($\Sigma_{\hat{Y}\hat{Y}}$) $= p$ there is no overall, or average, effect. However, consider the next problem.

   (b) Consider predicting a new response with covariate value x_0. What is the effect of multicollinearity?

$$\begin{aligned} Var(x_0\hat{\beta}) &= \sigma^2 x_0 (X^T X)^{-1} x_0^T \\ &= \sigma^2 x_0 \left(\sum \lambda_i^{-2} V_i V_i^T \right) x_0^T \\ &= \sigma^2 \sum \lambda_i^{-2} \|x_0 V_i\|^2 \end{aligned}$$

   So if x_0 has a large projection onto an eigenvector of $X^T X$ which has small eigenvalue, the prediction will have large variance. The variance will be largest when x_0 is proportional to the eigenvector with smallest eigenvalue and will be largest when proportional to the eigenvector with largest eigenvalue.

2. Consider a standard linear model with design matrix

$$X = \begin{pmatrix} -2 & -2 \\ -1 & -1 \\ 0 & 0 \\ 1 & 1 \\ 2 & 2 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}$$

   and unknown coefficients β_1, β_2.

   (a) Which linear combinations $\hat{\psi} = a_1\hat{\beta}_1 + a_2\hat{\beta}_2$ have largest and smallest variance, subject to $a_1^2 + a_2^2 = 1$?

   See the previous problem. The eigenvalues of $X^T X$ are 4 and 20, corresponding to vectors $(1/\sqrt{2}, -1/\sqrt{2})$ and $(1/\sqrt{2}, 1/\sqrt{2})$ respectively.

(b) If the columns of the design matrix are thought of as values continuous variables x_1 and x_2, then we are fitting a plane through the origin: $y = \beta_1 x_1 + \beta_2 x_2$. If $u = (u_1, u_2)$ is a direction vector, what is the estimated slope of the plane in that direction? In what direction is the estimated slope most variable? Least variable? Do the answers "make sense?"

If u is a unit vector the slope in the direction u is $\beta_1 U_1 + \beta_2 u_2$, so the results above apply. Plot the points in the plane and observe that they are spread out well along the line $x_1 = x_2$, but not as much in the orthogonal direction $x_1 = -x_2$. Intutively, then, there is better resolution in the former direction.

3. Suppose that $Y \sim MVN(0, \sigma^2 I_n)$. Show that $\bar{Y}$ is independent of $\sum_{i=1}^{n}(Y_i - \bar{Y})^2$.

Verify that $\Sigma_{\bar{Y}, Y-\bar{Y}} = 0$. Under normality, they are thus independent. Finally, use the fact that if U and V are independent, then U is independent of $g(V)$ for any function g.