Review of Linear Algebra for the Final

Additional Notes Summarized by Yourself

You can fill in this empty block to summarize the course contents that are not listed in this file.

Reduced Row Echelon Form (RREF)

A matrix is in Reduced Row Echelon Form if it satisfies the following 4 conditions

1. All zero rows are at the bottom.
2. The first non-zero entry of every non-zero row is a 1 (leading one).
3. Leading ones go from left to right.
4. All entries above and below any leading one are zero.

If a matrix satisfies only the first 3 conditions above then we say it is in Row Echelon Form (REF).

The Row Reduction Algorithm

Step 1: Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

Step 2: Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

Step 3: Use row replacement operations to create zeros in all positions below the pivot.

Step 4: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1 to 3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Step 5: Backward phase. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

Steps 1-4 produce a matrix in row echelon form (REF). A fifth step produces a matrix in reduced row echelon form (RREF).

Existence and Uniqueness Theorem

A linear system is consistent if and only if the rightmost column of the augmented matrix is not a pivot column, that is, if and only if an echelon form of the augmented matrix has no row of the form $\begin{bmatrix} 0 & \cdots & 0 & b \end{bmatrix}$ with b nonzero.

If a linear system is consistent, then the solution set contains either
(i) a unique solution, when there are no free variables, or
(ii) infinitely many solutions, when there is at least one free variable.

Using Row Reduction to Solve a Linear System

1. Write the augmented matrix of the system.
2. Use the row reduction algorithm to obtain an equivalent augmented matrix in echelon form. Decide whether the system is consistent. If there is no solution, stop; otherwise, go to the next step.
3. Continue row reduction to obtain the reduced echelon form.
4. Write the system of equations corresponding to the matrix obtained in step 3 .
5. Rewrite each nonzero equation from step 4 so that its one basic variable is expressed in terms of any free variables appearing in the equation.

The Matrix Equation $A\mathbf{x} = \mathbf{b}$

An equation in the form of $A\mathbf{x} = \mathbf{b}$ is called a *matrix equation.*

Theorem: If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}$$

Theorem: Let A be an $m \times n$ matrix. Then the following statements are logically equivalent.

1. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
2. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
3. The columns of A span $\mathbb{R}^m$.
4. A has a pivot position in every row.
5. $T(\mathbf{x}) = A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$

Linear Combination and Span

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \ldots, c_p$, the vector $\mathbf{y}$ defined by $\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ is called a *linear combination* of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ with weights $c_1, \ldots, c_p$.

If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by $\mathit{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$

Homogeneous Linear Systems

Definition: A system of linear equations is said to be homogeneous if it can be written in the form $A\mathbf{x} = \mathbf{0}$, where A is an $m \times n$ matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$.

Theorem: The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\iff$ the equation has at least one free variable

Parametric Vector Form

Summary: Writing a solution set (of a consistent system) in parametric vector form

1. Row reduce the augmented matrix to reduced echelon form.
2. Express each basic variable in terms of any free variables appearing in an equation.
3. Write a typical solution $\mathbf{x}$ as a vector whose entries depend on the free variables, if any.
4. Decompose $\mathbf{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

Linear Independence

Definition: A set $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ of vectors in $\mathbb{R}^n$ is said to be *linearly independent* if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

is $c_1 = c_2 = \cdots = c_p = 0$. Otherwise the vectors are called *linearly dependant* (which also means that at least one of them can be written as a linear combination of the others).

Theorems:
1. A set containing only one vector $\mathbf{v}$ is linearly independent if and only if $\mathbf{v}$ is not the zero vector.
2. A set of two vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent if at least one of the vectors is a multiple of the other.
3. An indexed set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others.
4. If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.
5. If a set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Sums and Scalar Multiples

Let A, B, and C be matrices of the same size, and let r and s be scalars.

- $A+B=B+A$
- $r(A+B)=rA+rB$
- $(A+B)+C=A+(B+C)$
- $(r+s)A=rA+sA$
- $A+\mathbf{0}=A$
- $r(sA)=(rs)A$

Properties of Matrix Multiplication

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

- $A(BC)=(AB)C$ (associative law of multiplication)
- $A(B+C)=AB+AC$ (left distributive law)
- $(B+C)A=BA+CA$ (right distributive law)
- $r(AB)=(rA)B=A(rB)$ for any scalar r
- $I_m A=A=AI_n$ (identity for matrix multiplication)

Transpose of a Matrix

Definition: Given an $m \times n$ matrix A, the *transpose* of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

Properties:
- $\left(A^T\right)^T=A$
- $(A+B)^T=A^T+B^T$
- For any scalar $r, (rA)^T=rA^T$
- $(AB)^T=B^T A^T$

Transformation, Domain, Codomain, Image and Range

A *transformation* (or function or mapping) T from a vector space V into a vector space W is a rule that assigns to each vector $\mathbf{x}$ in V a vector $T(\mathbf{x})$ in W.

The set V is called the *domain* of T, and W is called the *codomain* of T.
For $\mathbf{x}$ in V, the vector $T(\mathbf{x})$ in W is called the *image* of $\mathbf{x}$.

The set of all images $T(\mathbf{x})$ is called the *range* of T.

The *kernel* of such a T is the set of all u in V such that $T(\mathbf{u})=\mathbf{0}$ (the zero vector in W).

Linear Transformations

Definition: A transformation (or mapping) T is *linear* if
(1) $T(\mathbf{u}+\mathbf{v})=T(\mathbf{u})+T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;
(2) $T(c\mathbf{u})=cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

Properties: If T is a linear transformation, then
(1) $T(\mathbf{0})=\mathbf{0}$
(2) $T\left(c_1\mathbf{v}_1+\cdots+c_p\mathbf{v}_p\right)=c_1T\left(\mathbf{v}_1\right)+\cdots+c_pT\left(\mathbf{v}_p\right)$

Standard Matrix for the Linear Transformation

Let $T:\mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. The *standard matrix for the linear transformation* T is

$$A=\left[\begin{array}{lll} T\left(\mathbf{e}_1\right) & \cdots & T\left(\mathbf{e}_n\right) \end{array}\right]$$

where $\mathbf{e}_j$ is the j th column of the identity matrix in $\mathbb{R}^n$. A is the $m \times n$ matrix and

$$T(\mathbf{x})=A\mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n.$$

Examples:
- Counterclockwise rotation about the origin for a positive angle φ: $\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$
- Reflection through the x_1-axis: $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
- Reflection through the line $x_2=x_1$: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Onto and One-to-One Linear Transformations

Onto:
- A mapping $T:\mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be onto $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$. This is an existence question.
- Let A be the standard matrix for T, then T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$ (if and only if A has a pivot position in every row).

One-to-One:
- A mapping $T:\mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be one-to-one if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$. This is a uniqueness question.
- T is one-to-one if and only if the equation $T(\mathbf{x})=\mathbf{0}$ has only the trivial solution.
- Let A be the standard matrix for T, then T is one-to-one if and only if the columns of A are linearly independent.

Coordinate Systems

Suppose the set $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for a subspace H. For each $\mathbf{x}$ in H, the coordinates of $\mathbf{x}$ relative to the basis $\mathcal{B}$ are the weights $c_1, \dots, c_p$ such that $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p$, and the vector in $\mathbb{R}^p$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

is called the coordinate vector of $\mathbf{x}$ (relative to $\mathcal{B}$) or the $\mathcal{B}$-coordinate vector of $\mathbf{x}$.

The Inverse of a Matrix

Definition: Given a square matrix A its *inverse* (if it exists) is the matrix denoted by A^{-1} such that $AA^{-1} = A^{-1}A = I$.

Find A^{-1}: (1) If the matrix is a 2×2 matrix, we use the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

(2) For a matrix of higher dimensions, row reduce the augmented matrix $[A \quad I]$ to get $\left[I \quad A^{-1}\right]$. If the matrix is not invertible, we will not get the identity on the left side after applying the row reduction process.

(3) We can also use the formula

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} = \frac{1}{\det A}\operatorname{adj} A,$$

where C_{ji} is a cofactor of A.

In particular, we have the (i, j)-entry of A^{-1} given by

$$(A^{-1})_{i,j} = \frac{1}{\det(A)} C_{j,i}.$$

Properties: If A is an invertible $n \times n$ matrix, then

- then A^{-1} is invertible and $\left(A^{-1}\right)^{-1} = A$
- if B is $n \times n$ invertible, then so is AB, and $(AB)^{-1} = B^{-1}A^{-1}$
- A^T is also invertible and $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$
- The Invertible Matrix Theorem (next box).

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent.

1. A is an invertible matrix.
2. A is row equivalent to the $n \times n$ identity matrix.
3. A has n pivot positions.
4. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
5. The columns of A form a linearly independent set.
6. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
7. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
8. The columns of A span $\mathbb{R}^n$.
9. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
10. There is an $n \times n$ matrix C such that $CA = I$.
11. There is an $n \times n$ matrix D such that $AD = I$.
12. A^T is an invertible matrix.
13. The columns of A form a basis of $\mathbb{R}^n$.
14. $\operatorname{Col} A = \mathbb{R}^n$.
15. $\operatorname{rank} A = n$.
16. $\dim \operatorname{Nul} A = 0$.
17. $\operatorname{Nul} A = \{\mathbf{0}\}$.
18. $\det A \neq 0$.
19. The number 0 is not an eigenvalue of A.

Determinant

Minor: Given $A_{n\times n}$, the minor of entry ij is denoted by A_{ij} and is the determinant of the matrix obtained from A by removing row i and column j.

Cofactor: $C_{ij} = (-1)^{i+j}\det A_{ij}$

Determinant: Given an $n \times n$ matrix $A (n \geqslant 2)$

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

by expanding along the i^{th} row.

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

by expanding along the j^{th} column.

Properties: Given an $n \times n$ matrix A,

- if A has a zero row or zero column then $\det(A) = 0$.
- if we get matrix B by interchanging two rows of A then $\det(B) = -\det(A)$.
- if we get matrix B by multipying one row of A by $k \neq 0$ then $\det(B) = k\det(A)$.
- if we get matrix B by adding a multiple of a row to another of matrix A then $\det(B) = \det(A)$
- $\det(kA) = k^n \det(A)$
- $\det\left(A^T\right) = \det(A)$
- $\det(AB) = \det(A)\det(B)$
- $\det\left(A^{-1}\right) = \dfrac{1}{\det(A)}$

Vector Spaces

Definition: A *vector space* is a non-empty set V of objects, called vectors, on which are defined two operations, called addition and multiplication by scalars (real numbers), subject to the ten axioms listed below. The axioms must hold for all vectors $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ in V and for all scalars c and d.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a zero vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$, is in V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

Examples:

1. The spaces $\mathbb{R}^n$, where $n \geq 1$.
2. The set $\mathbb{P}_n$ of polynomials of degree at most n, where $n \geq 0$.
3. The set $M_{m\times n}$ of all $m \times n$ matrices with real entries, where m and n are positive integers.
4. The set of all real-valued functions defined on a set $\mathbb{D}$.

Subspaces

Definition: A *subspace* of a vector space V is a subset H of V that has three properties:

1. The zero vector of V is in H.
2. H is closed under vector addition. That is, for each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
3. H is closed under multiplication by scalars. That is, for each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

Examples:

1. In every vector space V, the subsets $\{\mathbf{0}\}$ and V are subspaces.
2. A line through the origin in $\mathbb{R}^2$ or $\mathbb{R}^3$.
3. A plane through the origin in $\mathbb{R}^3$. For example, the solutions to the homogeneous equation $3x + 4y + 5z = 0$ is a plane through the origin in $\mathbb{R}^3$.
4. All polynomials in $\mathbb{P}_n$ such that $\mathbf{p}(a) = 0$ for some fixed $a \in \mathbb{R}$ and positive integer n.
5. The set of all 3×3 symmetric matrices. Note we say an $n \times n$ matrix A is said to be symmetric if $A^T = A$. (Exercise 7 in the Lecture Notes §4.1).

Subspaces (continued)

Non-Examples:
1. A line in $\mathbb{R}^2$ or $\mathbb{R}^3$ not containing the origin.
2. A plane in $\mathbb{R}^3$ not containing the origin. For example, the solutions to the non-homogeneous equation $3x+4y+5z=6$ is a plane not containing the origin in $\mathbb{R}^3$. It is not a subspace of $\mathbb{R}^3$.
3. The first quadrant in $\mathbb{R}^2$.
4. All polynomials in $\mathbb{P}_n$ such that $\mathbf{p}(a)=3$ for some fixed $a \in \mathbb{R}$ and positive integer n.

Basis, Dimension

Basis: A *basis* for a subspace H of a vector space V is a linearly independent set in H that spans H.

Dimension: If a vector space V is spanned by a finite set, then V is said to be finite-dimensional, and the dimension of V, written as $\dim V$, is the number of vectors in a basis for V.

Col A, Nul A, Row A

Col A:
- The *column space* of a matrix A is the set $\operatorname{Col} A$ of all linear combinations of the columns of A.
- Col A is a subspace of $\mathbb{R}^m$ if A is $m \times n$.
- The pivot columns of a matrix A form a basis for the column space of A.
- $\operatorname{rank} A = \dim \operatorname{Col} A$

Nul A:
- The *null space* of a matrix A is the set Nul A of all solutions of the homogeneous equation $A\mathbf{x}=\mathbf{0}$.
- To test whether a given vector $\mathbf{v}$ is in Nul A, just compute $A\mathbf{v}$ to see whether $A\mathbf{v}$ is the zero vector.
- To find a basis for $\operatorname{Nul} A$, we solve the equation $A\mathbf{x}=\mathbf{0}$ and write the solution for $\mathbf{x}$ in parametric vector form. The vectors in the parametric form give us a basis for $\operatorname{Nul} A$.
- The nullity of a matrix A is the dimension of its $\operatorname{Nul} A$.

Row A:
- The set of all linear combinations of the row vectors of A is called the *row space* of A, and is denoted by Row A.
- Row A is a subspace of $\mathbb{R}^n$ if A is $m \times n$.
- If two matrices A and B are row equivalent, then their row spaces are the same. If B is in echelon form, the nonzero rows of B form a basis for the row space of A as well as for that of B.
- Row $A = \operatorname{Col} A^T$
- dim Row $A = \dim \operatorname{Col} A = \operatorname{rank} A$

Rank Thm: $\operatorname{rank} A+$ nullity $A =$ number of columns in A

Linearly Independent Sets

Definition: A set $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ of vectors in a vector space V is said to be *linearly independent* if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

is $c_1 = c_2 = \cdots = c_p = 0$. Otherwise the vectors are called *linearly dependant* (which also means that at least one of them can be written as a linear combination of the others).

Eigenvalues and Eigenvectors

Definition: A scalar λ is called an *eigenvalue* of A if $|\mathbf{A}-\lambda\mathbf{I}|=0$ (*characteristic equation*).
An *eigenvector* associated with the eigenvalue λ is a nonzero vector $\mathbf{v}$ such that $(\mathbf{A}-\lambda\mathbf{I})\mathbf{v}=\mathbf{0}$.

Eigenspace: Given a particular eigenvalue λ of the n by n matrix A, the set $E=\{\mathbf{v} : (A-\lambda I)\mathbf{v}=\mathbf{0}\}$ is called the *eigenspace* of A associated with λ.

Properties:
- If $A\mathbf{x}=\lambda\mathbf{x}$, then $A^k\mathbf{x}=\lambda^k\mathbf{x}$ for any positive integer k. So λ^k is an eigenvalue for A^k. Check Practice Problems # 2 on Page 279. Solutions are on Page 282.
- If $A\mathbf{x}=\lambda\mathbf{x}$, then $s\lambda$ is an eigenvalue of sA for any real number s.
- The eigenvalues of a triangular matrix are the entries on its main diagonal.
- If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

The Characteristic Polynomial of a 2×2 matrix

Let A be a 2×2 matrix, then its characteristic polynomial is

$$\lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$$

Remark: Recall the *trace* of a square matrix A is the sum of the diagonal entries in A and is denoted by $\operatorname{tr} A$.

Similarity

Definition: If A and B are $n \times n$ matrices, then A *is similar to* B if there is an invertible matrix P such that $P^{-1}AP = B$, or, equivalently, $A = PBP^{-1}$.

Properties:
- Any square matrix A is similar to itself. (Reflexivity)
- A is similar to B if and only if B is similar to A. (Symmetry)
- If A is similar to B and B is similar to C, then A is similar to C. (Transitivity)
- If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).
- If A and B are similar, then $\det A = \det B$ (Example 4 in §5.2).
- Similar matrices have the same rank.

Warnings:
- It is not true that if two matrices have the same eigenvalues implies they are similar. Check the Lecture Notes §5.2 for an example.
- Similarity is not the same as row equivalence. Row operations on a matrix usually change its eigenvalues.

Diagonalization

Definition: A square matrix A is said to be *diagonalizable* if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D.

Properties:
- An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
- An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.
- A symmetric matrix is always diagonalizable.

Diagonalizing A: Check Example 3 in Lecture Notes §5.3 as an exercise.
Step 1. Find the eigenvalues of A
Step 2. Find n linearly independent eigenvectors of A if A is $n \times n$. (A is not diagonalizable if this step fails.
Step 3. Construct P with the eigenvectors found in Step 2.
Step 4. Construct the diagonal matrix D with the the corresponding eigenvalues from columns of P.

Warnings:
- When A has fewer than n distinct eigenvalues, it is still possible to diagonalize A. (Example 3 in Lecture Notes §5.3)
- Diagonalizable $\not\Rightarrow$ Invertible. For example, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is diagonalizable but not invertible.
- Invertible $\not\Rightarrow$ Diagonalizable. For example, $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is invertible but not diagonalizable.
- Diagonalizable $\not\Rightarrow$ no zero eigenvalues. For example, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is diagonalizable and 0 is an eigenvalue.

Applications to Differential Equations

Constant Coeff. Homogeneous: $\mathbf{x}' = A\mathbf{x}$

Solution: $\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots,$
where $\mathbf{x}_i$ are fundamental solutions from eigenvalues & eigenvectors.
The method is described as below.

The Eigenvalue Method for $\mathbf{x}' = A\mathbf{x}$ in §5.7:

We consider A to be 2×2, then the general solution is $\mathbf{x}(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t)$, with the fundamental solutions $\mathbf{x}_1(t), \mathbf{x}_2(t)$ found has follows.

- Distinct Real Eigenvalues. $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda_1 t}, \mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda_2 t}$
- Complex Eigenvalues. $\lambda_{1,2} = p \pm qi$. (*suggestion: use an example to review the method*)

  If $\mathbf{v} = \mathbf{a} + i\mathbf{b}$ is an eigenvector associated with $\lambda = p + qi$, then

  $\mathbf{x}_1(t) = e^{pt}(\mathbf{a}\cos qt - \mathbf{b}\sin qt)$, $\mathbf{x}_2(t) = e^{pt}(\mathbf{b}\cos qt + \mathbf{a}\sin qt)$.

Trajectories for the System $\mathbf{x}' = A\mathbf{x}$:

- **attractor**: A has distinct negative real eigenvalues.
- **repeller**: A has distinct positive real eigenvalues.
- **saddle point**: A has real eigenvalues of opposite sign.
- **spiral point**: A has complex conjugate eigenvalues with nonzero real parts.
- **center**: A has purely imaginary eigenvalues.

Inner Product Spaces, Length, and Orthogonality

Inner Product: An inner product on a vector space V is a function that, to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V, associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ and satisfies the following axioms, for all $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ in V and all scalars c : (1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$. (2) $\langle \mathbf{u}+\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$. (3) $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$. (4)$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$. A vector space with an inner product is called an *inner product space.*

Length: The length (or norm) of a vector v is the scalar $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$

Distance: The *distance* between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$.

Orthogonality: Vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Orthogonal Sets

<u>Definitions:</u>
- A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be *an orthogonal set* if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.
- A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an *orthonormal set* if it is an orthogonal set of unit vectors.

<u>Properties:</u>
- If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.
- An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.
- Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then (1) $\|U\mathbf{x}\| = \|\mathbf{x}\|$, (2) $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$, (3) $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.
- An orthogonal matrix is a real square matrix whose columns and rows are orthonormal vectors. Equivalently, a matrix P is orthogonal if its transpose is equal to its inverse: $P^T = P^{-1}$.

Orthogonal Projections

<u>Theorem:</u> (The Orthogonal Decomposition Theorem) Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \operatorname{proj}_W \mathbf{y} + \mathbf{z}$$

where $\operatorname{proj}_W \mathbf{y}$ is in W and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\operatorname{proj}_W \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p$$

and $\mathbf{z} = \mathbf{y} - \operatorname{proj}_W \mathbf{y}$.

<u>Theorem:</u> (The Best Approximation Theorem) Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\operatorname{proj}_W \mathbf{y}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\operatorname{proj}_W \mathbf{y}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \operatorname{proj}_W \mathbf{y}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all $\mathbf{v}$ in W distinct from $\operatorname{proj}_W \mathbf{y}$. $\operatorname{proj}_W \mathbf{y}$ is called the best approximation to y by elements of W.

Least-Squares Problems

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a *least-squares solution* of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$.

The Gram-Schmidt Process

Given a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthogonal basis for W.
In addition, $\operatorname{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \operatorname{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ for $1 \leq k \leq p$.

The QR Factorization

<u>The QR Factorization:</u> If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $\operatorname{Col} A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

<u>Remark:</u> The matrix Q can be obtained from the Gram-Schmidt Process. Notice that $R = Q^T A$ since $U^T U = I$.

Examples of Inner Product Spaces Other Than $\mathbb{R}^n$

<u>An inner product on $\mathbb{P}_n$:</u> Let $t_0, \dots, t_n$ be distinct real numbers. For p and q in $\mathbb{P}_n$,
$\langle p, q \rangle = p(t_0)\,q(t_0) + p(t_1)\,q(t_1) + \cdots + p(t_n)\,q(t_n)$
defines an inner product on $\mathbb{P}_n$.

<u>An Inner Product on $C[a,b]$:</u> For f, g in $C[a,b]$,

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

defines an inner product on $C[a,b]$.

Diagonalization of Symmetric Matrices

A symmetric matrix is a matrix A such that $A^T = A$.

<u>Properties:</u>
- An $n \times n$ matrix A has n real eigenvalues, counting multiplicities.
- If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.
- An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.