Math 2940 Worksheet

Least Squares, SVD

Week 14

December 5th, 2019

This worksheet covers material from **Sections 6.4, 6.5,** and **7.4**. Please work in collaboration with your classmates to complete the following exercises - this means sharing ideas and asking each other questions.

Question 1.

(a) Apply the Gram-Schmidt process to the columns of $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}$

(b) Use your answer from part (a) to compute a QR factorization of A.

(c) Use your answer from part (b) to find a least-squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ for $\boldsymbol{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$.

Question 2. Suppose $A = QR$, where Q is an $m \times n$ matrix with orthogonal columns and R is an $n \times n$ matrix. Show that if the columns of A are linearly dependent, then R cannot be invertible.

Question 3. Let $A = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 5 \\ -3 \\ -3 \end{bmatrix}$.
Find a least-squares solution of $A\boldsymbol{x} = \boldsymbol{b}$, and compute the associated least-squares error.

Question 4.
Consider the set of vectors with unit length in $\mathbb{R}^2$, denoted by $D = \{\boldsymbol{v} \in \mathbb{R}^2 \ : \ \|\boldsymbol{v}\| = 1\}$. Describe the image of D under the linear transformation

$$T(\boldsymbol{v}) = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \boldsymbol{v}, \qquad \boldsymbol{v} \in D$$

Sketch the set D and its image $T(D)$. [*Hint*: use the SVD of the matrix of T.]

Answer to Question 1.

(a) We want to take the vectors $\boldsymbol{x}_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}$ and $\boldsymbol{x}_2 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix}$, and convert them into orthogonal vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

For the first vector, we choose

$$\boldsymbol{v}_1 = \boldsymbol{x}_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}$$

To make sure that $\boldsymbol{v}_2$ is orthogonal to $\boldsymbol{v}_1$, we want to take $\boldsymbol{x}_2$ and subtract the component in the $\boldsymbol{v}_1$ direction.

$$\begin{aligned} \boldsymbol{v}_2 = \boldsymbol{x}_2 - \text{proj}_{\boldsymbol{v}_1}(\boldsymbol{x}_2) = \boldsymbol{x}_2 - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \\ = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - \frac{\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix}}{5^2 + 1^2 + (-3)^2 + 1^2} = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - \frac{72}{36}\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix} \end{aligned}$$

Now that we have orthogonal vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, we can find their magnitudes:

$$\begin{aligned} \|\boldsymbol{v}_1\| &= \sqrt{5^2 + 1^2 + (-3)^2 + 1^2} = \sqrt{36} = 6 \\ \|\boldsymbol{v}_2\| &= \sqrt{(-1)^2 + 5^2 + 1^2 + 3^2} = \sqrt{36} = 6 \end{aligned}$$

So the orthonormal basis is

$$\boxed{\begin{aligned} \boldsymbol{u}_1 = \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} = \begin{bmatrix} 5/6 \\ 1/6 \\ -3/6 \\ 1/6 \end{bmatrix} \\ \boldsymbol{u}_2 = \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} = \begin{bmatrix} -1/6 \\ 5/6 \\ 1/6 \\ 3/6 \end{bmatrix} \end{aligned}}$$

(b) To find Q, we can just combine the colums we found in part (a)

$$Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$$

To find R, we note that

$$A = QR$$

Multiplying both sides on the left by Q^T,

$$Q^T A = Q^T QR = IR = R$$

So we can compute

$$R = Q^T A = \begin{bmatrix} 5/6 & 1/6 & -3/6 & 1/6 \\ -1/6 & 5/6 & 1/6 & 3/6 \end{bmatrix} \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$$

Therefore the QR factorization of A is:

$$\boxed{A = QR = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix} \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}}$$

(c) First we note that the least-squares system is equivalent to solving

$$A\boldsymbol{x} = QQ^T\boldsymbol{b}$$

Using the QR factorization,

$$QR\boldsymbol{x} = QQ^T\boldsymbol{b}$$

Multiplying both sides on the left by Q^T,

$$\begin{aligned} Q^T QR\boldsymbol{x} &= Q^T QQ^T\boldsymbol{b} \\ R\boldsymbol{x} &= Q^T\boldsymbol{b} \end{aligned}$$

First we compute

$$Q^T\boldsymbol{b} = \begin{bmatrix} 5/6 & 1/6 & -3/6 & 1/6 \\ -1/6 & 5/6 & 1/6 & 3/6 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{-5+2-3+6}{6} \\ \frac{1+10+1+18}{6} \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$

So we want to solve the system

$$\begin{aligned} R\boldsymbol{x} &= Q^T\boldsymbol{b} \\ \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix} \boldsymbol{x} &= \begin{bmatrix} 0 \\ 5 \end{bmatrix} \end{aligned}$$

Setting up the augmented matrix and row reducing,

$$\left[\begin{array}{cc|c} 6 & 12 & 0 \\ 0 & 6 & 5 \end{array}\right] \xrightarrow{\text{R1}-2\cdot\text{R2}} \left[\begin{array}{cc|c} 6 & 0 & -10 \\ 0 & 6 & 5 \end{array}\right] \xrightarrow{\frac{1}{6}\cdot\text{R2}} \left[\begin{array}{cc|c} 6 & 0 & -10 \\ 0 & 1 & 5/6 \end{array}\right] \xrightarrow{\frac{1}{6}\cdot\text{R1}} \left[\begin{array}{cc|c} 1 & 0 & -10/6 \\ 0 & 1 & 5/6 \end{array}\right]$$

Therefore the least-squares solution is $\boxed{\begin{bmatrix} -10/6 \\ 5/6 \end{bmatrix}}$.

Answer to Question 2. First, we know that

$$A = QR$$

Multiplying both sides on the left by Q^T,

$$Q^T A = Q^T Q R = IR = R$$

Since the columns of A are linearly dependent, we know that there exists a non-zero vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \mathbf{0}$. However, this means

$$R\boldsymbol{x} = Q^T A\boldsymbol{x} = Q^T(A\boldsymbol{x}) = Q^T\mathbf{0} = \mathbf{0}$$

Since there is a non-zero vector $\boldsymbol{x}$ such that $R\boldsymbol{x} = \mathbf{0}$, this implies that $\boxed{R \text{ is not invertible.}}$

Answer to Question 3. We can find least-squares solutions as solutions of the system $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$.
We can compute

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

and

$$A^T\boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

Setting up the augmented matrix for $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ and row reducing,

$$\left[\begin{array}{ccc|c} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{array}\right] \xrightarrow{\frac{1}{3}\cdot\text{R1}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{array}\right] \xrightarrow{\frac{1}{14}\cdot\text{R3}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 9 & 83 & 28 & -65 \\ 0 & 2 & 1 & -2 \end{array}\right]$$

$$\xrightarrow{\text{R2}-9\cdot\text{R1}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 2 & 1 & -2 \end{array}\right] \xrightarrow{\frac{1}{28}\cdot\text{R2}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 2 & 1 & -2 \\ 0 & 2 & 1 & -2 \end{array}\right] \xrightarrow{\text{R3}-\text{R2}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 2 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{2}\cdot\text{R2}} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\text{R1}-3\cdot\text{R2}} \left[\begin{array}{ccc|c} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since there is a free variable, we see that there are actually inifinitely many least-squares solutions to this problem.

To choose a specific one, we can set $x_3 = 0$ to get a least squares solution $\boxed{\widehat{\boldsymbol{x}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}}$.

We can also check that $A\widehat{\boldsymbol{x}} = \boldsymbol{b}$ exactly, so the associated least-squares error is zero.

Answer to Question 4. First we want to compute the SVD of the matrix $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$.

We compute

$$A^T A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

To find the singular values of A, we can compute the eigenvalues of $A^T A$ as follows

$$\begin{aligned} \det(A^T A - \lambda I) = \det \left(\begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix} \right) &= (17-\lambda)^2 - 8^2 = 0 \\ (17-\lambda)^2 &= 8^2 \\ 17 - \lambda &= \pm 8 \\ \lambda &= 25, 9 \end{aligned}$$

So the singular values are $\sigma_1 = \sqrt{25} = 5$ and $\sigma_2 = \sqrt{9} = 3$, and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$.

To find the right singular vectors, we want to find the eigenvectors of $A^T A$. For the first eigenvector,

$$A - 25I = \begin{bmatrix} -8 & 8 \\ 8 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

Therefore the first right singular vector is proportional to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Normalizing, we get

$$\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

For the second eigenvector,

$$A - 9I = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Therefore the first right singular vector is proportional to $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Normalizing, we get

$$\boldsymbol{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

and therefore $V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$.

We can compute the first two singular vectors as

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1} A \boldsymbol{v}_1 = \frac{1}{5} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \boldsymbol{u}_2 \quad = \frac{1}{\sigma_2} A \boldsymbol{v}_2 = \frac{1}{3} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3\sqrt{2} \\ -1/3\sqrt{2} \\ 4/3\sqrt{2} \end{bmatrix}$$

For the third singular vector, we need to find a vector orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. This is equivalent to the system of equations

$$\left[\begin{array}{ccc|c} 1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 4 & 0 \\ 1 & 1 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 4 & 0 \\ 0 & 2 & -4 & 0 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right]$$

So we see that $x_1 = -2x_3$ and $x_2 = 2x_3$. Setting $x_3 = 1$, we see that $\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ is orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Normalizing,

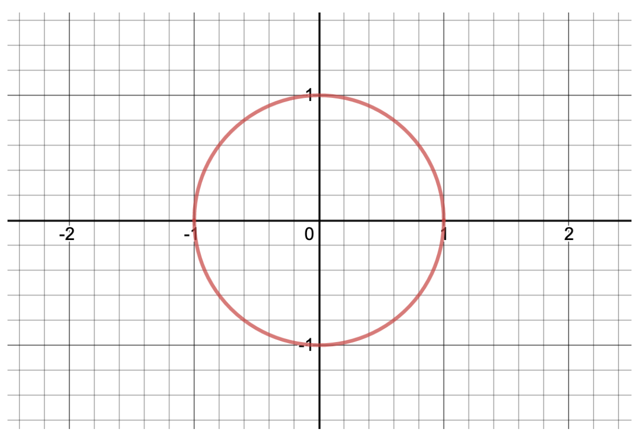
$$\boldsymbol{u}_3 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

And therefore $U = \begin{bmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & -2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{bmatrix}$.

So the singular value decomposition is

$$\boxed{A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & -2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}$$

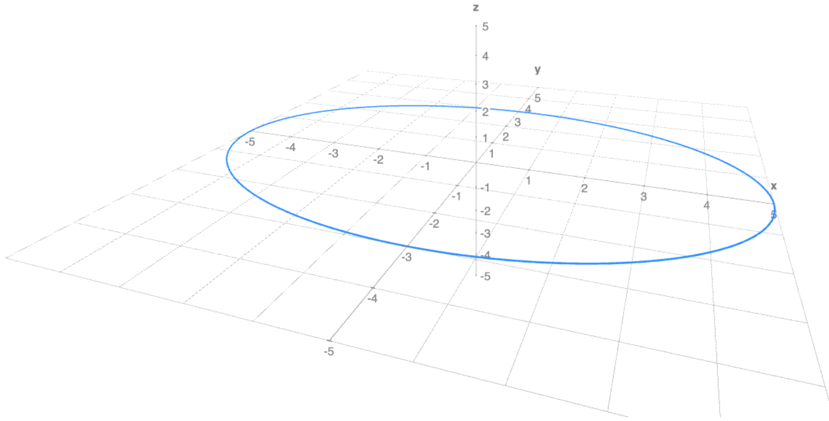
We recall that the set $D = \{\boldsymbol{v} \in \mathbb{R}^2 : \|\boldsymbol{v}\| = 1\}$, which is just the unit circle in $\mathbb{R}^2$:



To determine what the transformation T does to the set D, we will figure out what each individual part of the SVD of T does to the set D.

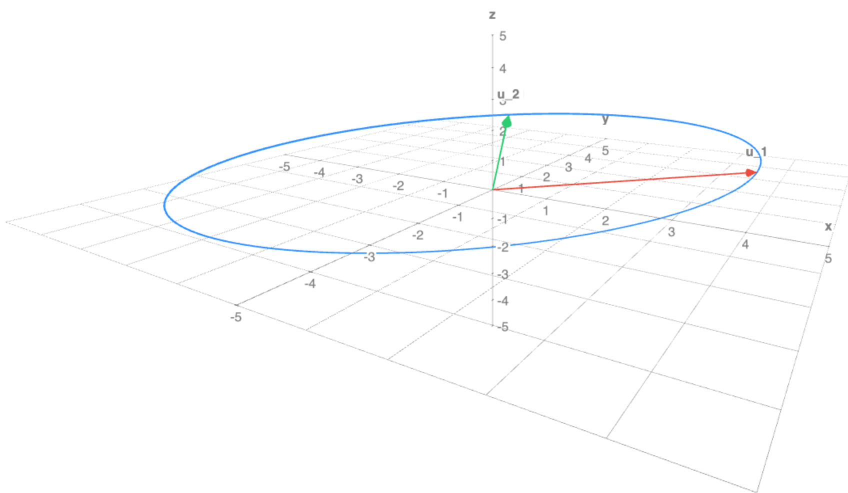
First, since $V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ corresponds to a rotation, this won't change the set D.

Then, $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$ will transform D from a circle in $\mathbb{R}^2$ to an ellipse in $\mathbb{R}^3$. The ellipse will actually lie entirely in the xy-plane, with a major axis of $\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$ and a minor axis of $\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$:



You can find an interactive version of this plot at https://www.math3d.org/6PgGSXbQ

Finally, $U = \begin{bmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & -2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{bmatrix}$ will rotate the major axis to the vector $\boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$, and rotate the minor axis to the vector $\boldsymbol{u}_1 = \begin{bmatrix} 1/3\sqrt{2} \\ -1/3\sqrt{2} \\ 4/3\sqrt{2} \end{bmatrix}$:



You can find an interactive version of this plot at https://www.math3d.org/Vq5qsBSx