Marcus Rose
Math 126 HW 6
Due 3/16/23

---

1. Q1

   (a) Q1a

   Answer Note the definition of Least Squares (LS): An $x \in \mathbb{R}^n$ that minimizes $\|b-Ax\|_2$ is called a least squares solution of $Ax = b$. Then, for

   $$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \implies \min_x [(x_1 + x_2 - 2)^2 + (x_1 - x_2)^2 + (x_1 + 2x_2 - 2)^2]$$

   The x_1, x_2 solution that minimizes the above system is minimal iff

   $$A^T Ax = A^T b,$$

   and while the least squares solution x here might not need to be unique, Ax is indeed unique. This is also the solution is the sum of the elements that span the null space of A. So,

   $$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix},$$

   $$A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \implies \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} x = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \implies x = \begin{bmatrix} \frac{6}{7} \\ \frac{5}{7} \end{bmatrix}.$$

   (b) Q1b

   Answer Note as mentioned above, while x need not be unique; however, Ax is unique. In order for there to exist a LS solution, $A^T Ax = A^T b$ must hold. Because there is only one transformation Ax such that x is a LS solution, this x that minimizes the aforementioned expression must be the only LS solution to the system.

2. If $Ax = b$, then the set of solutions of $Ax = b$ must be equal to the set of solutions of $A^T Ax = A^T b$. Then verify this statement explicitly for Example 14.7.

   Proof Let x_1 be a solution of $A^T Ax = A^T b$. Let x_2 be a solution of $Ax = b$. WTS $\text{Null}(A) = x_1 - x_2$. If x_2 is a solution of $A^T Ax = A^T b$ as well as the given solution for $Ax = b$, then that is to say that $x_1 - x_2$ is in the nullspace of $A^T A$. I.e.

   $$(x_1 - x_2) \in (A) \iff A(x_1 - x_2) = 0 \iff A(x_1) = A(x_2) = b$$

   Therefore, x_1 is a solution of $Ax = b$. To see why the nullspace of $A^T A$ is equal to the nullspace of A^T, see the auxiliary proof below. $\square$

Sub-Proof Let x be a vector in the nullspace of A^TA. Then,

$$A^TAx = 0 \implies x^TA^TAx = \|A^Tx\|_2,$$

and x is also a vector in the nullspace of A^T. It follows naturally that $\text{rank}(A) = \text{rank}(A^T) = \text{rank}(A^TA)$, as will be shown in the following problem. $\square$

In example 14.7, note

$$x = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}.$$

Take the first vector in this set of solutions. Then, $x = \begin{bmatrix} 0 & 1 & -2 \end{bmatrix}^T$ is a solution of

$$A^TAx = A^Tb.$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \implies \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Because this x can have infinitely many solutions because of the 0 in the x_2 position, note that the set of solutions for $A^TAx = A^Tb$ is still the set of solutions to the normal equations.

3. Q3

   (a) Q3a

   Answer This matrix is full rank and the rank is 2. All of the columns in the matrix fully span the column space, i.e. all columns are linearly independent. A^TA is also full rank with rank 2.

   (b) Q3b

   Answer With $p = 4$, the new matrix under the float point number system $\mathbb{F}_4$ is

   $$A_{p=4} \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, A_{p=4}^TA_{p=4} = \begin{bmatrix} 1.1 \times 10^{-8} & 1 \\ 1 & 1.1 \times 10^{-8} \end{bmatrix}$$

   and this matrix is full rank. Its columns while close to being dependent, are still linearly independent. Now with $p = 9$,

   $$A_{p=9} \begin{bmatrix} 1 & 1 \\ 0.001000000 & 0 \\ 0 & 0.001000000 \end{bmatrix}, A_{p=9}^TA_{p=9} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

   and we're back to the rank 1 matrix. It has rank 1 because of the equivalent vectors in the column space.

   (c) Q3c

Answer Comparing the two ranks under different floating point number regimes, we can see that computing least squares via the normal equations might result in an extremely small value that is read as zero in some floating point system $\mathbb{F}$. Therefore, we should exercise caution not to accidentally run into a dimensionality reduction of A due to a transformation that effects one of the basis vectors *just slightly.*

4. Q4

   (a) 4a

   Answer Given

$$u_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

we show that these two column vectors form an orthonormal basis for $col(A)$ first by showing orthonormality then by showing the spanning condition. $u_1 \perp_{nor} u_2$ if $u_1 \cdot u_2 = 0, i = j, u_1 \cdot u_2 = 1, i \neq j$, for vector indices i, j. So,

$$u_1 \cdot u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0.$$

Hence the two vectors are orthonormal. Next for the spanning condition, note

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & -1 \end{bmatrix}.$$

Vectors u_1, u_2 are in the column space of A if there exists an x_1, x_2 such that

$$Ax = u,$$

for both instances of u. Then for u_1,

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & -1 \end{bmatrix}^T \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix} = \sqrt{5} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and for u_2,

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

meaning the condition is satisfied for both instances of u.

   (b) Compute the orthogonal projection, $\hat{b}$, of b onto col(A).

Answer Note the definition of the orthogonal projection $\hat{b}$ of b onto Col(A),

$$\hat{b} = \text{proj}_b\text{Col}(A) = \frac{b\text{Col}(A)}{b \cdot b}b = \frac{ba_1}{b \cdot b}b + \ldots + \frac{ba_m}{b \cdot b}b,$$

for a matrix A spanning m columns. There are two vectors that span the column space of A, so the procedure will be done for each column of A, a_1 and a_2:

$$\hat{b} = \frac{1}{5}\begin{bmatrix}1 & 2 & 0\end{bmatrix}\begin{bmatrix}4\\1\\2\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix} + \begin{bmatrix}0 & 0 & 1\end{bmatrix}\begin{bmatrix}4\\1\\2\end{bmatrix}\begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}\frac{6}{5}\\\frac{12}{5}\\2\end{bmatrix}$$

(c) Compute the solution to $A\hat{x} = \hat{b}$ by row reduction on the augmented matrix.

Answer

$$\begin{bmatrix}1 & 1 & \frac{6}{5}\\2 & 2 & \frac{12}{5}\\0 & -1 & 2\end{bmatrix} \sim \begin{bmatrix}1 & 1 & \frac{6}{5}\\2 & 2 & \frac{12}{5}\\0 & -1 & 2\end{bmatrix} \sim \begin{bmatrix}1 & 0 & \frac{16}{5}\\0 & 1 & -2\\0 & 0 & 0\end{bmatrix}.$$

Hence, $\hat{x} = \begin{bmatrix}\frac{16}{5}\\-2\\0\end{bmatrix}$.

(d) Compute the residual $r = b - A\hat{x}$ and verify it is orthogonal to col(A).

Answer

$$r = b - A\hat{x} = \begin{bmatrix}4\\1\\2\end{bmatrix} - \begin{bmatrix}1 & 1\\2 & 2\\0 & -1\end{bmatrix}\begin{bmatrix}\frac{16}{5}\\-2\end{bmatrix} = \begin{bmatrix}\frac{14}{5}\\\frac{-7}{5}\\0\end{bmatrix}.$$

Then to show orthogonality,

$$A^Tr = \begin{bmatrix}1 & 1\\2 & 2\\0 & -1\end{bmatrix}^T\begin{bmatrix}\frac{14}{5}\\\frac{-7}{5}\\0\end{bmatrix} = \mathbf{0}$$

(e) Verify this $\hat{x}$ is also a solution to the normal equations by working out the normal equations and showing that $\hat{x}$ satisfies $A^TAx = A^Tb$.

Answer Note $\hat{x} = \begin{bmatrix}\frac{16}{5}\\-2\\0\end{bmatrix}$. Then,

$$\begin{bmatrix}1 & 1\\2 & 2\\0 & -1\end{bmatrix}^T\begin{bmatrix}1 & 1\\2 & 2\\0 & -1\end{bmatrix}\begin{bmatrix}\frac{14}{5}\\\frac{-7}{5}\\0\end{bmatrix} = \begin{bmatrix}1 & 1\\2 & 2\\0 & -1\end{bmatrix}^T\begin{bmatrix}4\\1\\2\end{bmatrix} \implies \begin{bmatrix}5 & 5\\5 & 6\end{bmatrix}\begin{bmatrix}\frac{14}{5}\\\frac{-7}{5}\\0\end{bmatrix} = A^Tb$$

5. Q5

(a) Verify that each of these matrices satisfies the 2 conditions for being orthogonal projectors.

Answer A square matrix B is an orthogonal projector, if $B^T = B$ and $B^2 = B$. Verifying the first condition,

$$B = Q_1Q_1^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B^T = Q_2Q_2^T = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}^T = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

Note that both matrix products equality to the identity matrix with respective dimensions 2 and 1, and this implies that the first equality is true. Next showing the second condition,

$$B^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B^2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}^2,$$

and note that,

$$B^2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

(b) Give the rank of each of the matrices B.

Answer B_1 is full rank with rank 2. Each column is linearly independent. B_2 has rank 1 because each column is equivalent and there is total linear dependence among the columns.

(c) Now describe geometrically where each of these projects (that is, if v is any arbitrary vector in $\mathbb{R}^3$, where will BV live?).

Answer Note that given B_1 as described above and an arbitrary vector $x = [x_1, x_2, x_3]^T$, the transformation will leave x_1 and x_3 dimensions unchanged, however, the x_2 dimension is reduced to 0.

In terms of B_2, each of the component of the matrix is equal to one third, so the transformation B_2 maps each of the original coordinates to the simple average between the three values of the x vector, i.e. the new vector after the transformation spans $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]^T$.

6. Q6

The claim is that the orthogonal projection $\hat{b}$ previously computed is equivalent to UU^Tb, or that UU^T is a transformation equivalent to an orthogonal projection.

Proof WTS that UU^T is an orthogonal projection transformation.

Note that U is full column rank, which means that the column spaces for UU^T and U are equivalent. If UU^T is the orthogonal projection transformation, it must project onto the column space of U, I.e.,

$$(UU^T)U = U \implies U^T(UU^T)U = U^TU \implies U^TU = I.$$

This means that the columns of U^TU (and more prudently U as proved earier) are orthonormal. As such, UU^T is an orthogonal projection projection matrix it must an orthogonal projector. $\square$

7. Q7

(a) Show the normal equations of the matrix A^TA are symmetric.

Proof WTS $(A^TA)^T = A^TA$. Note by expansion,

$$(A^TA)^T = A^T(A^T)^T = A^TA,$$

with middle equality due to matrix-product associativity.

$\square$

(b) Show A^TA is positive semi-definite.

Proof WTS $\forall y \in \mathbb{R}^m, y^T(A^TA)y \geq 0$. Let $y \in \mathbb{R}^m$. Then,

$$y^T(A^TA)y = (y^TA^T)Ay = (Ay)^TAy = \|Ay\|_2^2,$$

with second to last equality from transpose products. Then, because the square of the Euclidean distance is always greater than zero and zero when $Ay = 0$ (i.e. positive semi-definite), and we just showed equality between the two norm and $y^T(A^TA)y$, $y^T(A^TA)y$ must also be positive semi-definite.

$\square$

(c) Show A^TA is positive definite if rank(A) is m.

Proof If A has rank m, then that means it has at least full column rank and its columns are all linearly independent. Then, Ay is a linear combination of the columns of A, and $y \neq 0 \implies Ay \neq 0$, as linear combination produces a non trivial solution. Therefore for $y \neq 0$,

$$y^TA^TAy = (Ay)^TAy = y^ty = \|y\|_2 > 0,$$

i.e. y^TA^TAy is positive definite.

$\square$

Acknowledgements: I worked with Lulu Rajguru and Luca Grossman on all problems.