Problem Set 5
Math 146 Spring 2023
Due: Friday, March 24 , 11:59 PM

Note: Numbers 1-8 are "pen and paper" problem for which you should show work. Number 9 is a Matlab problem that requires code submissions. All answers should be submitted in a write-up.

1. (34 points) Consider the following linear system $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} \tag{1}$$

   (a) By performing the Gaussian Elimination, show that the system does not have a solution.
   (b) Construct the normal system and find the least-squares solution $\mathbf{x}_{\text{lsq}}$.
   (c) Compute the residual vector $\mathbf{r} = \mathbf{b} - A\mathbf{x}_{\text{lsq}}$.
   (d) Calculate the relative size of the residual,

$$\epsilon_{\text{rel}} = \frac{\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2}{\|\mathbf{b}\|_2} \tag{2}$$

   (e) Show that the residual is orthogonal to each column vector of $A = [\mathbf{a}_1, \mathbf{a}_2]$.
   (f) Compute $\|A\mathbf{x}_{\text{lsq}}\|_2^2$ and $\|\mathbf{b}\|_2^2$ and show

$$\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2^2 + \|A\mathbf{x}_{\text{lsq}}\|_2^2 = \|\mathbf{b}\|_2^2. \tag{3}$$

   (g) With a picture, explain the geometric meaning of the relation (3). Hint: The relation has the same form as the Pythagorean theorem.

2. (13 points) Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}. \tag{4}$$

   (a) Show that $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3\}$ forms an orthogonal basis for $\mathbb{R}^3$.
   (b) By normalizing $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, construct an orthonormal basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \mathbf{q}_3\}$ for $\mathbb{R}^3$.

3. (22 points) Consider the following basis for $\mathbb{R}^2$:

$$\mathbf{a}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \tag{5}$$

   (a) Use the Gram Schmidt orthogonalization process to find the set $\{\mathbf{q}_1, \mathbf{q}_2\}$ that forms an orthonormal basis for $\mathbb{R}^2$:

(b) By plotting $\mathbf{q}_1$ and $\mathbf{q}_2$ in the $\mathbb{R}^2$ plane, explain how geometrically that $\{\mathbf{q}_1,\ \mathbf{q}_2\}$ forms an orthonormal basis for $\mathbb{R}^2$.

(c) For $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, compute its coordinates with respect to the basis $\{\mathbf{v}_1,\ \mathbf{v}_2\}$. In other words, compute $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ satisfying $\mathbf{w} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2$.

4. (10 points) Let $\{\mathbf{q}_1,\ \mathbf{q}_2\}$ be an orthonormal basis for $\mathbb{R}^2$. Consider $\mathbf{w} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2$.

   (a) Express $\mathbf{w}$ in terms of the matrix $Q = [\mathbf{q}_1\ \mathbf{q}_2]$ and vector $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$.

   (b) By using the result of (a) and the fact that $Q^T Q = QQ^T = I$, express α in terms of Q^T and $\mathbf{w}$.

5. (10 points) A square matrix Q is orthogonal if $Q^\intercal = Q^{-1}$. Show that the condition number (in the 2-norm) of an orthogonal matrix is 1. Note: This helps to explain why, unlike in using the normal equations, the conditioning of the linear problem is not worsened by using QR Factorization.

6. (10 points) Explain what may happen during the course of using the Gram-Schmidt process to solve the least squares problem if the matrix A is rank deficient (i.e. if the columns are not linearly independent). What kind of fix could you put in your code to remedy this?

7. (12 points) In class, we claimed that $P = \widehat{Q}\widehat{Q}^\intercal$ is an orthogonal projector, where $\widehat{Q}$ is defined as the matrix whose columns, $\{\mathbf{q}_1, ...\mathbf{q}_n\}$ ($\mathbf{q}_i \in \mathbb{R}^m$) give an orthonormal basis for the range of matrix A. Prove this fact about P.

8. (12 points) In class, we defined a reflection defined by $F = I - 2P_v$, where P_v is the orthogonal projection onto vector $\mathbf{v}$. Prove that F is an orthogonal matrix.

9. (a) (25 points) Write functions to perform the reduced QR factorization using the classical Gram-Schmidt algorithm and the modified Gram-Schmidt algorithm. Your code should take as input $m \times n$ matrix A with linearly independent column vectors (Note: this means you do NOT need to include your fix from Number (6)) and return matrices $\widehat{Q}$ and $\widehat{R}$. Also write functions to use these factorizations to solve the least squares problem. These functions should take as input matrices A and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your codes using a random matrix of size 50×10, and explain how you validated them. Note: Matlab's backslash performs least squares if $m > n$ for input matrix A.

   20 Bonus points for making a version of the Classical Gram-Schmidt least squares function that includes your fix from Number (6), validating it, and explaining how you

validated it. (Note: If you are checking if something is equal to 0, since there is numerical error, you may instead want to check if it is below some tolerance.)

(b) (12 points) Write a function to solve the least squares problem using the Householder method. For this you may want to simply write one function, taking as input the $m \times n$ matrix A and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your code and explain how you validated it.

(c) (25 points) Find an 11th degree polynomial that approximates the function $\cos(4x)$ evaluated at 50 equally spaced points on the interval $[0, 1]$ by solving the discrete least squares problem. You will be finding the 12 coefficients of $p(x) = c_0 + c_1x + ... + c_{11}x^{11}$.

The Vandermonde matrix that arises in the discrete problem is 50×12. Set up the linear leqst squares problem for the coefficients, and solve it using the methods below:

i. Form the normal equations and solve them.
ii. Use a QR factorization generated by classical Gram-Schmidt.
iii. Use a QR factorization generated by modified Gram-Schmidt.
iv. Use a QR factorization generated by Householder reflectors.
v. Use Matlab's least squares solver by `A\b`. This is also based on QR, but it includes pivoting for added stability.

For each method, give the 2-norm of the residual. Additionally, estimate what the relative error of the coefficients will be using condition numbers of the appropriate matrices. (You can skip this step for the Matlab function.) Give your results in a table, with one row for each method.

For reference, using quad precision (128 bits vs the 53 bits of double precision), this problem can be solved with a residual norm of `7.999154576455076e-09`, giving the coefficients provided in the table below, as well as in the file coefficients.mat.

c_0	1.000000000996606e+00
c_1	-4.227430949815150e-07
c_2	-7.999981235683346e+00
c_3	-3.187632625738558e-04
c_4	1.066943079610163e+01
c_5	-1.382028878048870e-02
c_6	-5.647075625417684e+00
c_7	-7.531602738192263e-02
c_8	1.693606966623459e+00
c_9	6.032106743884792e-03
c_{10}	-3.742417027133638e-01
c_{11}	8.804057595513443e-02

(d) (15 points) Comment on the performance of each method. Things to keep in mind/discuss: In class we discussed that the classical Gram-Schmidt method is unstable. Additionally, the Householder method is more stable than the modified Gram-Schmidt method. Lastly, using the condition numbers only allows you to *predict/estimate* the relative error of your coefficients. (Is there one error prediction that you should not believe? Which one and why?)