Math 241: Final Exam Review

Practice Problems Solutions

First: problems from Exam 1 review, Exam 2 review, and both the midterms, as well as homework.

For material presented since that time:

1. The vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}$ are orthogonal in $\mathbb{R}^3$. Find a third vector $\mathbf{v}_3$ that is orthogonal to both $\mathbf{v}_1, \mathbf{v}_2$, such that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ form a basis for $\mathbb{R}^3$.

   Solution: $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, or any multiple of this vector.

2. Suppose $\mathbf{v} \in \operatorname{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$, where the $\mathbf{u}_i$ are orthonormal. Prove that $\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k$.

   Solution: Since $\mathbf{v} \in \operatorname{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$, it must be the case that there exists $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k$.
   Now, for $1 \leq i \leq k$, note that

   $$\mathbf{v} \cdot \mathbf{u}_i = (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k) \cdot \mathbf{u}_i = \alpha_1\mathbf{u}_1 \cdot \mathbf{u}_i + \alpha_2\mathbf{u}_2 \cdot \mathbf{u}_i + \cdots + \alpha_k\mathbf{u}_k \cdot \mathbf{u}_i.$$

   Moreover, since the $\mathbf{u}$ vectors are orthonormal, $\mathbf{u}_j \cdot \mathbf{u}_i = 0$ for $i \neq j$ and $= 1$ for $i = j$. Hence,

   $$\mathbf{v} \cdot \mathbf{u}_i = \alpha_i.$$

   Plugging this in to the original expression for $\mathbf{v}$ yields the result.

3. Let $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \mid y - 2z = 0 \right\}$. Let $\mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}$. Calculate $\operatorname{Proj}_W \mathbf{v}$.

   Solution: $\operatorname{Proj}_W \mathbf{v} = \begin{bmatrix} 2 \\ 18/5 \\ 9/5 \end{bmatrix}$.

4. Let $\mathbf{v}_1, \mathbf{v}_2$ be vectors in a vector space V, and let W be a subspace of V. Suppose that $\operatorname{Proj}_W \mathbf{v}_1 = \operatorname{Proj}_W \mathbf{v}_2$. Must $\mathbf{v}_1 = \mathbf{v}_2$?

   Solution: No. Suppose that $V = \mathbb{R}^3$, W is the xy-plane, and $\mathbf{v}_1, \mathbf{v}_2$ have the same x and y coordinates. Regardless of their z-coordinate, it must be the case that their projection onto W is the same.

5. Let $\mathbf{v}_1, \mathbf{v}_2$ be vectors in a vector space V, and let W be a subspace of V. Suppose that $\text{Proj}_W \mathbf{v}_1 = \text{Proj}_W \mathbf{v}_2$, and also $\text{Proj}_{W^\perp} \mathbf{v}_1 = \text{Proj}_{W^\perp} \mathbf{v}_2$. Must $\mathbf{v}_1 = \mathbf{v}_2$?

Solution: Yes. Recall that $\mathbf{v}_1$ can be uniquely written in the form $\mathbf{v}_1 = \mathbf{w}_a + \mathbf{w}_b$, where $\mathbf{w}_a \in W$, namely, $\mathbf{w}_a = \text{Proj}_W \mathbf{v}_1$, and $\mathbf{w}_b \in W^\perp$, and namely, $\mathbf{w}_b = \text{Proj}_{W^\perp} \mathbf{v}_1$. Likewise, $\mathbf{v}_2$ can be uniquely written in the from $\mathbf{v}_2 = \mathbf{w}'_a + \mathbf{w}'_b$, where $\mathbf{w}'_a \in W$ and $\mathbf{w}'_a = \text{Proj}_W \mathbf{v}_2$, and $\mathbf{w}'_b \in W^\perp$, and $\mathbf{w}'_b = \text{Proj}_{W^\perp} \mathbf{v}_2$. The hypothesis states that $\mathbf{w}_a = \mathbf{w}'_a$, and $\mathbf{w}_b = \mathbf{w}'_b$. Thus, $\mathbf{v}_1 = \mathbf{v}_2$.

6. Find a singular value decomposition of the matrix $A = \begin{bmatrix} 1 & 1 \\ -2 & 2 \\ -1 & -1 \end{bmatrix}$.

Solution: There is more than one possible answer. Here is one.

$$A = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -1 & 0 & 0 \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{8} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

7. Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & -2 & -3 & -4 \end{bmatrix}$. How many nonzero singular values does A have? How do you know?

Solution: A has precisely 1 nonzero singular value. This is because A is rank 1, since every column is a multiple of every other.

8. Show that every rank 1 matrix of dimension $m \times n$ can be uniquely represented as $c\mathbf{u}\mathbf{v}^T$, where $\mathbf{u}$ is a $m \times 1$ vector with $\|\mathbf{u}\| = 1$, and $\mathbf{v}$ is an $n \times 1$ vector with $\|\mathbf{v}\| = 1$.

Solution: Since A is rank 1, every column of A is a multiple of every other column of A. Let $\mathbf{a}_i$ be the first nonzero column of A, and let $\mathbf{v} = \frac{1}{\|\mathbf{a}_i\|}\mathbf{a}_i$, so that $\mathbf{a}_i = \|\mathbf{a}_i\|\mathbf{v}$. We put $\|\mathbf{a}_i\| = c_i$, and note that for every column in A, by setting $c_j = \|\mathbf{a}_j\|$, we obtain that $\mathbf{a}_j = c_j\mathbf{v}$. Note further that this representation is unique up to sign of $\mathbf{v}$; that is, the only such vectors $\mathbf{v}$ that could have been chosen are $\mathbf{v}$ or $-\mathbf{v}$. Hence we may write

$$\begin{aligned} A &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \\ &= \begin{bmatrix} c_1\mathbf{v} & c_2\mathbf{v} & \dots & c_n\mathbf{v} \end{bmatrix} \\ &= \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix} \mathbf{v} \end{aligned}$$

Define $\mathbf{w} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$, and define $c = \|\mathbf{w}\|$ and $\mathbf{u} = \frac{1}{c}\mathbf{w}$. Then we immediately have $A = c\mathbf{u}^T\mathbf{v}$ from the above calculation. Note that this representation is unique up to signs; that is, we may move a negative sign from one term to another, but as the selection of $\mathbf{v}$ was unique up to signs, each other vector is also unique up to signs.

9. Suppose U is a matrix with orthonormal columns. Show that $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ for which the product is defined.

   Solution: Write $U\mathbf{x} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$, where the x_i are the coordinates of $\mathbf{x}$ and the $\mathbf{u}_i$ are the columns of U. Then we have

$$\begin{aligned}\|U\mathbf{x}\|^2 &= (U\mathbf{x}) \cdot (U\mathbf{x}) \\ &= (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) \cdot (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) \\ &= x_1\mathbf{u}_1 \cdot (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) + x_2\mathbf{u}_2 \cdot (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) + \cdots + x_n\mathbf{u}_n \cdot (x_1\mathbf{u}_1 + \cdots + \\ &= x_1^2 + x_2^2 + \cdots + x_n^2 = \|x\|^2,\end{aligned}$$

   since $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ when $i \neq j$ and 1 when $i = j$.

10. A SVD for the matrix A is as follows (where we have rounded to 2 decimal places for convenience):

$$A = \begin{bmatrix} .40 & -.78 & .47 \\ .37 & -.33 & -.87 \\ -.84 & -.52 & -.16 \end{bmatrix} \begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} .3 & -.51 & -.81 \\ .76 & .64 & -.12 \\ .58 & -.58 & .58 \end{bmatrix}$$

   Using this decomposition for A, answer the following (without doing ANY arithmetic):

   (a) What is the rank of A?

   Solution: 2

   (b) Find a basis for $\operatorname{Col} A$. Find a basis for $\operatorname{Null} A$.

   Solution: A basis for $\operatorname{Col} A$ is $\left\{ \begin{bmatrix} .4 \\ .37 \\ -.84 \end{bmatrix}, \begin{bmatrix} -.78 \\ -.33 \\ -.52 \end{bmatrix} \right\}$. A basis for $\operatorname{Null} A$ is $\left\{ \begin{bmatrix} .58 \\ -.58 \\ .58 \end{bmatrix} \right\}$.

   (c) Find a unit vector $\mathbf{v}$ so that $\|A\mathbf{v}\|$ is maximal.

   Solution: $\mathbf{v} = \begin{bmatrix} .3 \\ -.51 \\ -.81 \end{bmatrix}$

11. Suppose you have a collection of data $\{(t_i, y_i) \mid 1 \leq i \leq N\}$, where each $t_i, y_i \in \mathbb{R}$, and you would like to approximate these data with a function that takes the form $y \approx f(t) = at^2 + bt + c$. How would you set up a least squares problem to accomplish your goal?

Solution: For each data point, we obtain the desired equation $y_i = at_i^2 + bt_i + c$. We then form a system of equation $A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{y}$, where each row of A takes the form $[t_i^2, t_i, 1]$ and each coordinate of $\mathbf{y}$ is exactly y_i. We can then solve the least squares problem to obtain optimal approximate values for a, b, c.

12. Suppose you wish to approximate a collection of data $\{(t_i, y_i) \mid 1 \leq i \leq N\}$, where each $t_i, y_i \in \mathbb{R}$, using least squares, with a horizontal line $f(t) = c$. What is c? Show that your answer is correct.

Solution: For each data point, we obtain the equation $y_i = c$. Hence, our least squares problem takes the form

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [c] = \mathbf{y},$$

where $\mathbf{y}$ contains the values of y_i. Note that $A^T A = [N]$, so $(A^T A)^{-1} = [1/N]$, where both of these represent 1×1 matrices. Moreover, $A^T \mathbf{y} = \sum_{i=1}^N y_i$. Hence, $\hat{c} = \frac{1}{N} \sum_{i=1}^N y_i$, the mean of the y_i.

13. The lines $\mathcal{L}_1 = \left\{ \begin{bmatrix} x \\ x \\ x \end{bmatrix} \in \mathbb{R}^3 \right\}$ and $\mathcal{L}_2 = \left\{ \begin{bmatrix} y \\ 3y \\ -1 \end{bmatrix} \in \mathbb{R}^3 \right\}$ do not intersect. Write and solve a least squares problem to find the shortest line segment between these two lines.

Solution: We wish to find a choice of x and y so that $\left\| \begin{bmatrix} x \\ x \\ x \end{bmatrix} - \begin{bmatrix} y \\ 3y \\ -1 \end{bmatrix} \right\|$ is minimal. Writing this as a matrix equation, we wish to find a choice of $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ so that

$$\left\| \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 0 \end{bmatrix} \mathbf{x} - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \right\|$$

is minimal.

Using least squares, and taking $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 0 \end{bmatrix}$ we know that this is minimized when

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{5}{7} \\ \frac{2}{7} \end{bmatrix}$$

.

Thus, the line segment of shortest distance between the two lines is between the points $\begin{bmatrix} -\frac{5}{7} \\ -\frac{5}{7} \\ -\frac{5}{7} \end{bmatrix} \in \mathcal{L}_1$ and $\begin{bmatrix} \frac{2}{7} \\ \frac{6}{7} \\ -1 \end{bmatrix} \in \mathcal{L}_2$.

14. Why is A^TA noninvertible when A has linearly dependent columns?

Solution: When A has linearly dependent columns, it has rank less than n, so it has at most $n-1$ nonzero singular values. Note that A^TA is an $n \times n$ matrix whose nonzero eigenvalues are the squares of the nonzero singular values of A. Since there are at most $n-1$ nonzero singular values, it must be the case that any other eigenvalues of A^TA are 0, and hence A^TA is noninvertible, as 0 is not possible as an eigenvalue of an invertible matrix.

15. A university registrar keeps track of class attendance, measured as a percentage (so 100 means full attendance, and 30 means 30% of the students attend). For each class lecture, she records the attendance y, and several features:

- x_1 is the day of week, with Monday coded as 1 and Friday coded as 5.
- x_2 is the week of the quarter, coded as 1 for the first week, and 10 for the last week.
- x_3 is the hour of the lecture, with 8AM coded as 8, and 4PM coded as 16.
- $x_4 = \max\{T - 80, 0\}$, where T is the outside temperature (so x_4 is the number of degrees above 80°F ≈ 26.5°C).
- $x_5 = \max\{50 - T, 0\}$, where T is the outside temperature (so x_5 is the number of degrees below 50°F ≈ 10°C).

(These features were suggested by a professor who is an expert in the theory of class attendance.) A 241 student carefully fits the data with the following least squares regression model,

$$\hat{y} = -1.4x_1 - 0.3x_2 + 1.1x_3 - 0.6x_4 - 0.5x_5 + 68.2,$$

and validates it properly. Give a short story/explanation, in English, of this model.

Solution: Answers, of course, may vary. It appears that the most predictive piece of data for attendance is the day of the week, with attendance declining throughout the week. Later weeks in the quarter have lower attendance, generally, as well. The second most predictive feature is the time of the class. If the class is earlier in the day, attendance is less frequent than later classes. Finally, overly warm and overly cold temperatures are both negatively associated with attendance (presumably students find something else to do on warm days, and stay in bed on cold ones!)

16. Carefully prove, in the case that A is an $m \times 2$ matrix with linearly independent columns, that $A(A^TA)^{-1}A^T\mathbf{b} = \text{Proj}_{\text{Col}\,A}\,\mathbf{b}$.

Solution: Write the columns of A as $\mathbf{a}_1$ and $\mathbf{a}_2$. Using Gram-Schmidt, we have that we can obtain an orthogonal basis for the columns of A as

$$\mathbf{v}_1 = \mathbf{a}_1, \quad \mathbf{v}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_1 \cdot \mathbf{a}_2}{\mathbf{a}_1 \cdot \mathbf{a}_1}\mathbf{a}_1.$$

Normalizing this basis, we obtain an orthonormal basis for the columns of A as

$$\mathbf{q}_1 = \frac{1}{\mathbf{a}_1 \cdot \mathbf{a}_1}\mathbf{a}_1, \quad , \mathbf{q}_2 = \frac{\mathbf{a}_1 \cdot \mathbf{a}_1}{c}\mathbf{a}_2 - \frac{\mathbf{a}_1 \cdot \mathbf{a}_2}{c}\mathbf{a}_1,$$

where $c = (\mathbf{a}_1 \cdot \mathbf{a}_1)(\mathbf{a}_2 \cdot \mathbf{a}_2) - (\mathbf{a}_1 \cdot \mathbf{a}_2)^2$. We therefore can write

$$A = [\mathbf{q}_1 \quad \mathbf{q}_2] \begin{bmatrix} \frac{1}{\mathbf{a}_1 \cdot \mathbf{a}_1} & -\frac{\mathbf{a}_1 \cdot \mathbf{a}_2}{c} \\ 0 & \frac{\mathbf{a}_1 \cdot \mathbf{a}_1}{c} \end{bmatrix},$$

a QR-factorization of A. Now, recall that $\text{Proj}_{\text{Col}\,A}\,\mathbf{b} = (\mathbf{b} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{b} \cdot \mathbf{q}_2)\mathbf{q}_2$, and hence we have

$$\begin{aligned} \text{Proj}_{\text{Col}\,A}\,\mathbf{b} &= (\mathbf{b} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{b} \cdot \mathbf{q}_2)\mathbf{q}_2 \\ &= [\mathbf{q}_1 \quad \mathbf{q}_2] \begin{bmatrix} \mathbf{b} \cdot \mathbf{q}_1 \\ \mathbf{b} \cdot \mathbf{q}_2 \end{bmatrix} \\ &= Q \begin{matrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \end{matrix} \mathbf{b} \\ &= QQ^T\mathbf{b} \end{aligned}$$

On the other hand,

$$\begin{aligned} A(A^TA)^{-1}A^T\mathbf{b} &= QR(R^TQ^TQR)^{-1}R^TQ^T\mathbf{b} \\ &= QR(R^TR)^{-1}R^TQ^T\mathbf{b} \\ &= QRR^{-1}R^{-T}R^TQ^T\mathbf{b} \\ &= QQ^T\mathbf{b}. \end{aligned}$$

Hence, as desired, $A(A^TA)^{-1}A^T\mathbf{b} = \text{Proj}_{\text{Col}\,A}\,\mathbf{b}$.