Math 210 **Fall 2019**
Homework 5 **Drew Armstrong**

Problem 1. We say that P is a *projection matrix* if $P^T = P$ and $P^2 = P$.

(a) If P is a projection, show that $I - P$ is also a projection.
(b) Show that the projections P and $I - P$ satisfy $P(I - P) = 0$.
(c) Let A be any matrix of shape $m \times n$ so that $A^T A$ is square of shape $n \times n$. Assuming that the inverse $(A^T A)^{-1}$ exists, show that $P = A(A^T A)^{-1} A^T$ is a projection matrix. [We saw in class that this matrix projects onto the column space of A.]
(d) In the special case that A is a square and invertible, show that $P = A(A^T A)^{-1} A^T = I$. What does this mean?

(a) Let P be a projection matrix, so that $P^T = P$ and $P^2 = P$ and define $Q := I - P$. Then Q is also a projection matrix because

$$Q^T = (I - P)^T = I^T - P^T = I - P = Q$$

and

$$Q^2 = (I - P)^2 = II - IP - PI + P^2 = I - P - P + P = I - P = Q.$$

(b) Continuing from (a), we have $PQ = P(I - P) = PI - P^2 = P - P = 0$. (This notation represents the matrix of zeroes. I can't think of a better notation for it. Maybe O?)

(c) Let A be any matrix such that $(A^T A)^{-1}$ exists, and define $P = A(A^T A)^{-1} A^T$. Then

$$\begin{aligned} P^2 &= [A(A^T A)^{-1} A^T][A(A^T A)^{-1} A^T] \\ &= A(A^T A)^{-1}(A^T A)(A^T A)^{-1} A^T \\ &= A\cancel{(A^T A)^{-1}(A^T A)}(A^T A)^{-1} A^T \\ &= AI(A^T A)^{-1} A^T \\ &= A(A^T A)^{-1} A^T = P \end{aligned}$$

and

$$\begin{aligned} P^T &= [A(A^T A)^{-1} A^T]^T \\ &= (A^T)^T[(A^T A)^{-1}]^T A^T \\ &= A[(A^T A)^{-1}]^T A^T \\ &= A[(A^T A)^T]^{-1} A^T \\ &= A[A^T (A^T)^T]^{-1} A^T \\ &= A[A^T A]^{-1} A^T = P. \end{aligned}$$

Hence P is a projection matrix. In fact, one can show that every projection matrix has this form. (But we won't.)

(d) Continuing from (c), suppose that A is square and A^{-1} exists. Then

$$P = A(A^T A)^{-1} A^T = \cancel{AA^{-1}(A^T)^{-1}A^T} = II = I.$$

Explanation: The matrix $P = A(A^T A)^{-1} A^T$ projects onto the column space of A. If A is square and invertible then its column space is everything. We observe that

$$\text{project onto everything} \;=\; \text{do nothing.}$$

Problem 2. Consider the plane $x + 2y + 2z = 0$ with normal vector $\mathbf{a} = (1, 2, 2)$.

(a) Use the formula from 1(c) to find the 3×3 matrix P that projects onto the line $t\mathbf{a}$. [Hint: Just let $A = \mathbf{a}$.]

(b) Use the matrix P to project the vector $\mathbf{b} = (1, -1, 1)$ onto the line.

(c) Find two vectors in the plane $x + 2y + 2z = 0$ and then use the formula from 1(c) to find the 3×3 matrix Q that projects onto the plane. [Hint: Let A be the 3×2 matrix whose columns are the two vectors that you found.]

(d) Use the matrix Q to project the vector $\mathbf{b} = (1, -1, 1)$ onto the plane.

(e) Finally, check that $P + Q = I$. Does this surprise you?

(a) If $A = \mathbf{a}$ is a $n \times 1$ matrix then the column space is just the line $t\mathbf{a}$, and the matrix product $\mathbf{a}^T\mathbf{a} = \mathbf{a} \bullet \mathbf{a} = \|\mathbf{a}\|^2$ is just a number. The matrix that projects onto the line $t\mathbf{a}$ is

$$P = \mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a} = \mathbf{a}(\|\mathbf{a}\|^2)^{-1}\mathbf{a} = \frac{1}{\|\mathbf{a}\|^2}\mathbf{a}\mathbf{a}^T.$$

In the case of $\mathbf{a} = (1, 2, 2)$ we obtain

$$P = \frac{1}{9}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}.$$

(b) Then we project the vector $\mathbf{b} = (1, -1, 1)$ onto the line as follows:

$$P\mathbf{b} = \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

(c) Now consider the plane $\mathbf{a}^T\mathbf{x} = x + 2y + 2z = 0$, which is perpendicular to the line $t\mathbf{a}$ and let Q be the matrix that projects onto the plane. We know that $Q = B(B^TB)^{-1}B^T$, where $B = \begin{pmatrix} \mathbf{u} & \mathbf{v} \end{pmatrix}$ is any 3×2 matrix whose columns $\mathbf{u}$ and $\mathbf{v}$ span the plane. Let's pick $\mathbf{u} = (-2, 1, 0)$ and $\mathbf{v} = (-2, 0, 1)$.[1] Then we have

$$B^TB = \begin{pmatrix} -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}\begin{pmatrix} -2 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix},$$

and hence[2]

$$(B^TB)^{-1} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}^{-1} = \frac{1}{9}\begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}.$$

Finally, we compute

$$Q = B(B^TB)^{-1}B^T = \begin{pmatrix} -2 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{9}\begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}\begin{pmatrix} -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

(d) We project the vector $\mathbf{b} = (1, -1, 1)$ onto the plane as follows:

$$Q\mathbf{b} = \frac{1}{9}\begin{pmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 8 \\ -11 \\ 7 \end{pmatrix}.$$

[1] These are the vectors you get by letting $y = s$ and $z = t$ be parameters.

[2] Use Gaussian elimination if you need to.

(e) We observe that $P\mathbf{b} + Q\mathbf{b} = \mathbf{b}$. Indeed, the same identity would hold for any vector $\mathbf{b}$ since the four points $\mathbf{b}, P\mathbf{b}, Q\mathbf{b}, \mathbf{0}$ lie at the vertices of a rectangle. It follows that $P + Q$ is the identity matrix. Remark: We could have used this as a shortcut to compute Q. See the next problem.

Problem 3. Shortcut. Let $\mathbf{a} = (1, 2, -1, 1)$ and consider the following hyperplane in $\mathbb{R}^4$:

$$\mathbf{a}^T\mathbf{x} = 1x_1 + 2x_2 - 1x_3 + 1x_4 = 0.$$

(a) Use 1(c) to compute the matrix P that projects onto the line $t\mathbf{a}$.
(b) We could also use 1(c) to compute the matrix Q that projects onto the hyperplane, but this would take too long. Instead, use the shortcut formula $Q = I - P$.
(c) Project the point $(1, 2, 3, 4)$ onto the hyperplane.

(a) The matrix that projects onto the line is

$$P = \frac{1}{\|\mathbf{a}\|^2}\mathbf{a}\mathbf{a}^T = \frac{1}{7}\begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 & 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -2 & 2 \\ -1 & -2 & 1 & -1 \\ 1 & 2 & -1 & 1 \end{pmatrix}.$$

(b) The matrix that projects onto the hyperplane is

$$Q = I - P = \frac{1}{7}\begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} - \frac{1}{7}\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -2 & 2 \\ -1 & -2 & 1 & -1 \\ 1 & 2 & -1 & 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 6 & -2 & 1 & -1 \\ -2 & 3 & 2 & -2 \\ 1 & 2 & 6 & 1 \\ -1 & -2 & 1 & 6 \end{pmatrix}$$

(c) We project the point $\mathbf{b} = (1, 2, 3, 4)$ onto the hyperplane as follows:

$$Q\mathbf{b} = \frac{1}{7}\begin{pmatrix} 6 & -2 & 1 & -1 \\ -2 & 3 & 2 & -2 \\ 1 & 2 & 6 & 1 \\ -1 & -2 & 1 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 \\ 2 \\ 27 \\ 22 \end{pmatrix}.$$

Problem 4. Find the best fit line $C + tD = b$ for the data points

$$\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

using the following steps:

(a) Write down the matrix equation $A\mathbf{x} = \mathbf{b}$ that **would be** true if all four points were on the same line $C + tD = b$. This equation has no solution.
(b) Now write down the normal equation $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ and solve it to find the least squares approximation $\hat{\mathbf{x}} = (C, D)$.
(c) Compute the error vector $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$.
(d) Finally, draw the four data points along with their best fit line. Label the vertical errors with the entries of the error vector $\mathbf{e}$.

(a) Here is the unsolvable equation $A\mathbf{x} = \mathbf{b}$:

$$\left\{\begin{array}{rcl} C - 1D &=& 3 \\ C + 0D &=& 2 \\ C + 1D &=& 2 \\ C + 2D &=& 1 \end{array}\right\} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

(b) The (solvable) normal equation is $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 8 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} C \\ D \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 6 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 8 \\ 1 \end{pmatrix}$$

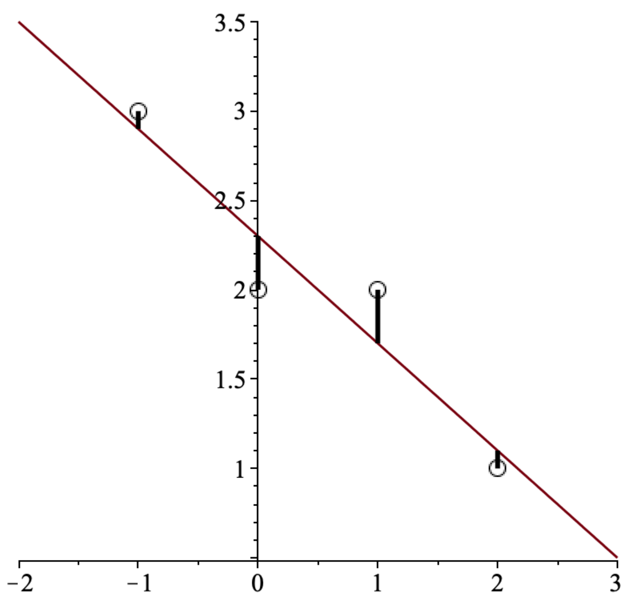
$$\begin{pmatrix} C \\ D \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 46 \\ -12 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 23 \\ -6 \end{pmatrix}$$

We conclude that the best fit line is $b = \frac{23}{10} - \frac{6}{10}t$.

(c) The error vector (height of data points minus height of the best fit line) is

$$\mathbf{b} - P\mathbf{b} = \mathbf{b} - A\hat{\mathbf{x}} = \begin{pmatrix} 3 \\ 2 \\ 2 \\ 1 \end{pmatrix} - \frac{1}{10} \begin{pmatrix} 29 \\ 23 \\ 17 \\ 11 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}.$$

(d) Picture:



Problem 5. Find the best fit parabola $C + tD + Et^2 = b$ for the data points

$$\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

using the following steps:

(a) Write down the matrix equation $A\mathbf{x} = \mathbf{b}$ that **would be** true if all four points were on the same parabola $C + tD + t^2E = b$. This equation has no solution.
(b) Now write down the normal equation $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ and solve it to find the least squares approximation $\hat{\mathbf{x}} = (C, D, E)$.
(c) Compute the error vector $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$.
(d) Finally, draw the four data points along with their best fit parabola. Label the vertical errors with the entries of the error vector $\mathbf{e}$.

(a) Here is the unsolvable equation $A\mathbf{x} = \mathbf{b}$:

$$\left\{ \begin{array}{rcl} C - 1D + 1E &=& 3 \\ C + 0D + 0E &=& 0 \\ C + 1D + 1E &=& 0 \\ C + 2D + 4E &=& 1 \end{array} \right\} \Leftrightarrow \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

(b) The (solvable) normal equation is $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix}$$

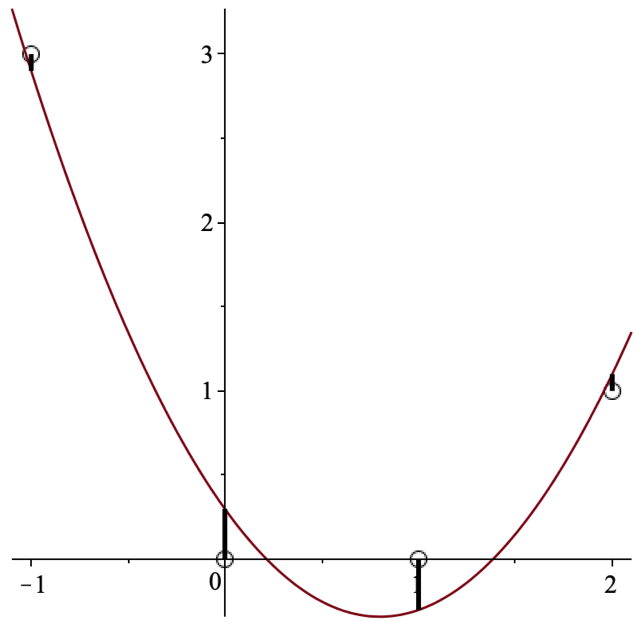
$$\begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix} = \begin{pmatrix} 3/10 \\ -8/5 \\ 1 \end{pmatrix}.$$

(I used a computer in the final step.) We conclude that the best fit parabola is $b = \frac{3}{10} - \frac{8}{5}t + t^2$.

(c) The error vector (height of data points minus height of the best fit parabola) is

$$\mathbf{b} - P\mathbf{b} = \mathbf{b} - A\hat{\mathbf{x}} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{10} \begin{pmatrix} 29 \\ 3 \\ -3 \\ 11 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}.$$

(d) Picture:


3
2
1
−1
0
1
2