HW 3

You may (and should) talk about problems with each other and with me, providing attribution for any good ideas you might get. Your final write-up should be your own.

1: Funny formulations Suppose that C is symmetric and positive definite. Show that solving the generalized least squares problem

$$\text{minimize } \|Ax - b\|_{C^{-1}}^2$$

is equivalent to solving the linear system

$$\begin{bmatrix} C & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} v \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}.$$

2: Criss-cross Suppose $A \in \mathbb{R}^{m \times n}$ is full rank and $A = QR$ is an economy QR decomposition. The *leverage score* for row i is the squared norm of row i, i.e. $\nu_i = \sum_{j=1}^m q_{ij}^2$. In this exercise, we will show that if $r = Ax - b$ is the residual in a least squares problem with $x = A^\dagger b$, then the leave-one-out cross-validation (LOOCV) error for row i is $r_i/(1 - \nu_i)$.

1. Without loss of generality, consider $i = n$, and write A, Q, b, and $r = b - QQ^T b$ as
$$A = \begin{bmatrix} A_1 \\ a_2 \end{bmatrix}, \quad Q = \begin{bmatrix} U \\ v^T \end{bmatrix}, \quad b = \begin{bmatrix} c \\ \beta \end{bmatrix}, \quad r = \begin{bmatrix} s \\ \rho \end{bmatrix}.$$
Argue that the $\tilde{\beta} = a_2 A_1^\dagger c$ (the predicted value of the last row based on a model fit to all the other rows) can also be written as $\tilde{\beta} = v^T U^\dagger c$.

2. Argue that if $y = Q^T b$ and $\hat{y} = U^\dagger c$, then
$$\hat{y} = (I - vv^T)^{-1}(y - v\beta).$$

3. The LOOCV error for the last row is $\tilde{\beta} - \beta$. Show how to write this as $\tilde{\beta} - \beta = -\rho/(1 - \nu)$, where $\nu = \|v\|^2$.

Usually, the total LOOCV error is written as the mean squared prediction error over the leave-one-out models, i.e. (for a linear model)

$$\text{LOOCV} = \frac{1}{m} \sum_{i=1}^m \left(\frac{r_i}{1 - \nu_i} \right)^2.$$

Computing the LOOCV statistic is more expensive for nonlinear models!

3: Perturbed projectors Suppose we have two least squares systems

$$\begin{aligned} Ax &= b + r, & A^T r &= 0, \\ \hat{A}\hat{x} &= \hat{b} + \hat{r}, & \hat{A}^T \hat{r} &= 0, \end{aligned}$$

and both systems are full rank.

1. Show that
$$\frac{\|\hat{r} - r\|}{\|b\|} \leq \frac{\|\hat{b} - b\|}{\|b\|} + \|\hat{\Pi} - \Pi\|$$
where $\Pi = AA^\dagger$ and $\hat{\Pi} = \hat{A}\hat{A}^\dagger$.

2. Show that if $X, Y \in \mathbb{R}^{m \times n}$ are any matrices such that $X^T Y = 0$, then $\|X + Y\|^2 \leq \|X\|^2 + \|Y\|^2$.

3. Also show (using an SVD) that if Z is any square matrix, then $\|ZZ^T - I\| = \|Z^T Z - I\|$.

4. Show that if $A = QR$ and $\hat{A} = \hat{Q}\hat{R}$ and $\hat{Q} = QZ + E$ for $Z = Q^T \hat{Q}$, then
$$\|\hat{\Pi} - \Pi\| \leq \sqrt{3}\|E\|$$
Hint: Write $\hat{\Pi} - \Pi = Q(ZZ^T - I)Q^T + QZE^T + E\hat{Q}^T$ and exploit the previous two points. You will need to argue that $\|Z\| \leq 1$ and that the range spaces of Q and of E are orthogonal.

We can also give a geometric interpretation, as the norm $\|E\|$ is the sine of the largest *canonical angle* between the range spaces of $\hat{Q}$ and Q.

4: Contemplate constraints Consider the constrained problem

$$\text{minimize } \|Ax - b\|^2 \text{ s.t. } C^T x = 0$$

and suppose that we are given an economy factorization $A = QR$. Using the Lagrange multiplier constraint formulation from class, show that Rx is the residual in the unconstrained least squares problem

$$\text{minimize } \|R^{-T} C\lambda - Q^T b\|^2.$$

Hint: Look to the formulation from problem 1.