Math 210 **Fall 2020**
Homework 5 **Drew Armstrong**

Problem 1. Special Matrices. For any angle θ we define the following matrices:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad F_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \quad P_\theta = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

(a) Describe what each matrix does geometrically.
(b) Compute the determinant of each matrix.
(c) For each matrix that is invertible, compute the inverse.

(a): The matrix R_θ **rotates** counterclockwise by angle θ. The matrix F_θ **reflects** across the line with angle $\theta/2$. [See the lecture notes for dicsussion.] The matrix P_θ **projects** onto the line with angle θ. Indeed, the matrix that projects onto the line $t\mathbf{a} = t(\cos\theta, \sin\theta)$ is

$$P = \frac{1}{\|\mathbf{a}\|^2}\mathbf{a}\mathbf{a}^T = \frac{1}{\cos^2\theta + \sin^2\theta}\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \end{pmatrix} = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

(b): To save space we will write $c = \cos\theta$ and $s = \sin\theta$. Then the determinants are

$$\begin{aligned} \det(R_\theta) &= c^2 + s^2 = 1, \\ \det(F_\theta) &= -c^2 - s^2 = -1, \\ \det(P_\theta) &= c^2s^2 - cscs = 0. \end{aligned}$$

Note that these determinants do not depend on the angle θ. [Remark: It is a general phenomenon that rotations have determinant 1, reflections have determinant -1 and projections have determinant 0.]

(c): Recall that a matrix is invertible if and only if its determinant is not zero. Thus P_θ is not invertible. The inverses of R_θ and F_θ are given by

$$R_\theta^{-1} = \frac{1}{\det(R_\theta)}\begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

and

$$F_\theta^{-1} = \frac{1}{\det(R_\theta)}\begin{pmatrix} -c & -s \\ -s & c \end{pmatrix} = \begin{pmatrix} c & s \\ s & -c \end{pmatrix}.$$

[Remark: We observe that $R_\theta^{-1} = R_{-\theta}$ because rotating by $-\theta$ is the opposite of rotating by θ. And we observe that $F_\theta^{-1} = F_\theta$ because reflecting twice is the same as doing nothing.]

Problem 2. Projections in General.[1] We call P a *projection* if $P^T = P$ and $P^2 = P$.

(a) If P is a projection, show that $Q = I - P$ is also a projection.
(b) Show that the projections P and Q from part (a) satisfy $PQ = 0$.
(c) Let A be any matrix (possibly non-square), so that A^TA is a square matrix. Assuming that $(A^TA)^{-1}$ exists, show that $P = A(A^TA)^{-1}A^T$ is a projection. [We saw in class that this matrix projects orthogonally onto the **column space** of A.]
(d) In the special case that A is invertible, show that $P = A(A^TA)^{-1}A^T = I$. What does this mean? [Hint: The column space of an invertible matrix is the whole space.]

---

[1]Technically, these matrices are called *orthogonal projections* because they project at right angles.

(a): Let P be a projection so that $P^T = P$ and $P^2 = P$. Then we have

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

and

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

so that Q is also a projection.

(b): We have $PQ = P(I-P) = PI - P^2 = P - P = 0$. [Similarly, we have $QP = 0$.] Geometric Meaning of (a) and (b): P and Q are projections onto a pair of orthogonal subspaces. See the lecture notes for discussion.

(c): To show that $P = A(A^TA)^{-1}A^T$ is a projection we first observe that $P^2 = P$:

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

To show that $P^T = P$ we will use the matrix identities $(ABC)^T = C^TB^TA^T$, $(A^T)^T = A$ and $(B^{-1})^T = (B^T)^{-1}$:

$$\begin{aligned} P^T &= [A(A^TA)^{-1}A^T]^T & \\ &= (A^T)^T[(A^TA)^{-1}]^TA^T & (ABC)^T = C^TB^TA^T \\ &= A[(A^TA)^{-1}]^TA^T & (A^T)^T = A \\ &= A[(A^TA)^T]^{-1}A^T & (B^{-1})^T = (B^T)^{-1} \\ &= A[A^T(A^T)^T]^{-1}A^T & (BA)^T = A^TB^T \\ &= A[A^TA]^{-1}A^T & (A^T)^T = A \\ &= P. & \end{aligned}$$

(d): If A^{-1} exists then we use the identity $(BA)^{-1} = A^{-1}B^{-1}$ to observe that

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

Geometric Meaning: The matrix P is the projection onto the column space of A. If A is invertible then the columns of A are independent, hence the column space of A is **the whole space**. To project a point into the whole space we **do nothing** because the point is already in the whole space. Geometrically we would never consider this case; we only do it to check that the algebra makes sense.

Problem 3. Specific Projections. Consider the following matrices:

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

(a) Compute the 3×3 matrix $P = \mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T$ that projects onto the column space of $\mathbf{a}$, i.e., the matrix that projects onto the line $t(1, 1, -1)$.

(b) Compute the 3×3 matrix $Q = A(A^TA)^{-1}A^T$ that projects onto the column space of A, i.e., the matrix that projects onto the plane $s(1, 2, 3) + t(1, 1, 2)$.

(c) Check that $P+Q=I$ and $PQ=0$. Why does this happen? [Hint: How are the line from part (a) and the plane from part (b) related to each other?]

(a): The matrix that projects onto the line $t\mathbf{a}=t(1,1,-1)$ is

$$\begin{aligned}
P &= \mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T \\
&= \begin{pmatrix}1\\1\\-1\end{pmatrix}\left(\begin{pmatrix}1 & 1 & -1\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}\right)^{-1}\begin{pmatrix}1 & 1 & -1\end{pmatrix} \\
&= \begin{pmatrix}1\\1\\-1\end{pmatrix}(3)^{-1}\begin{pmatrix}1 & 1 & -1\end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix}1\\1\\-1\end{pmatrix}\begin{pmatrix}1 & 1 & -1\end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix}1 & 1 & -1\\1 & 1 & -1\\-1 & -1 & 1\end{pmatrix}.
\end{aligned}$$

(b): The matrix that projects onto the plane $C(A)=s(1,2,3)+t(1,1,2)$ is

$$\begin{aligned}
Q &= A(A^TA)^{-1}A^T \\
&= \begin{pmatrix}1 & 1\\2 & 1\\3 & 2\end{pmatrix}\left[\begin{pmatrix}1 & 2 & 3\\1 & 1 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\2 & 1\\3 & 2\end{pmatrix}\right]^{-1}\begin{pmatrix}1 & 2 & 3\\1 & 1 & 2\end{pmatrix} \\
&= \begin{pmatrix}1 & 1\\2 & 1\\3 & 2\end{pmatrix}\begin{pmatrix}14 & 9\\9 & 6\end{pmatrix}^{-1}\begin{pmatrix}1 & 2 & 3\\1 & 1 & 2\end{pmatrix} \\
&= \begin{pmatrix}1 & 1\\2 & 1\\3 & 2\end{pmatrix}\frac{1}{3}\begin{pmatrix}6 & -9\\-9 & 14\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\1 & 1 & 2\end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix}1 & 1\\2 & 1\\3 & 2\end{pmatrix}\begin{pmatrix}-3 & 3 & 0\\5 & -4 & 1\end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix}2 & -1 & 1\\-1 & 2 & 1\\1 & 1 & 2\end{pmatrix}.
\end{aligned}$$

(c): We have

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

and

$$PQ=\frac{1}{3}\begin{pmatrix}1 & 1 & -1\\1 & 1 & -1\\-1 & -1 & 1\end{pmatrix}\frac{1}{3}\begin{pmatrix}2 & -1 & 1\\-1 & 2 & 1\\1 & 1 & 2\end{pmatrix}=\frac{1}{9}\begin{pmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}=0.$$

This happens because the line in part (a) and the plane in part (b) are orthogonal complements. [I deliberately chose them that way. First I picked the columns of A and then I let $\mathbf{a}$ be their cross product. Projections onto some random line and plane would not satisfy this.]

Problem 4. Least Squares Approximation. Consider the following two lines in $\mathbb{R}^3$:

$$L_1 : (x, y, z) = (0, 0, 0) + s(1, 1, 1), \quad L_2 : (x, y, z) = (1, 0, 0) + t(-1, 1, 0).$$

(a) Write down the system of three linear equations in s, t that expresses the intersection of the two lines. [This system has no solution because the lines do **not** intersect.]
(b) Find the OLS approximations $\hat{s}$ and $\hat{t}$ for the system in part (a).
(c) Use your answer from (b) to compute the minimum distance between the two lines.

(a): A general point of L_1 has the form $(x, y, z) = (s, s, s)$ and a general point of L_2 has the form $(x, y, z) = (1 - t, t, 0)$. If the two lines intersect them we will have $(s, s, s) = (1 - t, t, 0)$, which gives a system of 3 linear equations in the 2 unknowns s, t:

$$\begin{cases} s &= 1 - t, \\ s &= t, \\ s &= 0. \end{cases} \quad \Rightarrow \quad \begin{cases} s + t &= 1, \\ s - t &= 0 \\ s + 0 &= 0. \end{cases}$$

(b): To find approximate solutions $\hat{s}, \hat{t}$ we consider the normal equation:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{s} \\ \hat{t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \hat{s} \\ \hat{t} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} \hat{s} \\ \hat{t} \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/2 \end{pmatrix}.$$

(c): The points on L_1 and L_2 that come closest to each other are $(\hat{s}, \hat{s}, \hat{s}) = (1/3, 1/3, 1/3)$ and $(1 - \hat{t}, \hat{t}, 0) = (1/2, 1/2, 0)$. The distance between these points is

$$\left\| \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} \right\| = \sqrt{(1/3 - 1/2)^3 + (1/3 - 1/2)^2 + (1/3 - 0)^2} = \sqrt{1/6}.$$

See the lecture notes for a picture.

Problem 5. Least Squares Regression. Consider four data points:

$$(x, y) = (1, 1), (2, 1), (3, 3), (4, 5).$$

(a) Find the OLS best fit line $y = mx + b$ for these points. Draw your answer.
(b) Find the OLS best fit parabola $y = ax^2 + bx + c$ for the same points. Draw your answer.

[I recommend using a computer algebra system to solve the normal equations.]

(a): Each data point gives a linear equation in m and b. This system of 4 linear equations in 2 unknowns has no solution, so we solve the normal equation:

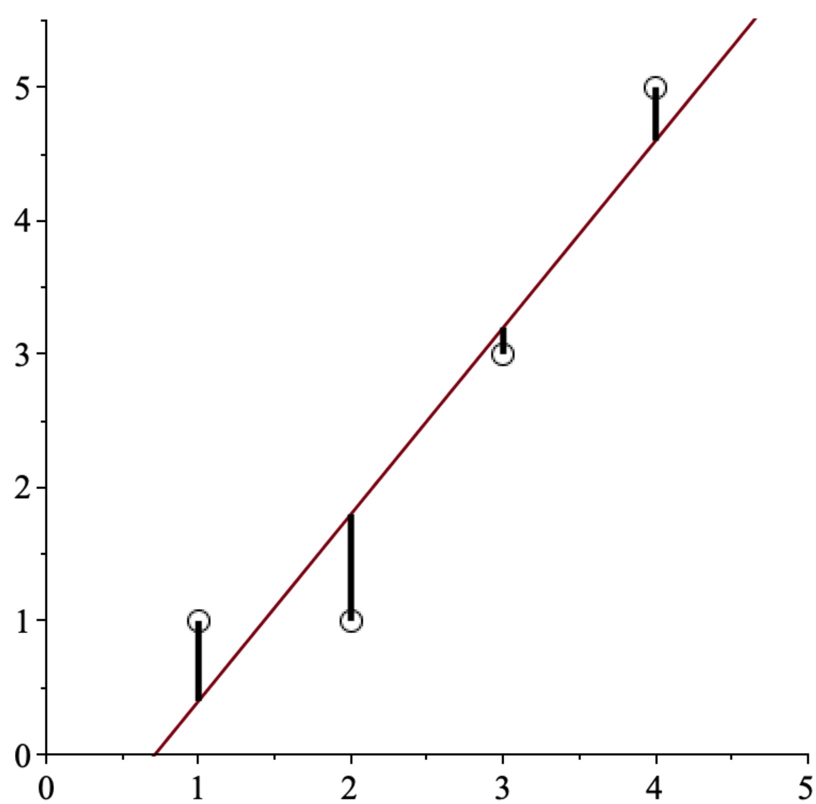
$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} \hat{m} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} \hat{m} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 32 \\ 10 \end{pmatrix}$$

$$\begin{pmatrix} \hat{m} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 7/5 \\ -1 \end{pmatrix}.$$

The best fit line is $y = \hat{m}x + \hat{b} = (7/5)x - 1$. Here is a picture:



(b): Each data point gives a linear equation in a, b, c. This system of 4 linear equations in 3 unknowns has no solution, so we solve the normal equation:

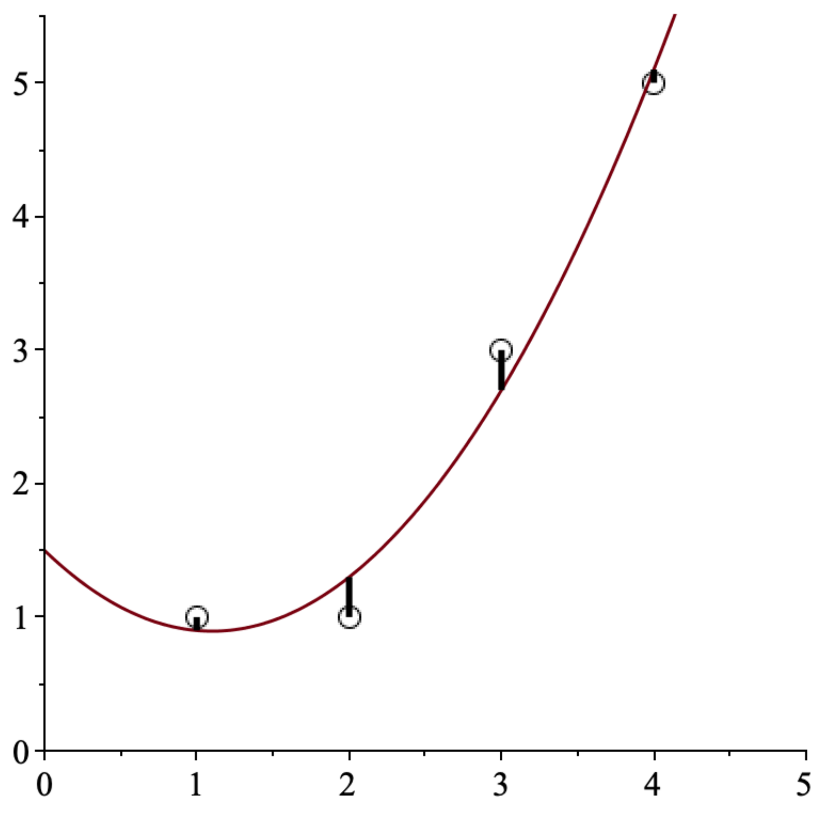
$$\begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} 112 \\ 32 \\ 10 \end{pmatrix}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} 1/2 \\ -11/10 \\ 3/2 \end{pmatrix}$$

The best fit parabola is $y = \hat{a}x^2 + \hat{b}x + \hat{c} = (1/2)x^2 - (11/10)x + (3/2)$. Here is a picture:



Observe that this parabola is a "better fit" than the best fit line. That is, the sum of the squares of the vertical errors is smaller. In fact, one can check that these sums are 6/5 in (a) and 1/5 in (b). So I guess (b) is six times "better" than (a).