Math 241: Final Exam Review

Key Concepts

- Understanding of vectors, matrices, and associated arithmetic, including properties of addition, scalar multiplication, matrix-vector products, matrix-matrix products, inverses, inner products, and norms
- Connections between systems of linear equations and matrix equations
- Spans and linear combinations, and how these relate to planes/lines/etc, and to solutions to systems of linear equations
- Vector spaces: what they are, how to determine if a set of vectors is a space or not
- Interpretations of matrix-vector multiplication: dot products, linear combinations of the columns of A, system of linear equations
- Null space of a matrix A: how to compute, relationship to the matrix equation $A\mathbf{x} = \mathbf{b}$
- Linear functions: what they are, how to determine if a function is linear. Associated function terms (injective, surjective, bijective, etc.)
- Definition of invertible linear function, matrix, understanding how to compute inverse of an $n \times n$ matrix A.
- Definitions of linearly dependent, independent, basis. Know how to construct a basis from a spanning set, how to change basis, and theorems surrounding bases, including dimension of a space, how to determine it, and related theory
- Rank-Nullity Theorem, for matrices and for linear transformations
- Determinant: Definition, how to compute, purpose
- Eigenvalues/eigenvectors: definition, how to compute, purpose. Includes: awareness and understanding of diagonalization
- Projection onto a subspace, including how to compute using Gram-Schmidt orthogonalization and using least squares
- Singular value decomposition: how to compute, how to express A as $U\Sigma V^T$ and what the parts of this decomposition mean. Understand what is meant by a "best rank k approximation" and how to get it from the SVD.
- Least Squares: understand how to calculate a least squares solution to an unsolvable system of equations, by QR-decomposition, pseudoinverse, and orthogonal projection. Understand how least squares can be used to fit functions to data, and basics of how to construct least squares problems.

Terms

- Vector (associated terms: elements, entries, length)
- Standard unit vector
- Linear combination (associated terms: affine combination, convex combination)
- System of linear equations
- Span, Spanning set (aka generating set)
- Homogeneous linear system (also non-homogeneous)
- Vector space, vector subspace
- Matrix (associated terms: dimension, elements, square, column, row, upper triangular, transpose)
- Column space
- Null space
- Linear function (aka linear transformation)
- Kernel
- Injective (aka one-to-one)
- Image (aka range)
- Surjective (aka onto)
- Bijective
- Inverse (for a function or a matrix)
- Linearly independent/dependent
- Basis
- Dimension
- Rank (of a matrix, or a set S)
- Nullity
- Determinant
- Eigenvalue, eigenvector, eigenspace
- Characteristic equation
- Algebraic multiplicity, geometric multiplicity
- Diagonalizable
- Similar
- Norm
- Inner product, associated norm ($\|v\| = \sqrt{\langle v, v \rangle}$)
- Orthogonal, orthonormal
- Projection onto a subpsace
- Singular value, right singular vector, left singular vector, singular value decomposition
- Least Squares problem, least squares solution
- Residual
- Time Series

Key Theorems

- If V is a vector space and $W \subseteq V$, then W is a vector subspace of V if and only if
  1. W contains $\mathbf{0}$
  2. W is closed under addition
  3. W is closed under scalar multiplication
- If V is a vector space and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$, then $\operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a vector subspace of V.
- Let A be an $m \times n$ matrix, and let $\mathbf{b} \in \mathbb{R}^m$. If $\mathbf{x}_1$ is a particular solution to the equation $A\mathbf{x} = \mathbf{b}$, then the set of solutions to the equation takes the form $\{\mathbf{x}_1 + \mathbf{y} \mid \mathbf{y} \in \operatorname{Null}(A)\}$.
- Let $f : V \to W$ be a linear function on vector spaces V and W. Then $\operatorname{Ker}(f)$ is a vector subpsace of V, and $\operatorname{Im}(f)$ is a vector subspace of W.
- Let $f : V \to W$ be a linear function on vector spaces V and W. Then f is injective if and only if $\operatorname{Ker}(f) = \{\mathbf{0}\}$.
- Let $f : V \to W$ be a linear function on vector spaces V and W. Then f is surjective if and only if $\operatorname{Im}(f) = W$.
- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear function, represented as $f(\mathbf{x}) = A\mathbf{x}$. Then $\operatorname{Ker}(f) = \operatorname{Null}(A)$, and $\operatorname{Im} f = \operatorname{Col}(A)$.
- The inverse of a linear function is linear.
- If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a spanning set for V, then there exists $\mathcal{B} \subseteq S$ that is a basis for V. Moreover, $\mathcal{B}$ can be found from S by iteratively eliminating vectors $\mathbf{v}_k \in S$ that can be written as a linear combination of vectors in $S \backslash \{\mathbf{v}_k\}$.
- If S is a generating set for a vector space V and $\mathcal{B}$ is a finite linearly independent set in V, then $|\mathcal{B}| \leq |S|$.
- If $\mathcal{B}_1$ and $\mathcal{B}_2$ are finite bases for a vector space V, then $|\mathcal{B}_1| = |\mathcal{B}_2|$.
- if S is a generating set for V, and $|S|$ is smallest among all possible generating sets for V, then S is a basis for V.
- If V is a finite-dimensional vector space, with $\dim V = n$, and $A \subseteq V$ is linearly independent, then there exists a basis $\mathcal{B}$ for V with $A \subseteq \mathcal{B}$.
- If V is a finite-dimensional vector space, with $\dim V = n$, and $A \subseteq V$ has $|A| = n$ and A is linearly independent, then A is a basis for V.
- (Rank Nullity Theorem v1) If $f : V \to W$ is a linear function, then

$$\dim \operatorname{Ker}(f) + \dim \operatorname{Im}(f) = \dim V$$

- (Rank Nullity Theorem v2) If A is an $m \times n$ matrix, then

$$\operatorname{rank} A + \operatorname{nullity} A = n.$$

- (Gram-Schmidt orthogonalization) Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a basis for a vector space V. Let $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$, where

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \operatorname{Proj}_{\mathbf{w}_1} \mathbf{v}_2 \\
&\cdots \\
\mathbf{w}_k &= \mathbf{v}_k - \operatorname{Proj}_{\mathbf{w}_1} \mathbf{v}_k - \operatorname{Proj}_{\mathbf{w}_2} \mathbf{v}_k - \cdots - \operatorname{Proj}_{\mathbf{w}_{k-1}} \mathbf{v}_k
\end{aligned}$$

  Then $\mathcal{C}$ is an orthogonal basis for V.

- If $\mathbf{w} = \operatorname{Proj}_W \mathbf{v}$ is the projection of $\mathbf{v}$ onto a subspace W, then $\|\mathbf{v} - \mathbf{w}\|$ is minimal among all elements of W.

- Let $A = U\Sigma V^T$ be a singular value decomposition for A, where U is $m \times m$, Σ is $m \times n$, and V is $n \times n$. Take the columns of V as $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and the nonzero singular values as $\sigma_1, \ldots, \sigma_k$. Then the vector space spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$ is the best dimension t approximation to the rows of A, in the sense that it minimizes

$$\sum_{i=1}^{m} \|\mathbf{a}_i - \operatorname{Proj}_W \mathbf{a}_i\|^2,$$

  where the $\mathbf{a}_i$ are the rows of A and the minimization is over all dimension t subspaces.

- Let A be a matrix with linearly independent columns. Then for any $\mathbf{b}$, the least squares solution $\hat{\mathbf{x}}$ to the equation $A\mathbf{x} = \mathbf{b}$ is given by $(A^TA)^{-1}A^T\mathbf{b}$.

- Let A be a matrix with linearly independent columns. Then for any $\mathbf{b}$, the least squares solution $\hat{\mathbf{x}}$ to the equation $A\mathbf{x} = \mathbf{b}$ satisfies $A\hat{\mathbf{x}} = \operatorname{Proj}_{\operatorname{Col} A} \mathbf{b}$.

Practice Problems

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 $\mathbb{R}^3$.

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

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 $\text{Proj}_W \mathbf{v}$.

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 $\text{Proj}_W \mathbf{v}_1 = \text{Proj}_W \mathbf{v}_2$. Must $\mathbf{v}_1 = \mathbf{v}_2$?

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 $\mathbf{v}_1 = \mathbf{v}_2$?

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

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?

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 $\|\mathbf{v}\| = 1$.

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.

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?

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

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

11. Suppose you have a collection of data $\{(t_i, y_i) \mid 1 \le i \le N\}$, 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?

12. Suppose you wish to approximate a collection of data $\{(t_i, y_i) \mid 1 \le i \le N\}$, 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.

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.

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

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.

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