Linear Equations in Linear Algebra

1 Definitions and Terms

1.1 Systems of Linear Equations

A **linear equation** in the variables x_1, x_2, ..., x_n is an equation that can be written in the form $a_1x_1 + a_2x_2 + \ldots a_nx_n = b$, where a_1, ..., a_n are the **coefficients**. A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations involving the same variables. A **solution** of a linear system is a list of numbers that makes each equation a true statement. 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. A linear system is said to be consistent, if it has either one solution or infinitely many solutions. A system is inconsistent if it has no solutions.

1.2 Matrices

The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**. A matrix containing only the coefficients of a linear system is called the **coefficient matrix**, while a matrix also including the constant at the end of a linear equation, is called an **augmented matrix**. The **size** of a matrix tells how many columns and rows it has. An **m** $\times$ **n matrix** has m rows and n columns.

There are three elementary row operations. **Replacement** adds to one row a multiple of another. **Interchange** interchanges two rows. **Scaling** multiplies all entries in a row by a nonzero constant. Two matrices are **row equivalent** if there is a sequence of row operations that transforms one matrix into the other. If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

1.3 Matrix Types

A **leading entry** of a row is the leftmost nonzero entry in the row. A rectangular matrix is in **echelon form** (and thus called an **echelon matrix**) if all nonzero rows are above any rows of all zeros, if each leading entry of a row is in a column to the right to the leading entry of the row above it, and all entries in a column below a leading entry are zeros. A matrix in echelon form is in **reduced echelon form** if also the leading entry in each nonzero row is 1, and each leading 1 is the only nonzero entry in its column. If a matrix A is row equivalent to an echelon matrix U, we call U **an echelon form of** A.

A **pivot position** in a matrix A is a location in A the corresponds to a leading 1 in the reduced echelon form of A. A **pivot column** is a column of A that contains a pivot position. Variables corresponding to pivot columns in the matrix are called **basic variables**. The other variables are called **free variables**. A **general solution** of a linear system gives an explicit description of all solutions.

1.4 Vectors

A matrix with only one column is called a **vector**. Two vectors are **equal** if, and only if, their corresponding entries are equal. A vector whose entries are all zero is called the **zero vector**, and is denoted by $\mathbf{0}$. 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 Span$\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$ and is called the **subset of $\mathbb{R}^n$ spanned by $\mathbf{v_1}, \ldots, \mathbf{v_p}$**. So Span$\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$ is the collection of all vectors that can be written in the form $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_p\mathbf{v_p}$ with $c_1, c_2, \ldots, c_p$ scalars.

1.5 Matrix Equations

If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \ldots, \mathbf{a_n}$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of A and $\mathbf{x}$**, denoted by $A\mathbf{x}$, is the **linear combination of the columns of A using the corresponding entries in $\mathbf{x}$ as weights**. That is, $A\mathbf{x} = [\mathbf{a_1} \;\; \mathbf{a_2} \;\; \ldots \;\; \mathbf{a_n}]\,\mathbf{x} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n}$. $A\mathbf{x}$ is a vector in $\mathbb{R}^m$. An equation in the form $A\mathbf{x} = b$ is called a **matrix equation**.

I is called an **identity matrix**, and has 1's on the diagonal and 0's elsewhere. I_n is the identity matrix of size $n \times n$. It is always true that $I_n\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^n$.

1.6 Solution Sets of Linear Systems

A system of linear equations is said to be **homogeneous** if it can be written in the form $A\mathbf{x} = \mathbf{0}$. Such a system always has the solution $\mathbf{x} = \mathbf{0}$, which is called the trivial solution. The important question is whether there are **nontrivial solutions**, that is, a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = 0$. The total set of solutions can be described by a **parametric vector equation**, which is in the form $\mathbf{x} = a_1\mathbf{u_1} + a_2\mathbf{u_2} + \ldots + a_n\mathbf{u_n}$.

1.7 Linear Independence

An indexed set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation $x_1\mathbf{v_1} + x_2\mathbf{v_2} + \ldots + x_p\mathbf{v_p} = \mathbf{0}$ has only the trivial solution. The set is said to be linearly dependent if there exist weights $c_1, c_2, \ldots, c_p$, not all zero, such that $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_p\mathbf{v_p} = \mathbf{0}$. This equation is called a **linear dependence relation** among $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}$. Also, the columns of a matrix A are linearly independent if, and only if, the equation $A\mathbf{x} = 0$ has only the trivial solution.

1.8 Linear Transformations

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

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ 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 $\mathbb{R}^n$. That is, if the range and the codomain coincide. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ 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 $\mathbb{R}^n$. If a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is both onto $\mathbb{R}^m$ and one-to-one, then for every $\mathbf{b}$ in $\mathbb{R}^m$ $A\mathbf{x} = \mathbf{b}$ has a unique solution. That is, there is exactly 1 $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$.

2 Theorems

1. Each matrix is row equivalent to one, and only one, reduced echelon matrix.
2. A linear system is consistent if, and only if the rightmost column of the augmented matrix is not a pivot column.
3. If a linear system is consistent, and if there are no free variables, there exists only 1 solution. If there are free variables, the solution set contains infinitely many solutions.
4. A vector equation $x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n} = \mathbf{b}$ has the same solution set as the linear system whose augmented matrix is $[\mathbf{a_1}\ \ \mathbf{a_2}\ \ \ldots\ \ \mathbf{a_n}\ \ \mathbf{b}]$.
5. A vector $\mathbf{b}$ is in $\text{Span}\{\mathbf{v_1},\ \ldots,\ \mathbf{v_p}\}$ if, and only if the linear system with augmented matrix $[\mathbf{v_1}\ \ \mathbf{v_2}\ \ \ldots\ \ \mathbf{v_p}\ \ \mathbf{b}]$ has a solution.
6. If A is an $m \times n$ matrix, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution set as the linear system whose augmented matrix is $[\mathbf{a_1}\ \ \mathbf{a_2}\ \ \ldots\ \ \mathbf{a_n}\ \ \mathbf{b}]$.
7. The following four statements are equivalent for a particular $m \times n$ coefficient matrix A. That is, if one is true, then all are true, and if one is false, then all are false:
   (a) For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
   (b) Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
   (c) The columns of A span $\mathbb{R}^m$.
   (d) A has a pivot position in every row.
8. The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if, and only if the equation has at least one free variable.
9. If the reduced echelon form of A has d free variables, then the solution set consists of a d-dimensional plane (that is, a line is a 1-dimensional plane, a plane is a 2-dimensional plane), which can be described by the parametric vector equation $\mathbf{x} = a_1\mathbf{u_1} + a_2\mathbf{u_2} + \ldots + a_d\mathbf{u_d}$.
10. If $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and if $A\mathbf{p} = \mathbf{b}$, then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors $\mathbf{w} = \mathbf{p} + \mathbf{v}$ where $\mathbf{v}$ is any solution of $A\mathbf{x} = \mathbf{0}$.
11. A indexed set $S = \{\mathbf{v_1},\ \mathbf{v_2},\ \ldots,\ \mathbf{v_p}\}$ is linearly dependent if, and only if at least one of the vectors in S is a linear combination of the others.
12. If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v_1},\ \mathbf{v_2},\ \ldots,\ \mathbf{v_p}\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.
13. If a set $S = \{\mathbf{v_1},\ \mathbf{v_2},\ \ldots,\ \mathbf{v_p}\}$ contains the zero vector $\mathbf{0}$, then the set is linearly dependent.
14. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then there exists a unique matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. In fact, $A = [\ T(\mathbf{e_1})\ \ T(\mathbf{e_2})\ \ \ldots\ \ T(\mathbf{e_n})\ ]$.
15. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and $T(\mathbf{x}) = A\mathbf{x}$, then:
    (a) T is one-to-one if, and only if the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.
    (b) T is one-to-one if, and only if the columns of A are linearly independent.
    (c) T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if, and only if the columns of A span $\mathbb{R}^m$.
16. If A and B are equally sized square matrices, and $AB = I$, then A and B are both invertible, and $A = B^{-1}$ and $B = A^{-1}$.

3 Calculation Rules

3.1 Vectors

Define the vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ as follows:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \tag{1}$$

If c is a scalar, then the following rules apply:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \tag{2}$$

$$c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix} \tag{3}$$

3.2 Matrices

The product of a matrix A with size $m \times n$ and a vector $\mathbf{x}$ in $\mathbb{R}^n$ is defined as:

$$A\mathbf{x} = [\mathbf{a_1} \;\; \mathbf{a_2} \;\; \ldots \;\; \mathbf{a_n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n} \tag{4}$$

Now the following rules apply:

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \tag{5}$$

$$A(c\mathbf{u}) = c(A\mathbf{u}) \tag{6}$$

3.3 Linear Transformations

If a transformation (or mapping) T is linear, then:

$$T(\mathbf{0}) = \mathbf{0} \tag{7}$$

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{8}$$

Or, more general:

$$T(c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_p\mathbf{v_p}) = c_1T(\mathbf{v_1}) + c_2T(\mathbf{v_2}) + \ldots + c_pT(\mathbf{v_p}) \tag{9}$$