Linear Algebra

Hechen Hu

December 2, 2017

Contents

1

Solving Linear Equations

1.1 Systems of Linear Equations

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

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

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 $x_1, \cdots, x_n$ 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*

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

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 $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \cdots, c_p$, 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* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ *using weights* $c_1, c_2, \cdots, c_p$.

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

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

with $c_1, c_2, \cdots, c_p$ scalars.

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

Definition. If $\mathbf{A}$ is an $m \times n$ matrix, with columns $a_1, \cdots, a_n$, 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,

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

$\mathbf{Ax}$ 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* $a_1, \cdots, a_n$, *and* $\mathbf{b}$ *is in* $\mathbb{R}^m$, *the matrix equation*

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

has the same solution set as the vector equation

$$x_1a_1 + x_2a_1 + \cdots + x_na_n = \mathbf{b}$$

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

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

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

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

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

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 $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be *linearly independent* if the vector equation

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

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

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

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

Theorem 1.6.1. *The columns of a matrix $\mathbf{A}$ are linearly independent iff the equation $\mathbf{Ax} = \mathbf{0}$ 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 $S = \{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ 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 $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$(Same as the criterion for the existence of solutions in a system of equations).*

Theorem 1.6.5. *If a set $S = \{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ 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 $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$. T is called a contraction when $0 \leqslant r \leqslant 1$ and a dilation when $r > 1$.*

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

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

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

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

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

Theorem 1.7.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ 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 $T : \mathbb{R}^n \to \mathbb{R}^m$ 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

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

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

2

Vector Spaces and Subspaces

3

Orthogonality

4

Determinants

5

Eigenvalues and Eigenvectors

6

The Singular Value Decomposition(SVD)

7

Linear Transformations

8

Complex Vectors and Matrices