21-241 Matrices and Linear Transformations **Fall 2019**
Review of Topics **Friday, September 27**

1 Definitions

You should be able to define each of the following terms

- field
- associative
- commutative
- distributive
- scalar
- vector
- vector addition
- scalar multiplication
- linear combination
- dot product
- matrix
- matrix addition
- matrix multiplication
- diagonal matrix
- identity matrix
- permutation matrix
- standard basis
- transpose
- upper triangular matrix
- lower triangular matrix
- linear equation
- system of linear equations
- solution to a system of linear equations

- Gaussian elimination
- row operation
- pivot (in Gaussian elimination)
- elimination matrix
- LU decomposition
- square matrix
- forward substitution
- back substitution
- invertible matrix/nonsingular matrix
- noninvertible matrix/singular matrix
- affine combination
- symmetric matrix
- block matrix
- inverse matrix
- vector space
- span
- subspace
- column space
- linear transformation
- linear independence
- basis
- null space
- Reduced row echelon form (RREF)
- rank of a matrix
- full row rank
- full column rank
- tall matrix
- wide matrix
- dimension

- row space
- left null space
- length of a vector
- unit vector
- orthogonal vectors
- orthogonal subspaces
- orthogonal complement
- projection of a vector onto a subspace
- best approximate solution to a linear system
- projection matrix
- line of best fit
- least squares approximation
- orthonormal set of vectors
- orthonormal basis
- orthogonal matrix
- QR decomposition

2 Computations

You should be able to

- Add vectors and multiply them by scalars (take linear combinations) and interpret geometrically.
- Compute dot products of vectors.
- Add matrices.
- Multiply matrices with vectors and other matrices.
- Multiply a scalar by a matrix.
- Use rules on page 73 to manipulate matrix expressions.
- Translate between a system of linear equations and a matrix equation of the form $A\mathbf{x} = \mathbf{b}$.
- Translate between the three row operations and multiplication on the left by appropriate elimination matrices.

- Perform Gaussian elimination to find the solution to a system of equations with n equations and n variables, or find that no solution exists.
- Use the steps of Gaussian elimination to give an LU decomposition for a square matrix.
- Interpret matrices as blocks of submatrices and compute products using block multiplication.
- Given an $n \times n$ matrix A, use Gauss-Jordan to find the inverse or verify that none exists.
- Check if a given set of vectors is a vector space.
- Use row operations to find RREF for any matrix.
- Use RREF to find a basis for the nullspace of a matrix.
- Use RREF to find a particular solution to $A\mathbf{x} = \mathbf{b}$ or determine that none exist.
- Use RREF to find a description of all solutions to $A\mathbf{x} = \mathbf{b}$ or determine that none exist.
- Use RREF to find a basis for the column space.
- Use RREF to find a basis for the row space.
- Use RREF of A^T to find a basis for the left null space.
- Compute the (Euclidean) length of a vector in $\mathbb{R}^n$ using the dot product.
- Find the angle between two vectors in $\mathbb{R}^n$ using the dot product. Determine if two vectors are orthogonal.
- Given a basis for a subspace $S \subseteq \mathbb{R}^n$, find the projection $\mathbf{p}$ of any vector $\mathbf{b} \in \mathbb{R}^n$ onto S.
- Given a basis for a subspace $S \subseteq \mathbb{R}^n$, find the projection matrix P so that the projection of any vector $\mathbf{b} \in \mathbb{R}^n$ onto S is $P\mathbf{b}$.
- Given an $m \times n$ matrix A and a vector $\mathbf{b} \in \mathbb{R}^m$, find the best approximate solution $\hat{\mathbf{x}} \in \mathbb{R}^n$ to the system $A\mathbf{x} = \mathbf{b}$.
- Find the line of best (least squares) fit for a set of data.
- Find the parabola of best fit for a set of data.
- Given a basis for a vector space V, use the Gram-Schmidt algorithm to find an orthonormal basis for V.
- Given an $m \times n$ matrix A, find an $m \times n$ matrix Q with orthogonal columns and an $n \times n$ upper triangular matrix R such that $A = QR$.

3 Theorems and proofs

You should be able to prove or explain why

- Vector addition is associative and commutative (using axioms for a field)
- Scalar multiplication distributes over vector addition (using axioms for a field)
- The product of two lower (upper) triangular matrices is lower (upper) triangular
- Matrix multiplication is associative
- Matrix equations on Page 73 hold (e.g. $A + B = B + A$).
- Row operations do not change the set of solutions for a linear system of equations.
- Gaussian elimination either finds the solution to a square system of equations or shows that none exists.
- A square matrix A can be written as $A = LU$, or $PA = LU$.
- An $n \times n$ system of linear equations has exactly one solution if and only if all pivots in Gaussian elimination are nonzero.
- A system of linear equations (in $\mathbb{R}$ or $\mathbb{C}$) always has 0, 1, or ∞ solutions.
- $A^T A$ is symmetric
- The product $A\mathbf{x}$ is a linear combination of the columns of A.
- If A is $n \times n$ and contains a nonzero vector in the nullspace, it is not invertible.
- The Gauss-Jordan algorithm works (i.e. it computes the inverse of a square matrix or determines that none exists).
- A square matrix has an inverse if and only if all pivots in Gaussian elimination are nonzero.
- If A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$
- If A has an inverse, it is unique
- Every subspace of $\mathbb{R}^n$ is a hyperplane through the origin.
- The span of a set of vectors in V is a subspace of V.
- The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b} \in C(A)$.
- The column space is a subspace of $\mathbb{R}^m$.
- An $m \times n$ matrix is a linear transformation with domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$.
- The column space of AB is a subspace of the column space of A.

- The intersection of two vector spaces is a vector space.
- The union of two vector spaces is not a vector space unless one contains the other.
- Given a basis for a vector space V, for any vector $\mathbf{v} \in V$ there is exactly one way to express $\mathbf{v}$ as a linear combination of the vectors in the basis.
- The null space is a vector space.
- The null space of A is $\{\mathbf{0}\}$ if and only if the columns of A are linearly independent.
- If $A\mathbf{x} = A\mathbf{y}$, then $(\mathbf{x} - \mathbf{y}) \in N(A)$.
- An $n \times n$ matrix A is invertible if and only if $C(A) = \mathbb{R}^n$ and $N(A) = \{\mathbf{0}\}$.
- An $n \times n$ matrix A is invertible if and only if its vectors form a basis for $\mathbb{R}^n$.
- If A has full column rank then $A\mathbf{x} = \mathbf{b}$ has zero solutions or one solution.
- If A has full row rank then $A\mathbf{x} = \mathbf{b}$ has ∞ solutions.
- An $n \times n$ matrix is invertible if and only if it has rank n.
- Every basis for a vector space has the same number of vectors.
- The dimension of the row space equals the rank of the matrix.
- The dimension of the column space equals the rank of the matrix.
- The dimension of the null space of an $m \times n$ matrix with rank r is $n - r$.
- The dimension of the left null space of an $m \times n$ matrix with rank r is $m - r$.
- For any $m \times n$ matrix A, $N(A)$ and $C(A^T)$ are orthogonal complements in $\mathbb{R}^n$.
- For any $m \times n$ matrix A, any $\mathbf{x} \in \mathbb{R}^n$ can be written as $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$, where $\mathbf{x}_r \in C(A^T)$ and $\mathbf{x}_n \in N(A)$.
- For any $m \times n$ matrix A, $N(A^T)$ and $C(A)$ are orthogonal complements in $\mathbb{R}^m$.
- For any $m \times n$ matrix A, any $\mathbf{b} \in \mathbb{R}^m$ can be written (uniquely) as $\mathbf{b} = \mathbf{p} + \mathbf{e}$, where $\mathbf{p} \in C(A)$ and $\mathbf{e} \in N(A^T)$.
- If A has full column rank, then $A^T A$ is invertible.
- For any $m \times n$ matrix A, the best approximate solution to $A\mathbf{x} = \mathbf{b}$ is given by $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$.
- For any $m \times n$ matrix A, then the projection of $\mathbf{b} \in \mathbb{R}^m$ onto $C(A)$ is $\mathbf{p} = A(A^T A)^{-1} A^T \mathbf{b}$.
- For any $m \times n$ matrix A, the projection matrix for the projection onto $C(A)$ is $P = A(A^T A)^{-1} A^T$.

- If P is a projection matrix, then $P^2 = P$.
- If A is an invertible matrix, then the projection of any vector $\mathbf{b}$ onto $C(A)$ is $\mathbf{b}$. (Equivalently, the projection matrix P is I.)
- The coefficients for the line of best fit for a set of data is given by the best approximate solution to a corresponding linear system.
- If the $m \times n$ matrix Q has orthonormal columns, then $Q^TQ = I_n$.
- If Q has orthonormal columns, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ (Q preserves lengths).
- If Q has orthonormal columns, then $(Q\mathbf{x})^TQ\mathbf{y} = \mathbf{x}^T\mathbf{y}$ (Q preserves dot products, hence angles).
- If the $m \times n$ matrix Q has orthonormal columns, then the best approximate solution to $Q\mathbf{x} = \mathbf{b}$ is given by $\hat{\mathbf{x}} = Q^T\mathbf{b}$.
- If the $m \times n$ matrix Q has orthonormal columns, then the projection of $\mathbf{b} \in \mathbb{R}^m$ onto $C(Q)$ is $\mathbf{p} = QQ^T\mathbf{b}$.
- If the $m \times n$ matrix Q has orthonormal columns, then the projection matrix for the projection onto $C(Q)$ is $P = QQ^T$.
- If $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then any $\mathbf{x} \in R^n$ can be uniquely expressed as
$$\mathbf{x} = c_1\mathbf{q}_1 + \cdots + c_n\mathbf{q}_n,$$
where $c_i = \mathbf{q}_i^T\mathbf{x}$.

4 Equivalent statements about square matrices

Suppose A is an $n \times n$ matrix. Then each column contains equivalent statements.

A ia nonsingular	A is singular
A is invertible	A is not invertible
The columns are independent	The columns are dependent
The rows are independent	The rows are dependent
$A\mathbf{x} = \mathbf{0}$ has one solution, $\mathbf{x} = \mathbf{0}$	$A\mathbf{x} = \mathbf{0}$ has many solutions
$A\mathbf{x} = \mathbf{b}$ has one solution $\mathbf{x} = A^{-1}\mathbf{b}$	$A\mathbf{x} = \mathbf{b}$ has no solution or infinitely many
A has n (nonzero) pivots	A has $r < n$ pivots
A has full rank $r = n$	A has rank $r < n$
The reduced row echelon form is $R = I$	R has at least one zero row
The column space is all of $\mathbb{R}^n$	The column space has dimension $r < n$
The row space is all of $\mathbb{R}^n$	The row space has dimension $r < n$

5 Algorithms

Gaussian Elimination / LU decomposition

INPUT: An $n \times n$ matrix A, a vector $\mathbf{b} \in \mathbb{R}^n$.

OUTPUT (GE): A vector $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$, or a certificate that none exists.

OUTPUT (LU): A lower triangular L and an upper triangular U (with a permutation matrix P, if necessary), such that $A = LU$ (or $PA = LU$).

Procedure: Transform system $A\mathbf{x} = \mathbf{b}$ using 3 types of row operations until system is upper triangular, then solve using back substitution.

Why it works: Row operations do not change the set of solutions to a system of equations.

Theoretical insights:

- $A\mathbf{x} = \mathbf{b}$ has a unique solution for all $\mathbf{b} \in \mathbb{R}^n$ if and only if all pivots are nonzero (invertible/nonsingular case). Otherwise there may be 0 or ∞ solutions, depending on rows with zeros in pivot positions (noninvertible/singular case).

  Why: Every step of back substitution gives a unique value for x_i if and only if all pivots are nonzero.

- By interpreting row operations as multiplication (on the left) by elimination matrices, we use their inverses to obtain an $A = LU$ (or $PA = LU$ if row swaps are required) decomposition.

  This allows us to easily solve $A\mathbf{x} = \mathbf{b}$ for any $\mathbf{b}$ using forward substitution ($L\mathbf{c} = \mathbf{b}$) then back substitution ($U\mathbf{x} = \mathbf{c}$).

Gauss-Jordan

INPUT: An $n \times n$ matrix A

OUTPUT: An $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$, or a certificate that A has no inverse.

Procedure: Beginning with $[A|I]$, use row operations to transform to $[I|A^{-1}]$.

Why it works: If E is the product of elimination matrices that transforms A to I, then $EA = I$, so $E = A^{-1}$. By block multiplication, $E = A^{-1}$ appears on the right side of the augmented matrix at the end of the algorithm.

Theoretical insight:

- A is invertible if and only if there are n nonzero pivots.

  (If there are fewer nonzero pivots, it is impossible to transform the left side of the matrix to I).

RREF

INPUT: An $m \times n$ matrix A.

OUTPUT: A matrix $R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$ (possibly with columns permuted).

Procedure: Use row operations.

Theoretical insights:

- A basis for the null space of A is given by the $n - r$ columns of $N = \begin{bmatrix} -F \\ I \end{bmatrix}$ (with rows permuted to match column permutation of R).

  Why: Since each column has a 1 in a row where every other column has a 0, the columns are linearly independent. Since once the free variables are chosen, the pivot variables are uniquely determined, these columns span the null space.

- All solutions to $A\mathbf{x} = \mathbf{b}$ are given by $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_n$, where $\mathbf{x}_p$ is a particular solution to $A\mathbf{x} = \mathbf{b}$, and $\mathbf{x}_n$ is any element of the null space.

  Why: $A(\mathbf{x}_p + \mathbf{x}_n) = A\mathbf{x}_p + A\mathbf{x}_n = \mathbf{b} + \mathbf{0} = \mathbf{b}$. Further, if $A\mathbf{x} = A\mathbf{y}$, then $(\mathbf{x} - \mathbf{y}) \in N(A)$.

- $\dim C(A) = \dim C(A^T) = r$, where r is the rank of A.

- $\dim C(A) + \dim N(A) = n$ (Rank-nullity theorem)

  Why: The dimensions of the four fundamental subspaces are not altered by row operations.

- Characterization of solutions to $A\mathbf{x} = \mathbf{b}$ in four cases
  - $r = n < m$ (full column rank)
  - $r = m < n$ (full row rank)
  - $r = m = n$ (invertible)
  - $r < m, r < n$

Gram-Schmidt / QR decomposition

INPUT (G-S): A basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ for a vector space V.

OUTPUT (G-S): An orthonormal basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ for V.

INPUT (QR): An $m \times n$ matrix A with full column rank (linearly independent columns).

OUTPUT (QR): An $m \times n$ matrix Q with orthonormal columns and an $n \times n$ upper triangular matrix R such that $A = QR$.

Procedure (G-S): For each basis vector $\mathbf{a}_i$,

- From $\mathbf{a}_i$, generate $\mathbf{A}_i$ by subtracting the projection of $\mathbf{a}_i$ onto $\mathbf{A}_j$ for every $j < i$.

- Let $\mathbf{q}_i = \mathbf{A}_i / \|\mathbf{A}_i\|$.

Why it works: The first step produces vectors orthogonal to all previously produced vectors because $(\mathbf{w} - proj_{\mathbf{v}}\mathbf{w}) \perp \mathbf{v}$. The second step normalizes them to each have length 1.

Procedure (QR): Rewrite the system of equations generated by Gram-Schmidt to express each $\mathbf{a}_i$ in terms of $\mathbf{q}_i$'s. Combine these into a matrix equation $A = QR$ where $\mathbf{a}_i$ is the ith column of A, $\mathbf{q}_i$ is the ith column of Q, and $R_{ij} = \mathbf{q}_i^T \mathbf{a}_j$.

Theoretical insights:

- The projection of a vector $\mathbf{w}$ onto a vector $\mathbf{v}$ can be written as $\frac{\mathbf{v}^T\mathbf{w}}{\mathbf{v}^T\mathbf{v}}\mathbf{v}$.
- For all k, the span of $\{\mathbf{a}_1 \ldots, \mathbf{a}_k\}$ is the same as the span of $\{\mathbf{q}_1 \ldots, \mathbf{q}_k\}$.
- Since R is upper triangular (i.e. $R_{ij} = 0$ if $i > j$) and $R_{ij} = \mathbf{q}_i^T \mathbf{a}_j$, the basis vector $\mathbf{q}_i$ is orthogonal to every $\mathbf{a}_j$ where $j < i$ (i.e. all of the "previously considered vectors" in the basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$).