

Linear Algebra

HECHEN HU

December 27, 2017

Contents

1	Solving Linear Equations	1
1.1	Systems of Linear Equations	1
1.2	Row Reduction and Echelon Forms	2
1.3	Vector Equations	2
1.4	The Matrix Equation $\mathbf{Ax} = \mathbf{b}$	3
1.5	Solution Sets of Linear Systems	4
1.6	Linear Independence	4
1.7	Linear Transformations	5
2	Matrices	7
2.1	Matrices and Arithmetic Operations on Them	7
2.2	The Inverse of a Matrix	8
2.3	Subspaces of \mathbb{R}^n	11
3	Determinants	13
3.1	Determinants and some other Concepts	13
3.2	Properties of Determinants	14
4	Vector Spaces	17
4.1	Coordinate Systems	19
4.2	Rank	20
4.3	Change of Basis	20
5	Eigenvalues and Eigenvectors	21
5.1	Definition	21
5.2	Diagonalization	22
5.3	Complex Eigenvalues	22
6	Orthogonality	23
6.1	Orthogonal Sets	23
6.2	Orthogonal Projections	24
7	Symmetric Matrices and Quadratic Forms	27

1

Solving Linear Equations

1.1 Systems of Linear Equations

Definition. A *linear equation* in the variables x_1, \dots, x_n is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

where b and coefficients a_i are real or complex numbers. A *linear system* is a collection of one or more linear equations involving the same variables. A *solution* of the system is a list of numbers that makes each equation a true statement when their values are substituted for x_1, \dots, x_n respectively. The set of all possible solutions is called the *solution set* of the linear system. Two linear systems are called *equivalent* if they have the same solution set.

Definition. A linear system is *consistent* if it has either one solution or infinitely many solutions. If it has no solution, it is called *inconsistent*.

Definition. The *coefficient matrix* is the matrix where the coefficients of each variable in a system aligned in columns. If additionally the coefficient of the right-hand side of equations are added to the coefficient matrix, a new matrix called *augmented matrix* is generated.

Definition. *Elementary row operations* on a matrix include:

- (*Replacement*) Replace one row by the sum of itself and a multiple of another row.
- (*Interchange*) Interchange two rows.
- (*Scaling*) Multiply all entries in a row by a nonzero constant.

Two matrices are *row equivalent* if there is a sequence of elementary operations that transforms one matrix into the other.

Theorem 1.1.1. *If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.*

1.2 Row Reduction and Echelon Forms

Definition. A rectangular matrix is in *echelon form* if it has the following properties:

- All nonzero rows are above any rows of all zeros;
- Each leading entry of a row is in a column to the right of the leading entry of the row above it;
- All entries in a column below a leading entry are zero.

If a matrix in echelon form satisfies the following additional conditions, then it is in *reduced echelon form*:

- The leading entry in each nonzero row is 1;
- Each leading 1 is the only nonzero entry in its column.

Theorem 1.2.1. *Each matrix is row equivalent to an unique reduced echelon matrix.*

If a matrix A is row equivalent to an (reduced)echelon matrix U , U is called an *(reduced) echelon form of A* . The abbreviation RREF and REF are used for reduced (row) echelon form and (row) echelon form respectively.

Definition. A *pivot position* in a matrix A is a location in A that corresponds to a leading entry in an echelon form of A . A *pivot column* is a column of A that contains a pivot position.

Theorem 1.2.2. *A linear system is consistent iff the rightmost column of the augmented matrix **is not** a pivot column, that is, iff an echelon form of the augmented matrix has **no** row of the form*

$$\begin{bmatrix} 0 & \cdots & 0 & b \end{bmatrix} \quad \text{with } b \text{ nonzero}$$

If a linear system is consistent, then the solution set contains either

- *a unique solution, when there are no free variables.*
- *infinitely many solutions, when there is at least one free variable.*

1.3 Vector Equations

Definition. A matrix with only one column is called a *column vector*, or simply a *vector*.

Definition. Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ in \mathbb{R}^n and given scalars c_1, c_2, \dots, c_p , the vector \mathbf{y} defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p$$

is called a *linear combination* of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ using weights c_1, c_2, \dots, c_p .

Definition. If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ are in \mathbb{R}^n , then the set of all linear combinations of them is denoted by $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the *subset of \mathbb{R}^n spanned (or generated) by $\mathbf{v}_1, \dots, \mathbf{v}_p$* . That is, $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p$$

with c_1, c_2, \dots, c_p scalars.

1.4 The Matrix Equation $\mathbf{Ax} = \mathbf{b}$

Definition. If \mathbf{A} is an $m \times n$ matrix, with columns a_1, \dots, a_n , and if \mathbf{x} is in \mathbb{R}^n , then the *product of \mathbf{A} and \mathbf{x}* , denoted by \mathbf{Ax} , is the *linear combination of the columns of \mathbf{A} using the corresponding entries in \mathbf{x} as weights*, that is,

$$\mathbf{Ax} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} = x_1a_1 + x_2a_2 + \dots + x_na_n$$

\mathbf{Ax} is defined only if the number of columns of \mathbf{A} equals the number of entries in \mathbf{x} .

Definition. Equations having the form $\mathbf{Ax} = \mathbf{b}$ are called *matrix equations*.

Theorem 1.4.1. If \mathbf{A} is an $m \times n$ matrix, with columns a_1, \dots, a_n , and \mathbf{b} is in \mathbb{R}^m , the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1a_1 + x_2a_2 + \dots + x_na_n = \mathbf{b}$$

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

$$\begin{bmatrix} a_1 & a_2 & \dots & a_n & \mathbf{b} \end{bmatrix}$$

Definition. A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in \mathbb{R}^m *spans (or generates) \mathbb{R}^m* if $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\} = \mathbb{R}^m$.

Theorem 1.4.2. Let \mathbf{A} be an $m \times n$ coefficient matrix. Then the following statements are logically equivalent, that is, for a particular \mathbf{A} , either they are all true statements or they are all false.

- For each \mathbf{b} in \mathbb{R}^m , the equation $\mathbf{Ax} = \mathbf{b}$ has a solution.
- The columns of \mathbf{A} spans \mathbb{R}^m .
- \mathbf{A} has a pivot position in every row.

Theorem 1.4.3. If \mathbf{A} is an $m \times n$ matrix, \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , and c is a scalar, then

- $\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{Au} + \mathbf{Av}$.
- $\mathbf{A}(c\mathbf{u}) = c(\mathbf{Au})$.

1.5 Solution Sets of Linear Systems

Definition. A system of Linear equations is said to be *homogeneous* if it can be written in the form $\mathbf{Ax} = \mathbf{0}$. Such a system always has at least one solution, namely, $\mathbf{x} = \mathbf{0}$, and this solution is usually called the *trivial solution*. A homogeneous equation has a nontrivial solution iff the equation has at least one free variable.

Definition. Vector addition can be considered as a *translation*. e.g. the vector \mathbf{v} is *translated by* \mathbf{p} to $\mathbf{v} + \mathbf{p}$.

Definition. A *parametric vector equation* can be written as

$$\mathbf{x} = s\mathbf{u} + t\mathbf{v} \quad (s, t \in \mathbb{R})$$

which describes explicitly the spanned plane by \mathbf{u} and \mathbf{v} . Whenever a solution set is described explicitly with vectors, we say that the solution is in *parametric vector form*.

Theorem 1.5.1. Suppose the equation $\mathbf{Ax} = \mathbf{b}$ is consistent for some given \mathbf{b} , and let \mathbf{p} be a nonzero solution. Then the solution set of it is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where \mathbf{v}_h is any solution of the homogeneous equation $\mathbf{Ax} = \mathbf{0}$.

1.6 Linear Independence

Definition. An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in \mathbb{R}^n is said to be *linearly independent* if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be *linearly dependent* if there exist weights c_1, \dots, c_p , not all zero, such that

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

and this equation is called a *linear dependence relation* among $\mathbf{v}_1, \dots, \mathbf{v}_p$.

Theorem 1.6.1. *The columns of a matrix \mathbf{A} are linearly independent iff the equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ has **only** the trivial solution.*

Theorem 1.6.2. *A set of two vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent iff one of the vectors is a multiple of the other.*

Theorem 1.6.3. *An indexed set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of two or more vectors is linearly dependent iff at least one of the vectors in S is a linear combination of the others.*

Theorem 1.6.4. *Any set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in \mathbb{R}^n is linearly dependent if $p > n$ (Same as the criterion for the existence of solutions in a system of equations).*

Proof. Since $p > n$, there are more variables than equations, and therefore nontrivial solutions exist. \square

Theorem 1.6.5. *If a set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in \mathbb{R}^n contains the zero vector, then the set is linearly dependent.*

1.7 Linear Transformations

Definition. A *transformation* (or *function* or *mapping*) from \mathbb{R}^n to \mathbb{R}^m is a rule that assigns to each vector $\mathbf{x} \in \mathbb{R}^n$ a vector $T(\mathbf{x}) \in \mathbb{R}^m$. \mathbb{R}^n is called the *domain* of T , and \mathbb{R}^m is called the *codomain* of T . For $\mathbf{x} \in \mathbb{R}^n$, the vector $T(\mathbf{x}) \in \mathbb{R}^m$ is called the *image* of \mathbf{x} under T . The set of all images $T(\mathbf{x})$ is called the *range* of T .

Example 1.1. *Given a scalar r , define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$. T is called a *contraction* when $0 \leq r \leq 1$ and a *dilation* when $r > 1$.*

Theorem 1.7.1. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix \mathbf{A} such that*

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n$$

In fact, \mathbf{A} is the $m \times n$ matrix whose j th column is the vector $T(\mathbf{e}_j)$, where \mathbf{e}_j is the j th column of the identity matrix in \mathbb{R}^n .

$$\mathbf{A} = \begin{bmatrix} T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n) \end{bmatrix}$$

*The matrix \mathbf{A} is called the *standard matrix* for the linear transformation T .*

Theorem 1.7.2. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then T is injective iff the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.*

Theorem 1.7.3. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation and let \mathbf{A} be the standard matrix for T . Then*

- *T is surjective iff the columns of \mathbf{A} span \mathbb{R}^m ;*
- *T is injective iff the columns of \mathbf{A} are linearly independent.*

Definition. If there is a matrix \mathbf{A} such that

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \dots$$

then the equation above is called a *linear difference equation* (or *recurrence relation*).

2

Matrices

2.1 Matrices and Arithmetic Operations on Them

Definition. A *diagonal matrix* is a square matrix whose nondiagonal entries are zero.

Definition. Two matrices are equal if they have the same size and each entries are equal.

Definition. The *sum* of two matrices is the sum of each corresponding entries in these two matrices. Thus the sum is defined only when they have the same size.

Definition. The *scalar multiple* of a matrix has entries of the product of the scalar and each corresponding original entries.

Theorem 2.1.1. *The set of matrices of the same size with respect to matrix addition and scalar multiplication over the field of real numbers is a vector space.*

Definition. A square matrix is called *lower triangular* if all the entries above the main diagonal are zero. Similarly, a square matrix is called *upper triangular* if all the entries below the main diagonal are zero. A triangular matrix is one that is either lower triangular or upper triangular. A matrix that is both upper and lower triangular is called a *diagonal matrix*.

Definition. If \mathbf{A} is an $m \times n$ matrix, and if \mathbf{B} is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$, then the *product* \mathbf{AB} is the $m \times p$ matrix whose columns are $\mathbf{Ab}_1, \dots, \mathbf{Ab}_p$. Multiplication of matrices corresponds to composition of linear transformations.

Theorem 2.1.2. *The multiplication has the following properties:*

- *Associativity of multiplication;*

- *Left distribution;*
- *Right distribution;*
- *Associativity over scalar multiplication;*
- *Identity for matrix multiplication; i.e. If \mathbf{A} is a matrix of size $m \times n$, then*

$$\mathbf{I}_m \mathbf{A} = \mathbf{A} = \mathbf{A} \mathbf{I}_n$$

where \mathbf{I}_n is the $n \times n$ identity matrix.

Definition. In general, matrix multiplication is not commutative and the cancellation law do not hold. When two matrices' multiplication is commutative, they are said to be *commute* with one another. Also, if a product \mathbf{AB} is the zero matrix, in general it does not mean that either $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$.

Definition. If \mathbf{A} is an $m \times n$ matrix and k is a positive integer, then \mathbf{A}^k denoted the product of k copies of \mathbf{A} , i.e. the k th power of \mathbf{A} . The 0th power of a matrix is the identity matrix.

Definition. If \mathbf{A} is an $m \times n$ matrix, the *transpose* of \mathbf{A} is the $n \times m$ matrix, denoted \mathbf{A}^T , whose columns are formed from the corresponding rows of \mathbf{A} .

Theorem 2.1.3. *The transpose operation has the following properties:*

- $(\mathbf{A}^T)^T = \mathbf{A}$;
- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$;
- *Associativity with scalar multiplication;*
- $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$, that is, the transpose of a product of arbitrary number of matrices equals the product of their transpose in the reverse order.

2.2 The Inverse of a Matrix

Definition. If \mathbf{A} is an $n \times n$ matrix, then if

$$\mathbf{AA}^{-1} = \mathbf{I}_n$$

we say that \mathbf{A} is *invertible* and \mathbf{A}^{-1} an *inverse* of \mathbf{A} . The inverse of a matrix is unique. If a matrix is not invertible, it is called a *singular matrix*.

Theorem 2.2.1. *A matrix \mathbf{A} is invertible only if $\det(\mathbf{A}) \neq 0$, and in this case*

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{Adj}(\mathbf{A})$$

Theorem 2.2.2. *If \mathbf{A} is an invertible $n \times n$ matrix, then for each $\mathbf{b} \in \mathbb{R}^n$, the equation $\mathbf{Ax} = \mathbf{b}$ has the unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.*

Theorem 2.2.3. • *The inverse of the inverse of a invertible matrix is the matrix itself.*

- *The inverse of the product of arbitrary number of invertible square matrices is the product of the inverse of themselves multiplied in the reverse order.*
- *The transpose of a invertible matrix is also invertible. Moreover, the inverse of a matrix's transpose is the transpose of the matrix's inverse.*

Definition. An *elementary matrix* is a matrix obtained by performing a single elementary row operation on a identity matrix.

Theorem 2.2.4. *If an elementary row operations is performed on an $m \times n$ matrix \mathbf{A} , the resulting matrix can be written as \mathbf{EA} , where the $m \times m$ matrix \mathbf{E} is created by performing the same row operation on \mathbf{I}_m .*

Theorem 2.2.5. *Each elementary matrix \mathbf{E} is invertible. The inverse of \mathbf{E} is the elementary matrix of the same type that transforms \mathbf{E} back into \mathbf{I} .*

Theorem 2.2.6. *An $n \times n$ matrix \mathbf{A} is invertible iff \mathbf{A} is a row equivalent to \mathbf{I}_n , and in this case, any sequence of elementary row operations that reduces \mathbf{A} to \mathbf{I}_n also transforms \mathbf{I}_n into \mathbf{A}^{-1} .*

Theorem 2.2.7 (The Invertible Matrix Theorem). *Let \mathbf{A} be a square $n \times n$ matrix. Then the following statements are equivalent.*

- *\mathbf{A} is an invertible matrix.*
- *\mathbf{A} is row equivalent to the $n \times n$ identity matrix.*
- *\mathbf{A} has n pivot positions.*
- *The equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.*
- *The columns of \mathbf{A} form a linearly independent set.*
- *The linear transformation $\mathbf{x} \mapsto \mathbf{Ax}$ is injective.*
- *The equation $\mathbf{Ax} = \mathbf{b}$ has at least one solution for each \mathbf{b} in \mathbb{R}^n .*
- *The columns of \mathbf{A} span \mathbb{R}^n .*

- The linear transformation $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ maps \mathbb{R}^n onto \mathbb{R}^n .

- There is an $n \times n$ matrix \mathbf{C} such that $\mathbf{CA} = \mathbf{I}$.
- There is an $n \times n$ matrix \mathbf{D} such that $\mathbf{AD} = \mathbf{I}$.
- \mathbf{A}^T is an invertible matrix.

- The columns of \mathbf{A} form a basis of \mathbb{R}^n .

- $\text{Col } \mathbf{A} = \mathbb{R}^n$
- $\dim \text{Col } \mathbf{A} = n$
- $\text{rank } \mathbf{A} = n$
- $\text{Nul } \mathbf{A} = \{\mathbf{0}\}$
- $\dim \text{Nul } \mathbf{A} = 0$

Proposition. Let \mathbf{A} and \mathbf{B} be square matrices. If $\mathbf{AB} = \mathbf{I}$, then \mathbf{A} and \mathbf{B} are both invertible, with $\mathbf{B} = \mathbf{A}^{-1}$ and $\mathbf{A} = \mathbf{B}^{-1}$

Definition. A linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is *invertible* if there exists a function $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that

$$\begin{aligned} (\forall \mathbf{x} \in \mathbb{R}^n) \quad S(T(\mathbf{x})) &= \mathbf{x} \\ (\forall \mathbf{x} \in \mathbb{R}^n) \quad T(S(\mathbf{x})) &= \mathbf{x} \end{aligned}$$

and S is called the *inverse* of T and denoted T^{-1} .

Theorem 2.2.8. A linear transformation is invertible iff its standard matrix is invertible. In this case its inverse is unique.

Theorem 2.2.9. If \mathbf{A} is $m \times n$ and \mathbf{B} is $n \times p$, then

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \text{Col}_1(\mathbf{A}) & \text{Col}_2(\mathbf{A}) & \cdots & \text{Col}_n(\mathbf{A}) \end{bmatrix} \begin{bmatrix} \text{Row}_1(\mathbf{B}) \\ \text{Row}_2(\mathbf{B}) \\ \vdots \\ \text{Row}_n(\mathbf{B}) \end{bmatrix} \\ &= \text{Col}_1(\mathbf{A}) \text{Row}_1(\mathbf{B}) + \cdots + \text{Col}_n(\mathbf{A}) \text{Row}_n(\mathbf{B}) \end{aligned}$$

Definition. A *block matrix* is a partitioned matrix with zero blocks off the main diagonal. Such matrix is invertible iff each block on the diagonal is invertible.

Definition. A *factorization* of a matrix is an equation that expresses it as a product of two or more matrices.

Definition. An square matrix is said to be *strictly diagonally dominant* if the absolute of each diagonal entry exceeds the sum of the absolute values of the other entries in the same row.

2.3 Subspaces of \mathbb{R}^n

Definition. A *subspace* of \mathbb{R}^n is any set $H \in \mathbb{R}^n$ that has three properties:

- The zero vector is in H ;
- For each vector \mathbf{u} and \mathbf{v} in H , their sum is in H (addition is closed on H);
- For each \mathbf{u} in H and each scalar c , the vector $c\mathbf{u}$ is in H (scalar multiplication is closed on H).

Definition. The *column space* of a matrix \mathbf{A} is the set $\text{Col } \mathbf{A}$ of all linear combinations of the columns of \mathbf{A} .

Definition. The *null space* of a matrix \mathbf{A} is the set $\text{Nul } \mathbf{A}$ of all solutions to the homogeneous equation $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Theorem 2.3.1. *The null space of a $m \times n$ matrix is a subspace of \mathbb{R}^n .*

Definition. A *basis* for a subspace H of \mathbb{R}^n is a linearly independent set in H that spans H .

Example 2.1. *The standard basis for \mathbb{R}^n are vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$, where*

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Theorem 2.3.2. *The pivot columns of a matrix \mathbf{A} form a basis for the column space of \mathbf{A} .*

Definition. Suppose the set $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is the 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 $\mathbf{x} = c_1\mathbf{b}_1 + \dots + c_p\mathbf{b}_p$, 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}* .

Definition. The *dimension* of a nonzero subspace H , denoted by $\dim H$, is the number of vectors in any basis for H . The dimension of the zero subspace $\{\mathbf{0}\}$ is defined to be zero.

Definition. The *rank* of a matrix \mathbf{A} , denoted by $\text{rank } \mathbf{A}$, is the dimension of the column space of \mathbf{A} .

Theorem 2.3.3 (The Rank Theorem). *If a matrix \mathbf{A} has n columns, then $\text{rank } \mathbf{A} + \dim \text{Nul } \mathbf{A} = n$.*

Proof. The nonpivot columns correspond to the free variables in $\mathbf{Ax} = \mathbf{0}$, and since the number of pivot columns plus the number of nonpivot columns are the number of columns in the matrix, the proof completes. \square

Theorem 2.3.4 (The Basis Theorem). *Let H be a p -dimensional subspace of \mathbb{R}^n . Any linearly independent set of exactly p elements in H is automatically a basis for H . Also, any set of p elements of H that spans H is a basis for H .*

3

Determinants

3.1 Determinants and some other Concepts

Definition. The *determinant* of the matrix \mathbf{A}

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

denoted $\det \mathbf{A}$ and equals $ad - bc$. Determinant is only defined for a square matrix, but the procedure above can be repeated on higher dimension matrices, for example

$$\det \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix} = a \det \begin{bmatrix} f & g & h \\ j & k & l \\ n & o & p \end{bmatrix} - b \det \begin{bmatrix} e & g & h \\ i & k & l \\ m & o & p \end{bmatrix} + c \det \begin{bmatrix} e & f & h \\ i & j & l \\ m & n & p \end{bmatrix} - d \det \begin{bmatrix} e & f & g \\ i & j & k \\ m & n & o \end{bmatrix}$$

By the Leibniz formula for the determinant of an $n \times n$ matrix \mathbf{A} is

$$\det(\mathbf{A}) = \sum_{\sigma \in S_n} (\text{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma_i})$$

Here the sum is computed over all permutations σ of the set $\{1, 2, \dots, n\}$.

A permutation is a function that reorders this set of integers. The value in the i th position after the reordering σ is denoted by σ_i . For example, for $n = 3$, the original sequence $1, 2, 3$ might be reordered to $\sigma = [2, 3, 1]$, with $\sigma_1 = 2$, $\sigma_2 = 3$, and $\sigma_3 = 1$. The set of all such permutations (also known as the symmetric group on n elements) is denoted by S_n .

For each permutation σ , $\text{sgn}(\sigma)$ denotes the signature of σ , a value that is $+1$ whenever the reordering given by σ can be achieved by successively interchanging two entries an even number of times, and -1 whenever it can be achieved by an odd number of such interchanges.

Definition. If \mathbf{A} is a square matrix, then the *minor* of the entry in the i -th row and j -th column (also called the (i, j) *minor*, or a *first minor*) is the determinant of the submatrix formed by deleting the i -th row and j -th column. This number is often denoted $M_{i,j}$. The (i, j) *cofactor* is obtained by multiplying the minor by $(-1)^{i+j}$ and is denoted $C_{i,j}$.

In general, let \mathbf{A} be an $m \times n$ matrix and k an integer with $0 < k \leq m$, and $k \leq n$. A $k \times k$ minor of \mathbf{A} , also called minor determinant of order k of \mathbf{A} or, if $m = n$, $(n - k)$ th minor determinant of \mathbf{A} , is the determinant of a $k \times k$ matrix obtained from \mathbf{A} by deleting $m - k$ rows and $n - k$ columns.

Definition. The matrix formed by all of the cofactors of a square matrix \mathbf{A} is called the *cofactor matrix*.

Definition. The *adjugate* is the transpose of the cofactor matrix of it, that is, if \mathbf{A} is a matrix and \mathbf{C} is its cofactor matrix, then

$$\text{Adj}(\mathbf{A}) = \mathbf{C}^T$$

Theorem 3.1.1. For a invertible matrix $n \times n$ \mathbf{A}

$$\mathbf{A} \text{Adj}(\mathbf{A}) = \det(\mathbf{A})\mathbf{I}$$

or equivalently

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \text{Adj} \mathbf{A}$$

Theorem 3.1.2. The determinant of an square matrix can be computed by a cofactor expansion across any row or down any column. The expansion across the i th row is

$$\det \mathbf{A} = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$$

Theorem 3.1.3. If \mathbf{A} is a triangular matrix, then $\det \mathbf{A}$ is the product of the entries on the main diagonal of \mathbf{A} .

3.2 Properties of Determinants

Definition. An elementary matrix is called an *row replacement* if it is obtained from the identity matrix by adding a multiple of one row to another; it's called an *interchange* if it is obtained by interchanging two rows of identity; and it's called a *scale by r* if it is obtained by multiplying a row of identity by a nonzero scalar r .

Theorem 3.2.1. Let \mathbf{A} be a square matrix.

- If a multiple of one row of \mathbf{A} is added to another row to produce a matrix \mathbf{B} , then $\det \mathbf{A} = \det \mathbf{B}$.

- If two rows of \mathbf{A} are interchanged to produce \mathbf{B} , then $\det \mathbf{B} = -\det \mathbf{A}$.
- If one row of \mathbf{A} is multiplied by k to produce \mathbf{B} , then $\det \mathbf{B} = k \cdot \det \mathbf{A}$.

or, equivalently, if \mathbf{A} is an $n \times n$ matrix and \mathbf{E} is an $n \times n$ elementary matrix, then

$$\det \mathbf{EA} = (\det \mathbf{E})(\det \mathbf{A})$$

where $\det \mathbf{E}$ assumes $1, -1, r$ respectively for \mathbf{E} is a row replacement, an interchange, and a scale by r .

Theorem 3.2.2. If \mathbf{A} is an $n \times n$ matrix, then $\det \mathbf{A}^T = \det \mathbf{A}$.

Theorem 3.2.3. If \mathbf{A} and \mathbf{B} are $n \times n$ matrices, then $\det \mathbf{AB} = (\det \mathbf{A})(\det \mathbf{B})$.

Example 3.1. If all columns except one are held fixed in a square matrix, then its determinant is a linear function of that one (vector) variable.

Let $\mathbf{A}_i(\mathbf{b})$ denote the matrix obtained from \mathbf{A} by replacing column i by the vector \mathbf{b} .

Theorem 3.2.4. If \mathbf{A} is an invertible $n \times n$ matrix. For any $\mathbf{b} \in \mathbb{R}^n$, then unique solution \mathbf{x} of $\mathbf{Ax} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det \mathbf{A}_i(\mathbf{b})}{\det \mathbf{A}}, \quad i = 1, 2, \dots, n$$

Theorem 3.2.5. If \mathbf{A} is a 2×2 matrix, the area of the parallelogram determined by the columns of \mathbf{A} is $|\det \mathbf{A}|$. If \mathbf{A} is a 3×3 matrix, the volume of the parallelepiped determined by the columns of \mathbf{A} is $|\det \mathbf{A}|$.

Theorem 3.2.6. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation determined by a 2×2 matrix \mathbf{A} . If S is a parallelogram in \mathbb{R}^2 , then

$$\{\text{area of } T(S)\} = |\det \mathbf{A}| \cdot \{\text{area of } S\}$$

and similar, if in \mathbb{R}^3 S is a parallelepiped, then

$$\{\text{volume of } T(S)\} = |\det \mathbf{A}| \cdot \{\text{volume of } S\}$$

These conclusions hold whenever S has finite area or finite volume.

4

Vector Spaces

Definition. A *vector space* over a field F is a set V with two closed operations, *vector addition* or simply *addition* and *scalar multiplication*, that satisfy the following axioms:

1. Associativity of addition;
2. Commutativity of addition;
3. Identity element of addition;
4. Inverse elements of addition;
5. Compatibility of scalar multiplication with field multiplication;

$$a(b\mathbf{v}) = (ab)\mathbf{v}$$

6. Identity element of scalar multiplication;
7. Distributivity of scalar multiplication with respect to vector addition;
8. Distributivity of scalar multiplication with respect to field addition.

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

- The zero vector of V is in H .
- H is closed under vector addition.
- H is closed under multiplication by scalars.

If a subspace only contains the zero vector $\mathbf{0}$, it is called a *zero subspace* and written as $\{\mathbf{0}\}$.

Theorem 4.0.1. If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in a vector space V , then $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a subspace of V and is called the subspace spanned (or generated) by $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$. Given any subspace H of V , a spanning (or generating) set for H is a set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in H such that $H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

Definition. The null space of an $m \times n$ matrix \mathbf{A} , written as $\text{Nul } \mathbf{A}$, is the set of all solutions to the homogeneous equation $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Theorem 4.0.2. The null space of an $m \times n$ matrix \mathbf{A} is a subspace of \mathbb{R}^n .

Definition. The column space of an $m \times n$ matrix, written as $\text{Col } \mathbf{A}$, is the set of all linear combinations of the columns of \mathbf{A} .

Theorem 4.0.3. The column space of an $m \times n$ matrix is a subspace of \mathbb{R}^m .

Theorem 4.0.4. The column space of an $m \times n$ matrix \mathbf{A} is all of \mathbb{R}^m iff the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b} \in \mathbb{R}^m$.

Definition. For a linear transformation T from a vector space V into a vector space W , the kernel (or null space) of T is the set of all $\mathbf{u} \in V$ such that $T(\mathbf{u}) = \mathbf{0}$. The range of T is the set of all vectors in W of the form $T(\mathbf{x})$ for some $\mathbf{x} \in V$. If T can be written as a matrix transformation, then the kernel and the range of T are just the null space and the column space of that matrix. Kernel is a subspace of V , and range is a subspace of W .

Definition. An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\} \in V$ is said to be linearly independent if the vector equation

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

has only the trivial solution.

Theorem 4.0.5. An indexed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of two or more vectors, with $\mathbf{v}_1 = \mathbf{0}$, is linearly dependent iff some \mathbf{v}_j with $j > 1$ is a linear combination of the preceding vectors, $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$.

Definition. Let H be a subspace of a vector space V . An indexed set of vectors $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ in V is a basis for H if

1. \mathcal{B} is a linearly independent set;
2. the subspace spanned by \mathcal{B} coincides with H .

Theorem 4.0.6. Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be a set in V and let $H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

1. If one of the vectors in S is a linear combination of the remaining vectors in S , then the set formed from S by removing this vector still spans H .
2. If $H \neq \{\mathbf{0}\}$, some subset of S is a basis for H .

Theorem 4.0.7. The pivot columns of a matrix \mathbf{A} form a basis for $\text{Col } \mathbf{A}$.

4.1 Coordinate Systems

Theorem 4.1.1. *Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ be a basis for a vector space V . Then for each $\mathbf{x} \in V$, there exists a unique set of scalars c_1, \dots, c_n such that*

$$\mathbf{x} = c_1 \mathbf{b}_1 + \dots + c_n \mathbf{b}_n$$

Definition. Suppose the set $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for V and $\mathbf{x} \in V$. The *coordinate of \mathbf{x} relative to the basis \mathcal{B}* (or the *\mathcal{B} -coordinates of \mathbf{x}*) are the weights c_1, \dots, c_n such that $\mathbf{x} = c_1 \mathbf{b}_1 + \dots + c_n \mathbf{b}_n$.

The vector

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

is the *coordinate vector of \mathbf{x} (relative to \mathcal{B})*, or the *\mathcal{B} -coordinate vector of \mathbf{x}* . The mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is the *coordinate mapping (determined by \mathcal{B})*.

Definition. The matrix

$$\mathbf{P}_{\mathcal{B}} = [\mathbf{b}_1, \dots, \mathbf{b}_p]$$

is called the *change-of-coordinates matrix* from \mathcal{B} to the standard basis in \mathbb{R}^n , since for a vector $\mathbf{x} = c_1 \mathbf{b}_1 + \dots + c_n \mathbf{b}_n$ we obtain the relationship

$$\mathbf{x} = \mathbf{P}_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$$

Theorem 4.1.2. *Let \mathcal{B} be a basis for a vector space V . Then the coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is an injective linear transformation from V into \mathbb{R}^n .*

In general, an injective linear transformation from a vector space V onto another vector space W is called an *isomorphism* from V onto W .

Theorem 4.1.3. *If a vector space V has a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, then any set in V containing more than n vectors must be linearly dependent.*

Theorem 4.1.4. *If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.*

Definition. If 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 . If V is not spanned by a finite set, then V is said to be *infinite-dimensional*.

Theorem 4.1.5. *Let H be a subspace of a finite-dimensional vector space V . Any linearly independent set in H can be expanded, if necessary, to a basis for H . Also, H is finite-dimensional and*

$$\dim H \leq \dim V$$

Theorem 4.1.6 (The Basis Theorem). *Let V be a p -dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in V is automatically a basis for V . Any set of exactly p elements that spans V is a basis for V .*

The dimension of $\text{Nul } \mathbf{A}$ is the number of free variables in $\mathbf{A}\mathbf{x} = \mathbf{0}$, and the dimension of $\text{Col } \mathbf{A}$ is the number of pivot columns in \mathbf{A} .

4.2 Rank

Definition. The set of all linear combinations of the row vectors in \mathbf{A} is called the *row space* of \mathbf{A} and denoted $\text{Row } \mathbf{A}$.

Theorem 4.2.1. *If two matrices \mathbf{A} and \mathbf{B} are row equivalent, then their row spaces are the same. If \mathbf{B} is in echelon form, the nonzero rows of \mathbf{B} form a basis for the row space of \mathbf{A} as well as \mathbf{B} .*

Definition. The *rank* of \mathbf{A} is the dimension of the column space of \mathbf{A} .

Theorem 4.2.2 (The Rank Theorem). *The dimensions of the column space and the row space of an $m \times n$ matrix \mathbf{A} are equal. This common dimension, the rank of \mathbf{A} , also equals the number of pivot positions in \mathbf{A} and satisfies the equation*

$$\text{rank } \mathbf{A} + \dim \text{Nul } \mathbf{A} = n$$

4.3 Change of Basis

Theorem 4.3.1. *Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be bases of a vector space V . Then there is an $n \times n$ matrix $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$, called the *change-of-coordinates matrix* from \mathcal{B} to \mathcal{C} , such that*

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$$

The columns of $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$ are the \mathcal{C} -coordinate vectors of the vectors in the basis \mathcal{B} , that is

$$\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix}$$

Because the columns of this matrix are linearly independent, since they are the coordinate vectors of the linearly independent set \mathcal{B} , it follows that $\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}}$ is invertible, and we have

$$(\mathbf{P}_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = \mathbf{P}_{\mathcal{B} \leftarrow \mathcal{C}}$$

5

Eigenvalues and Eigenvectors

5.1 Definition

Definition. An *eigenvector* of an $n \times n$ matrix \mathbf{A} is a nonzero vector \mathbf{x} such that $\mathbf{Ax} = \lambda\mathbf{x}$ for some scalar λ . A scalar λ is called an *eigenvalue* of \mathbf{A} if there is a nontrivial solution \mathbf{x} of $\mathbf{Ax} = \lambda\mathbf{x}$; such an \mathbf{x} is called an *eigenvector corresponding to λ* .

Definition. The set of all solutions of

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

is a subspace of \mathbb{R}^n and is called the *eigenspace* of \mathbf{A} corresponding to λ .

Theorem 5.1.1. *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Theorem 5.1.2 (The Invertible Matrix Theorem). *Let \mathbf{A} be an $n \times n$ matrix. Then \mathbf{A} is invertible iff the number 0 is not an eigenvalue of \mathbf{A} .*

Theorem 5.1.3. *If $\mathbf{v}_1, \dots, \mathbf{v}_r$ are eigenvectors that corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_r$ of an $n \times n$ matrix \mathbf{A} , then the set $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is linearly independent.*

Definition. The scalar equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ is called the *characteristic equation* of \mathbf{A} . If \mathbf{A} is an $n \times n$ matrix, then $\det(\mathbf{A} - \lambda\mathbf{I})$ is a polynomial of degree n called the *characteristic polynomial* of \mathbf{A} .

Theorem 5.1.4. *A scalar λ is an eigenvalue of an $n \times n$ matrix \mathbf{A} iff λ satisfies the characteristic equation.*

Definition. If \mathbf{A} and \mathbf{B} are $n \times n$ matrices, then \mathbf{A} and \mathbf{B} are *similar* if there is an invertible matrix \mathbf{P} such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}$. Changing \mathbf{A} into $\mathbf{P}^{-1}\mathbf{AP}$ is called a *similarity transformation*.

Theorem 5.1.5. *If $n \times n$ matrices \mathbf{A} and \mathbf{B} are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).*

5.2 Diagonalization

Definition. A square matrix \mathbf{A} is said to be *diagonalizable* if \mathbf{A} is similar to a diagonal matrix.

Theorem 5.2.1. *An $n \times n$ matrix \mathbf{A} is diagonalizable iff \mathbf{A} has n linearly independent eigenvectors. In other words, \mathbf{A} is diagonalizable iff there are enough eigenvectors to form a basis of \mathbb{R}^n , and such basis is called an eigenvector basis.*

Theorem 5.2.2. *An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.*

Theorem 5.2.3. *Let \mathbf{A} be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \dots, \lambda_p$.*

1. *For $1 \leq k \leq p$, the dimension of the eigenspaces for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k .*
2. *The matrix \mathbf{A} is diagonalizable iff the sum of the dimensions of the distinct eigenspaces equals n , and this happens iff the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k .*
3. *If \mathbf{A} is diagonalizable and \mathcal{B}_k is a basis for the eigenspace corresponding to λ_k for each k , then the total collection of vectors in the sets $\mathcal{B}_1, \dots, \mathcal{B}_p$ forms an eigenvector basis for \mathbb{R}^n .*

Definition. The matrix

$$\mathbf{M} = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix}$$

where $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for the vector space V , \mathcal{C} is a basis in W , and T is a linear transformation from V to W , is called the *matrix for T relative to the bases \mathcal{B} and \mathcal{C}* . If W is the same as V and the basis \mathcal{C} is the same as \mathcal{B} , the matrix \mathbf{M} is called the *matrix for T relative to \mathcal{B}* or the *\mathcal{B} -matrix for T* , and denoted $[T]_{\mathcal{B}}$.

Theorem 5.2.4 (Diagonal Matrix Representation). *Suppose $\mathbf{A} = \mathbf{PDP}^{-1}$, where \mathbf{D} is a diagonal $n \times n$ matrix. If \mathcal{B} is the basis for \mathbb{R}^n formed from the columns of \mathbf{P} , then \mathbf{D} is the \mathcal{B} -matrix of the transformation $\mathbf{x} \mapsto \mathbf{Ax}$.*

5.3 Complex Eigenvalues

The theory of eigenvalues and eigenvectors developed for \mathbb{R}^n applies equally well on \mathbb{C}^n .

Theorem 5.3.1. *Let \mathbf{A} be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and associated eigenvector $\mathbf{v} \in \mathbb{C}^2$. Then*

$$\mathbf{A} = \mathbf{PCP}^{-1}, \quad \text{where} \quad \mathbf{P} = \begin{bmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

6

Orthogonality

Definition. If a vector \mathbf{z} is orthogonal to every vector in a subspace W of \mathbb{R}^n , then \mathbf{z} is said to be *orthogonal* to W . The set of all vectors \mathbf{z} that are orthogonal to W is called the *orthogonal complement* of W and denoted by W^\perp .

Theorem 6.0.1. 1. A vector \mathbf{x} is in W^\perp iff \mathbf{x} is orthogonal to every vector in a set that spans W .

2. W^\perp is a subspace of \mathbb{R}^n .

Proof. Left for Exercise □

Theorem 6.0.2. Let \mathbf{A} be an $m \times n$ matrix. Then the orthogonal complement of the row space of \mathbf{A} is the nullspace of \mathbf{A} , and the orthogonal complement of the column space of \mathbf{A} is the nullspace of \mathbf{A}^T :

$$(\text{Row } \mathbf{A})^\perp = \text{Nul } \mathbf{A}, \quad (\text{Col } \mathbf{A})^\perp = \text{Nul } \mathbf{A}^T$$

6.1 Orthogonal Sets

Definition. A set of vectors in \mathbb{R}^n is said to be an *orthogonal set* if each pair of distinct vectors from the set is orthogonal.

Theorem 6.1.1. 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 .

Definition. An *orthogonal basis* for a subspace W of \mathbb{R}^n is a basis for W that is also an orthogonal set.

Theorem 6.1.2. Each vector in a subspace of \mathbb{R}^n has a unique representation as a linear combination of its orthogonal basis.

Definition. The *orthogonal projection* of \mathbf{v} on an arbitrary non-zero vector \mathbf{b} can be written as:

$$\text{proj}_{\mathbf{b}} \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{b}}{\|\mathbf{b}\|^2} \mathbf{b} \quad (6.1)$$

Moreover, we can see that $\mathbf{v} - \text{proj}_{\mathbf{b}} \mathbf{v}$ is the vector component of \mathbf{v} orthogonal to \mathbf{b} . The projection $\text{proj}_{\mathbf{b}} \mathbf{v}$ is determined by the subspace $\text{Span}\{\mathbf{b}\}$, and we may call it the *orthogonal projection onto* $\text{Span}\{\mathbf{b}\}$.

Definition. A set is an *orthonormal set* if it is an orthogonal set of unit vectors. It is also an *orthonormal basis* for a subspace spanned by it.

Theorem 6.1.3. An $m \times n$ matrix \mathbf{U} has orthonormal columns iff $\mathbf{U}^T \mathbf{U} = \mathbf{I}$.

Theorem 6.1.4. Let \mathbf{U} be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then

1. $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$.
2. $\|\mathbf{U}\mathbf{x}\| \cdot \|\mathbf{U}\mathbf{y}\| = \mathbf{x} \cdot \mathbf{y}$, and it equals zero iff $\mathbf{x} \cdot \mathbf{y} = 0$.

Equivalently, they say that the linear mapping $\mathbf{x} \mapsto \mathbf{U}\mathbf{x}$ preserves lengths and orthogonality.

Definition. An *orthogonal matrix* is a square invertible matrix \mathbf{U} such that $\mathbf{U}^{-1} = \mathbf{U}^T$.

6.2 Orthogonal Projections

Theorem 6.2.1. Let W be a subspace of \mathbb{R}^n . Then each $\mathbf{y} \in \mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ 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

$$\hat{\mathbf{y}} = \text{proj}_{\mathbf{u}_1} \mathbf{y} + \dots + \text{proj}_{\mathbf{u}_p} \mathbf{y}$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

The vector $\hat{\mathbf{y}}$ is called the *orthogonal projection of \mathbf{y} onto W* and written as $\text{proj}_W \mathbf{y}$.

Theorem 6.2.2. Let W be a subspace of \mathbb{R}^n , $\mathbf{y} \in \mathbb{R}^n$, and $\hat{\mathbf{y}}$ the orthogonal projection of \mathbf{y} onto W . Then $\hat{\mathbf{y}}$ is the closest point in W to \mathbf{y} , in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all $\mathbf{v} \in W$ distinct from $\hat{\mathbf{y}}$.

Theorem 6.2.3. *If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace $W \in \mathbb{R}^n$, then*

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \dots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p$$

If $\mathbf{U} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_p]$, then

$$\text{proj}_W \mathbf{y} = \mathbf{U}\mathbf{U}^T \mathbf{y}$$

7

Symmetric Matrices and Quadratic Forms

Definition. A *symmetric matrix* is a matrix such that it equals to the transpose of itself.

Definition. A *quadratic form* on \mathbb{R}^n is a function Q defined on \mathbb{R}^n whose value at a vector $\mathbf{x} \in \mathbb{R}^n$ can be computed by an expression of the form $Q(\mathbf{x}) = \mathbf{x}\mathbf{A}^T\mathbf{x}$, where \mathbf{A} is an $n \times n$ symmetric matrix. The matrix \mathbf{A} is called the *matrix of the quadratic form*. The simplest example of a nonzero quadratic form is where the matrix of the quadratic form is the $n \times n$ identity matrix.

Definition. A quadratic form Q is:

1. *positive definite* if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$,
2. *negative definite* if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
3. *indefinite* if $Q(\mathbf{x})$ assumes both positive and negative values.

Theorem 7.0.1. Let \mathbf{A} be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}\mathbf{A}^T\mathbf{x}$ is:

1. *positive definite* iff the eigenvalues of \mathbf{A} are all positive,
2. *negative definite* iff the eigenvalues of \mathbf{A} are all negative,
3. *indefinite* iff the eigenvalues of \mathbf{A} has both positive and negative eigenvalues.

Definition. A *positive definite matrix* \mathbf{A} is a symmetric matrix for which the quadratic form is positive definite. The matrix is a *positive semidefinite matrix* if its quadratic form is nonnegative. Other terms are defined analogously.