1.1 Systems of Linear Equations

Linear Equations and Solutions

<u>**Definition:**</u> A **linear equation** in the variables $x_1, x_2, \dots, x_n$ is an equation that can be written in the form

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

where b and the **coefficients** $a_1, a_2, \dots, a_n$ are real or complex numbers, usually fixed and known in advance.

Note: The subscript n may be any positive integer. In textbook examples and exercises, n is normally between 2 and 5. In real-life problems n might be 50 or 5000, or even larger.

Examples:

<u>**Definition:**</u> A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations in the same variables—say, $x_1, x_2, \dots, x_n$.

A **solution** of a system is a list $(s_1, s_2, \dots, s_n)$ of n numbers [called an n-tuple] such that each equation is a TRUE statement when the values $s_1, s_2, \dots, s_n$ are substituted for $x_1, x_2, \dots, x_n$.

Examples:

Definition: The set of all possible solutions of a system of linear equations is called the **solution set** (of the linear system). Two linear systems are called **equivalent** if they have the same solution set. That is, each solution of the first system is a solution of the second system, and each solution of the second system is a solution of the first.

Note: Since, a **solution** of a system is the n-tuple such that each equation is a TRUE, the solution set is the list of n-tuples such that every equation is TRUE.

Examples:

Classifying Solution Sets

As we saw above, a system of linear equations may have:

1. **no** solution, or
2. exactly **one** solution, or
3. **infinitely many** solutions.

Definition: When a system of linear equations has no solutions, it is said to be **inconsistent**. Otherwise, if a linear system has one or infinitely many solutions, it is said to be **consistent.**

Matrices

Graphing is an appropriate tool when working with few variables and few equations. But we quickly realize the need for an efficient method to use when $n \gg 2$.

Definition: The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**. Given the system

$$\begin{array}{rrrl} x_1 & -2x_2 & +x_3 & = 0 \\ & 2x_2 & -8x_3 & = 8 \\ 5x_1 & & -5x_3 & = 10 \end{array}$$

with the coefficients of each variable aligned in columns, the matrix

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}$$

is called the **coefficient matrix** (or **matrix of coefficients**) of the system, and the matrix

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

is called the augmented matrix of the system.

Definition: The **size** of a matrix tells how many rows and columns it has (in that order). The augmented matrix above has 3 rows and 4 columns and is called a 3×4 (read "3 by 4") matrix. If m and n are positive integers, an $\boldsymbol{m} \times \boldsymbol{n}$ **matrix** is a rectuangular array of numbers with m rows and n columns (the number of rows comes first).

Examples:

Solving a Linear System with Matrices

Three basic operations are used to simplify a linear system: (1) replace one equation by the sum of itself and a multiple of another equation, (2) interchange two equations, and (3) multiply all the terms in an equation by a nonzero constant.

Definition: These three operations correspond to the following **row operations** on augmented matrices:

1. Replacement: Replace on row by the sum of itself and a multiple of another.
2. Interchange: Interchange or swap two rows.
3. Scaling: Multiply all entries in a row by a nonzero constant.

Note 1: Another way to think of *replacement* is to "add to one row a multiple of another row".

Note 2: Row operations are *reversible*: If two rows are interchanged, they can be interchanged again to revert it to the original matrix. If a row is scaled by a non-zero constant c, it can be scaled by $\frac{1}{c}$ to revert it to the original row. If a row is replaced with its sum and the multiple of another row, it can be reverted to the original row via the inverse operations (e.g., $+/-$; $\times/\div$).

Definition: Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms on matrix into the other. If two matrices A and B are row equivalent, we denote this $A \sim B$.

Theorem 0: Equivalence of Solution Sets

Theorem 0:

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

Example:

Existence and Uniqueness of Solutions

Answers to the following two questions will determine the nature of the solutions set for a linear system:

1. Existence: Is the system consistent; that is, does at least one solution *exist*?
2. Uniqueness: If a solution exists, is it the *only* one; that is, is the solution *unique*?

Examples:

Examples (cont.):

1.2 Row Reduction and Echelon Forms

Echelon Form

Definition: A **leading entry** of a row refers to the leftmost nonzero entry (in a nonzero row).

Definition: A rectangular matrix is in **echelon form** (or **row echelon form**) if it has the following three properties:

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

Definition: If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form**):

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

Definition: A(n) **(reduced) echelon matrix** is one that is in **(reduced) echelon form**.

Examples:

Note: Any nonzero matrix may be row reduced (that is, transformed by elementary row operations) into more than one matrix in echelon form, using different sequences of row operations.

Theorem 1: Uniqueness of the Reduced Echelon Form

Theorem 1:
Each matrix is row equivalent to one and only one reduced echelon matrix.

Definition: If a matrix A is row equivalent to an echelon matrix U, we call U *an* **echelon form of A** (or row echelon form of A); if U is in reduced echelon form, we call U *the* **reduced echelon form of A**.

Example:

Definition: A **pivot position** in a matrix A is a location in A that 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.

The Row Reduction Algorithm

Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form:

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

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 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Step 5
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.

Gauss-Jordan Elimination: The Row Reduction Algorithm

Algorithm: The process of producing the unique reduced echelon form of a matrix A is obtained by the following five-step 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 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Step 5: 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.

Note: The combination of steps 1-4, called the **forward phase** of the row reduction algorithm, is also known as **Gaussian Elimination**. Step 5, which produces the unique reduced echelon form, is called the **backward phase**.

Solutions of Linear Systems

The row reduction algorithm leads directly to an explicit description of the solution set of a linear system when the algorithm is applied to the augmented matrix of the system.

Suppose for example, that the augmented matrix of a linear system has been changed into the equivalent *reduced* echelon form

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There are three variables because the augmented matrix has four columns. The associated system of equations is

$$\begin{aligned} x_1 \qquad -5x^3 &= 1 \\ x_2 \quad +x^3 &= 4 \\ 0 &= 0 \end{aligned}$$

Definition: The variables x_1 and x_2 corresponding to pivot columns in the matrix are called **basic variables.** The other variable, x_3, is called a **free variable.**

Whenever a system is consistent, the solution set can be described explicitly by solving the *reduced* system of equations for the basic variables in terms of the free variables:

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3 \text{ is free} \end{cases}$$

Note 1: This operation is possible because the reduced echelon form places each basic variable in one and only one equation.

Note 2: "x_3 is free" means that you are free to choose any value for x_3. Once that is done, the formulae for x_1 and x_2 determine their corresponding values.

Note 3: Each different choice of x_3 determines a (different) solution of the system, and every solution of the system is determined by a choice of x_3.

Note 4: Whenever a system is consistent and has free variables, we can think of the free variables as "parameters." Solving a system amounts to finding a parametric description of the solution set or determining that the solution set is empty. Thus, when the solution set is described via parametric equations, we call that description a **parametric descriptions of solution sets.**

Example:

Back-Substitution vs. Reduced Row Echelon Form

Consider the following system, whose augment matrix is in echelon form but is not in reduced echelon form:

$$\begin{array}{rrrrrcr} x_1 & -7x_2 & +2x_3 & -5x_4 & +8x_5 & = & 10 \\ & x_2 & -3x_3 & +3x_4 & +\ x_5 & = & -5 \\ & & & x_4 & -\ x_5 & = & 4 \end{array}$$

We can proceed one of two ways:

1. We can proceed to the backward phase (Step 5) of the Gauss-Jordan Elimination Algorithm to find the reduced row echelon form of the matrix.
2. We can use back substitution.

Example:

<u>Note:</u> Both processes have the same number of arithmetic operations. But matrix format substantially reduces the likelihood of errors during hand computation. The best strategy is to use only the *reduced* echelon form to solve a system!

Existence and Uniqueness Questions (Revisited)

In Section 1.1, we asked two questions about linear systems:

1. Does at least one solution exist?
2. If there is at least one solution, is it a unique solution or not?

Once we have the augmented matrix of a system in nonreduced echelon form (a.k.a., the result of Step 4, the result of Gaussian Elimination, or the result of the forward phase), we can easily answer the questions above.

Example:

Theorem 2: Existence and Uniqueness Theorem

<u>Theorem 2:</u>
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

$$[0 \quad \dots \quad 0 \quad b]$$

with b nonzero. If a linear system is consistent, then the solution set contains either

(i) a **unique** solution, where there are no free variables, or
(ii) infinitely many solutions, where there is at least one free variable.

<u>Note:</u> The statement of Theorem 2 says "if *an* echelon form of the augmented matrix has no row…". That means if, at any point during Gaussian Elimination on an augmented matrix, you encounter a row of the form $[0 \quad \dots \quad 0 \quad b]$, the system is inconsistent.

Using Row Reduction to Solve a Linear System

The following procedure outlines how to find and describe all solutions of 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 (End of Gaussian Elimination; Step 4). Decide whether the system is consistent. If there is no solution, stop; otherwise, go to the next step.
3. Continue Gauss-Jordan Elimination (row reduction) to obtain the (unique) 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.

1.3 Vector Equations

Vectors and Operations in $\mathbb{R}^2$

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

Definition: The set of all vectors with two real number entries is denoted by $\mathbb{R}^2$ (read "**r-two**"). $\mathbb{R}$ stands for the real numbers that appear as entries in the vectors, and the exponent 2 indicates that each vector contains two entries.

Examples:

Note: When we name a vector with a letter, we can type it as, for example, $\boldsymbol{u}$ or $\boldsymbol{v}$. But it is difficult to write in boldface, so we usually denote a vector with an $\rightharpoonup$ (called a "vector bar" or a "rightwards harpoon") and we would instead write $\vec{u}$ or $\vec{v}$.

Definition: Two vectors are **equal** if and only if their corresponding entries are equal.

Examples:

Note: $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ are *not* equal because vectors in $\mathbb{R}^2$ are *ordered pairs* of real numbers.

Definition: Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^2$, their sum is the vector $\boldsymbol{u} + \boldsymbol{v}$ obtained by adding corresponding entries of $\boldsymbol{u}$ and $\boldsymbol{v}$.

Examples:

Definition: Given a vector $\boldsymbol{u}$ and real number c, the scalar multiple of $\mathbf{u}$ by c is the vector $c\boldsymbol{u}$ obtained by multiply each entry in $\boldsymbol{u}$ by c.

Note: The number c in $c\boldsymbol{u}$ is called a scalar; it is written in lightface type to distinguish it from the boldface vector $\boldsymbol{u}$. Handwritten, this would look like $c\vec{u}$ or $c \cdot \vec{u}$.

Examples:

Geometric Descriptions/Interpretations of $\mathbb{R}^2$

Object	Example(s)	Geometric Interpretation
Vectors as Points		
Vectors with Arrows		
Vector Sum		
Scalar Multiples		

Parallelogram Rule for Addition

Definition: If $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^2$ are represented as points in the plane, then $\boldsymbol{u} + \boldsymbol{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are $\boldsymbol{u}$, $\mathbf{0}$, and $\boldsymbol{v}$.

Note: The zero vector in $\mathbb{R}^2$, denoted $\mathbf{0}$ or $\vec{0}$, is the vector with 2 zeros. In $\mathbb{R}^2$, $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Examples:

Geometric Interpretation of Vectors in $\mathbb{R}^3$

Vectors in $\mathbb{R}^3$ are 3×1 column matrices with three entries. They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity.

Examples:

Vectors in $\mathbb{R}^n$

Definition: If n is a positive integer, $\mathbb{R}^{\boldsymbol{n}}$ (read "**r-n**") denotes the collection of all lists (or ordered n-tuples) of n real numbers, usually written as $n \times 1$ column matrices.

Note: Generally, the vector whose entries are all zero is called the **zero vector**. Usually, the number of entries in $\mathbf{0}$ will be clear from the context.

Examples:

Algebraic Properties of $\mathbb{R}^n$

For all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in $\mathbb{R}^n$ and all scalars c and d:

(i) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

(ii) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

(iii) $\boldsymbol{u} + \mathbf{0} = \mathbf{0} + \boldsymbol{u} = \boldsymbol{u}$

(iv) $\boldsymbol{u} + (-\boldsymbol{u}) = (-\boldsymbol{u}) + \boldsymbol{u} = 0$, where $-\boldsymbol{u}$ denotes $(-1)\boldsymbol{u}$

(v) $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$

(vi) $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$

(vii) $c(d\boldsymbol{u}) = (cd)\boldsymbol{u}$

(viii) $1\boldsymbol{u} = \boldsymbol{u}$

(ix) $0\boldsymbol{u} = \mathbf{0}$

Note: For simplicity of notation, a vector such as $\boldsymbol{u} + (-1)\boldsymbol{v}$ is often written as $\boldsymbol{u} - \boldsymbol{v}$.

Linear Combinations

<u>Definition:</u> Given vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_p}$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \dots, c_p$, the vector $\boldsymbol{y}$ defined by

$$\boldsymbol{y} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_p\boldsymbol{v_p}$$

is called a **linear combination** of $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_p}$ with **weights** $c_1, c_2, \dots, c_p$.

Examples:

Theorem B:

A vector equation

$$x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n} = \boldsymbol{b}$$

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

$$[\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \dots \quad \boldsymbol{a_n} \quad \boldsymbol{b}].$$

In particular, $\boldsymbol{b}$ can be generated by a linear combination of $\boldsymbol{a_1}, \dots, \boldsymbol{a_n}$ if and only if there exists a solution to the linear system corresponding to the augmented matrix above.

Span of a Set of Vectors

Definition: The set of all linear combinations of $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$ is denoted by Span$\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated) by** $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$. That is, Span$\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ is the collection of all vectors that can be written in the form

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_p\boldsymbol{v_p}$$

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

Note: Asking whether a vector $\boldsymbol{b}$ is in Span$\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ amounts to asking whether the vector equation $x_1\boldsymbol{v_1} + x_2\boldsymbol{v_2} + \cdots + x_p\boldsymbol{v_p} = \boldsymbol{b}$ has a solution, or equivalently, asking whether the linear system with augmented matrix $[\boldsymbol{v_1} \quad \cdots \quad \boldsymbol{v_p} \quad \boldsymbol{b}]$ has a solution.

Examples:

1.4 The Matrix Equation $A\boldsymbol{x} = \boldsymbol{b}$

__Definition:__ If A is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \dots, \boldsymbol{a_n}$, and if x is in $\mathbb{R}^n$, then the ***product of A and $\vec{x}$***, denoted by $A\boldsymbol{x}$, is the linear combination of the columns of A using the corresponding entires in $\boldsymbol{x}$ as weights; that is,

$$A\boldsymbol{x} = [\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \dots \quad \boldsymbol{a_4}]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n}$$

Note 1: Ax is defined only if the number of columns of A equals the number of entries in x.

Note 2: If A is an $m \times n$ matrix and $\boldsymbol{x}$ is a $n \times 1$ column vector, then $A\boldsymbol{x} = \boldsymbol{b}$ is an $m \times 1$ column vector.

__Definition:__ An equation of the form $A\boldsymbol{x} = \boldsymbol{b}$ is called a **matrix equation**.

Examples:

Theorem 3: Solution Equivalence

<u>Theorem 3:</u>
If A is an $m \times n$ matrix, with columns $\boldsymbol{a_1}, \dots, \boldsymbol{a_n}$, and if $\boldsymbol{b} \in \mathbb{R}^m$, the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ has the same solution set as the vector equation

$$x_1\boldsymbol{a_1} + x_2\,\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n} = \boldsymbol{b}$$

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

$$[\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \dots \quad \boldsymbol{a_n} \quad \boldsymbol{b}].$$

<u>Note:</u> Theorem 3 provides a powerful tool for gaining insight into problems in linear algebra, because a system of linear equations may be viewed in three different but equivalent ways: (1) as a matrix equation, (2) as a vector equation, or (3) as a system of linear equations.

Corollary 3.1:

<u>Corollary 3.1:</u> The equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if $\boldsymbol{b}$ is a linear combination of the columns of A.

Examples:

Theorem 4: Equivalent Conditions for the Existence of Solutions

<u>**Theorem 4:**</u>
Let A be an $m \times n$ matrix. Then, the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

(a) For each $\boldsymbol{b}$ in $\mathbb{R}^m$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.
(b) Each $\boldsymbol{b}$ 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.

<u>Note 1:</u> Theorem 4 is one of the most useful theorems in this chapter.

<u>Note 2:</u> Theorem 4 is about a coefficient matrix A, not an augmented matrix. If an augmented matrix $[A \quad \boldsymbol{b}]$ has a pivot position in every row, then the equation $A\boldsymbol{x} = \boldsymbol{b}$ may or may not be consistent.

Examples:

Row-Vector Rule for Computing $A\boldsymbol{x}$

If the product $A\boldsymbol{x}$ is defined, then the ith entry in $A\boldsymbol{x}$ is the sum of the products of corresponding entries from row i of A and from the vector $\boldsymbol{x}$.

In other words, if the matrix A and vector x are defined as follows,

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \ddots & a_{i,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

then, the ith entry in $A\boldsymbol{x}$=**b**, i.e., b_i, is given by

$$b_i = a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n$$

or, equivalently,

$$b_i = \prod_{k=1}^{n} a_{i,k}x_k.$$

Theorem 5: Properties of the Matrix-Vector Product $A\boldsymbol{x}$

Theorem 5:

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

a. $A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$

b. $A(c\boldsymbol{u}) = c(A\boldsymbol{u})$

Examples:

1.5 Solution Sets of Linear Systems

Homogeneous Linear Systems

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

<u>Definition:</u> Such a system $A\boldsymbol{x} = \mathbf{0}$ always has at least one solution, namely $\boldsymbol{x} = \mathbf{0}$ (the zero vector in $\mathbb{R}^n$). This zero solution is usually called the **trivial solution.**

<u>Definition:</u> For a given equation $A\boldsymbol{x} = \mathbf{0}$, the important question is whether there exists a **nontrivial solution**, that is, a nonzero vector **x** that satisfies A**x**=**0**.

Examples:

Corollary 2.1: Corollary to Theorem 2

The homogenous equation Ax=0 has a nontrivial solution if and only if the equation has at least one free variable.

Parametric Vector Form

<u>Definition:</u> Whenever a solution set is described explicitly with vectors, we say that a solution is in **parametric vector form**.

Examples:

Nonhomogeneous Linear Systems

A system of linear equations is said to be **nonhomogeneous** if it *<u>cannot</u>* be written in the form $A\boldsymbol{x} = \mathbf{0}$, where A is an $m \times n$ matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$.

When a nonhomogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding *<u>homogenous</u>* system.

Examples:

Theorem 6: Solutions of $A\boldsymbol{x} = \boldsymbol{b}$

Theorem 6:
Suppose the equation $A\boldsymbol{x} = \boldsymbol{b}$ is consistent for some given $\boldsymbol{b}$, and let $\boldsymbol{p}$ be a solution. Then the solution set of $A\boldsymbol{x} = \boldsymbol{b}$ is the set of all vectors of the form $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v_h}$, where $\boldsymbol{v_h}$ is any solution of the homogenous equation $A\boldsymbol{x} = \boldsymbol{0}$.

Note 1: Theorem 6 says that if Ax=b has a solution, then the solution set is obtained by translation the solution set of Ax=0, using any particular solution p of Ax=b for the translation.

Note 2: Theorem 6 applies only to an equation $A\boldsymbol{x} = \boldsymbol{b}$ that has *at least* one nonzero solution $\boldsymbol{p}$. When $A\boldsymbol{x} = \boldsymbol{b}$ has no solution, the solution set is empty.

Writing a Solution Set (of a Consistent System) in Parametric Vector Form

The following algorithm outlines the process of writing a solution set 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 $\boldsymbol{x}$ as a vector whose entries depend on the free variables, if any.
4. Decompose $\boldsymbol{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

Examples:

Examples (cont.):

1.7 Linear Independence

Homogenous equations can be studied from a different perspective by writing them as vector equations. In this way, the focus shifts from the unknown solutions of $A\boldsymbol{x} = \boldsymbol{0}$ to the vectors that appear in the vector equations.

For example, consider the equation

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This equation has a trivial solution, of course, where $x_1 = x_2 = x_3 = 0$. The main issue is whether the trivial solution is the *only one*.

Linear (In)Dependence

Definition: An indexed set of vectors $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\boldsymbol{v_1} + x_2\boldsymbol{v_2} + \cdots + x_p\boldsymbol{v_p} = \boldsymbol{0}$$

has only the trivial solution.

Definition: The set $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zeros, such that

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_p\boldsymbol{v_p} = \boldsymbol{0}.$$

When the weights $c_1, \dots, c_p$ are not all zero, the equation $c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_p\boldsymbol{v_p} = \boldsymbol{0}$ is called a **linear dependence relation** among $\boldsymbol{v_1}, \dots, \boldsymbol{v_p}$.

Examples:

Examples (cont.):

Linear Independence of Matrix Columns

Suppose that we begin with a matrix $A = [\boldsymbol{a_1} \quad \cdots \quad \boldsymbol{a_n}]$ instead of a set of vectors. The matrix equation $A\boldsymbol{x} = \boldsymbol{0}$ can be written as

$$x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n} = \boldsymbol{0}.$$

Each linear dependence relation among the columns of A corresponds to a nontrivial solution of Ax=0.

Theorem C:

The columns of a matrix A are linearly independent if and only if the equation $A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution.

Examples:

Special Case: Set of One Vector

A set containing only one vector—say, $\boldsymbol{v}$—is linearly independent if and only if $\boldsymbol{v}$ is not the zero vector. This is because the vector equation $x_1\boldsymbol{v} = \mathbf{0}$ has only the trivial solution when $\boldsymbol{v} \neq \mathbf{0}$.

The zero vector is linearly dependent because $x_1\mathbf{0} = \mathbf{0}$ has many nontrivial solutions.

Example:

Special case: Set of Two Vectors

A set of two vectors $\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ is linearly dependent if at least one of the vectors is a multiple of the other. The set is linearly independent if and only if neither of the vectors is a multiple of the other.

Theorem 7: Characterization of Linearly Dependent Sets

Theorem 7:
An indexed set $S = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ 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. In fact, if S is linearly dependent and $v_1 \neq 0$, then some $\boldsymbol{v_j}$ (with $j > 1$) is a linear combination of the preceding vectors, $\boldsymbol{v_1}, \dots, \boldsymbol{v_{j-1}}$.

<u>Note:</u> Theorem 7 does not say that every vector in a linearly dependent set is a linear combination of the preceding vectors. Rather, it says that in a linearly dependent set there exists a vector such that it is linear combination of the other vectors.

Example:

Theorem 8: Upper Bound on Linear Independence

Theorem 8:
If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is any set $\{v_1, \dots, v_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

<u>Note:</u> Theorem 8 says nothing about the case in which the number of vectors in the set does *not* exceed the number of entries in each vector.

Example:

Theorem 9: Linear Dependence and the Zero Vector

Theorem 9:
If a set $S = \{v_1, \dots, v_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Example:

1.8 Introduction to Linear Transformations

Transformations

The difference between a matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ and the associated vector equation $x_1\boldsymbol{a_1} + x_2\,\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n} = \boldsymbol{b}$ is a matter of notation. However, a matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ can arise in linear algebra in a way that does not directly connect to a linear combination of vectors. This happens when we think of the matrix A as an object that "acts" on a vector $\boldsymbol{x}$ by multiplication to produce a new vector called $A\boldsymbol{x}$.

Examples:

Matrix Transformations

From the point of view of transformations, solving the equation $A\boldsymbol{x} = \boldsymbol{b}$ amounts to finding all vectors $\boldsymbol{x}$ in $\mathbb{R}^n$ that are transformed into the vector $\boldsymbol{b}$ in $\mathbb{R}^m$ under the "action" of multiplication by A.

The correspondence from $\boldsymbol{x}$ to $A\boldsymbol{x}$ is a **function** from one set of vectors to another. This concept generalizes the common notion of a function as a rule that transforms one real number into another.

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

Note 1: We will be primarily dealing with the matrix transformation $T: \mathbb{R}^n \to \mathbb{R}^m, T(\boldsymbol{x}) = A\boldsymbol{x}$, where A is an $m \times n$ matrix. As a shorthand notation, we will denote this *matrix transformation* by $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

Note 2: If A is an $m \times n$ matrix, then the domain of $T: \boldsymbol{x} \mapsto A\boldsymbol{x}$ is $\mathbb{R}^n$ and the codomain of T is $\mathbb{R}^m$. In other words, the exponent on the domain is the number of columns of A and the exponent on the range is the number of rows of A.

Examples:

Examples (cont.):

Linear Transformation

Recall Theorem 5:
If A is an $m \times n$ matrix, u and $\boldsymbol{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:
a. $A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$
b. $A(c\boldsymbol{u}) = c(A\boldsymbol{u})$

These properties, written in function notation, identify the most important class of transformation in linear algebra.

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

Note 1: "Every matrix transformation is a linear transformation" is a corollary to Theorem 5. Thus, we might also call this definition, instead, *Corollary 5.1*.

Note 2: Linear transformations *preserve the operations of vector addition and scalar multiplication*.

Examples:

Properties of a Linear Transformation:

(1) $T(\boldsymbol{u}+\boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v}$ in the domain of T;
(2) $T(c\boldsymbol{u}) = c(Tu)$ for all scalars c and all $\boldsymbol{u}$ in the domain of T.
(3) $T(\boldsymbol{0^n}) = \boldsymbol{0^m}$
(4) $T(c\boldsymbol{u} + d\boldsymbol{v}) = cT(\boldsymbol{u}) + dT(\boldsymbol{v})$ for all scalars c, d and all $\boldsymbol{u}, \boldsymbol{v}$ in the domain of T.

Examples:

1.9 The Matrix of a Linear Transformation

Motivating Example

The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $\boldsymbol{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose T is a lienar transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ such that

$$T(\boldsymbol{e_1}) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \text{ and } T(\boldsymbol{e_2}) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}.$$

With no additional information, find a formula for the image of an arbitrary $\boldsymbol{x}$ in $\mathbb{R}^2$.

Example:

Theorem 10: Existence of a Matrix Equation for every Linear Transformation

Theorem 10:
Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then, there exists a unique matrix A such that
$$T(\boldsymbol{x}) = A\boldsymbol{x} \text{ for all } x \text{ in } \mathbb{R}^n.$$
In fact, A is the $m \times n$ matrix whose jth column is the vector $T(\boldsymbol{e_j})$, where $\boldsymbol{e_j}$ is the jth column of the identiy matrix in $\mathbb{R}^n$:
$$A = \begin{bmatrix} T(\boldsymbol{e_1}) & \dots & T(\boldsymbol{e_j}) & \dots & T(\boldsymbol{e_n}) \end{bmatrix}.$$

<u>Note 1:</u> The matrix A, when defined as above, is called the **standard matrix for the linear transformation $\boldsymbol{T}$**.

<u>Note 2:</u> Theorem 10 says two things, not only does A *exist*, but A is *unique*.

Proof: (If $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $\exists!\, A | T(\boldsymbol{x}) = A(\boldsymbol{x})\ \forall \boldsymbol{x} \in \mathbb{R}^n$)

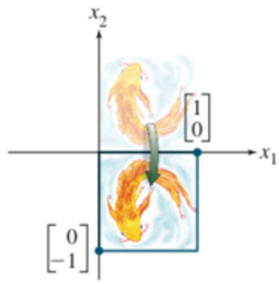
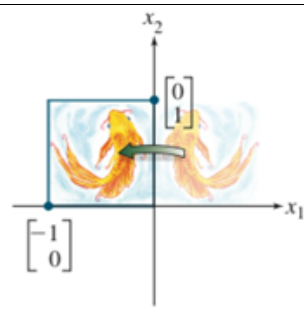
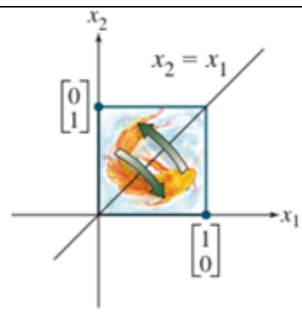
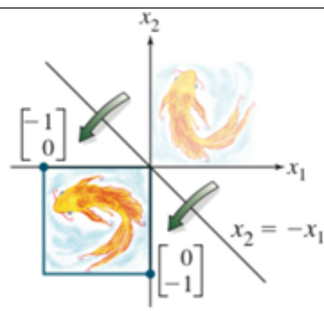
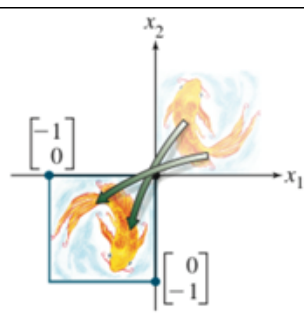
Examples:

Geometric Linear Transformations of $\mathbb{R}^2$

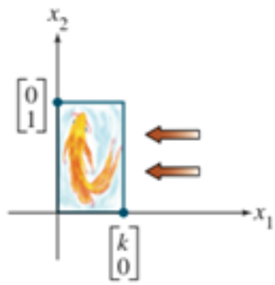
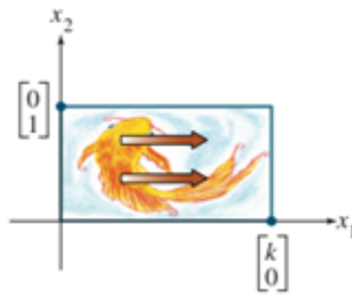
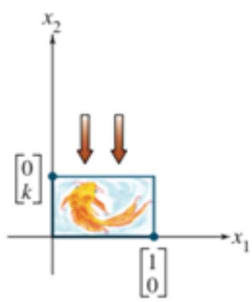
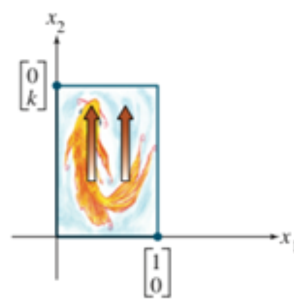
Definition: Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that rotations each point in $\mathbb{R}^2$ about the origin through an angle ϕ (lowercase-phi; handwritten as φ), with counterclockwise rotation for $\phi > 0$. The standard matrix A of this transformation is given by

$$A = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

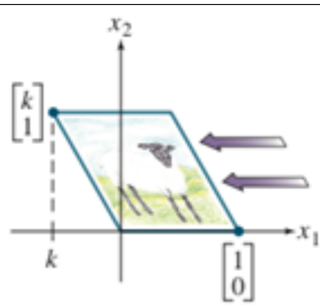
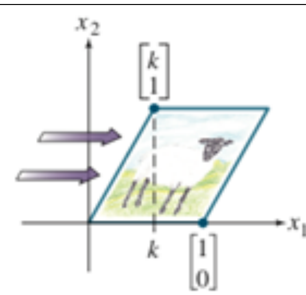
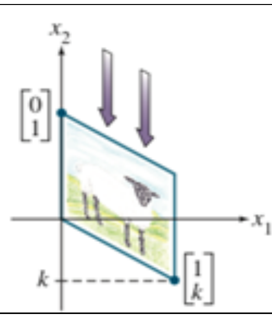
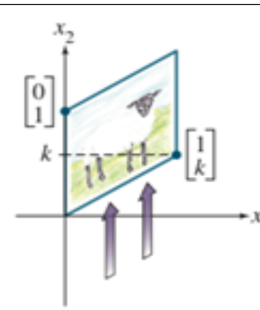
Reflections

Transformation	Image of the Unit Square	Standard Matrix
		
		
		
		
		

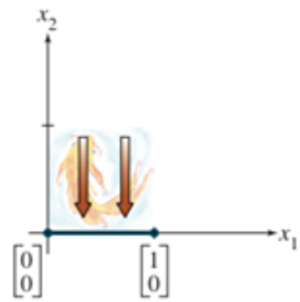
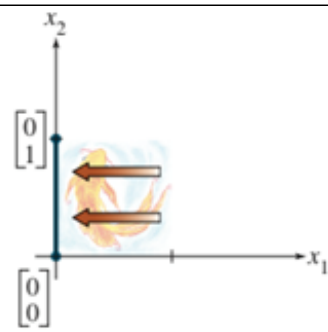
Contractions and Expansions

Transformation	Image of the Unit Square		Standard Matrix
	$0 < k < 1$	$k > 1$	
			
			

Shears

Transformation	Image of the Unit Square		Standard Matrix
	$k < 0$	$k > 0$	
			
			

Projections

Transformation	Image of the Unit Square	Standard Matrix
		
		

Existence and Uniqueness Question

The concept of a linear transformation provides a new way to understand the existence and uniqueness questions we asked earlier. Asking these questions requires some new terminology.

Definition: A mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ is said to be **injective** on $\mathbb{R}^n$ if each $\boldsymbol{x}$ in $\mathbb{R}^n$ is maps to one and only one $\boldsymbol{b}$ in $\mathbb{R}^m$.

Note 1: Some books (including ours) use the language **one-to-one** instead of injective. We will use them interchangeably but prefer injective.

Note 2: Another way to think of injectivity is that every element in the domain maps to only one element in the domain. Or, $\forall b_1, b_2 \in \mathbb{R}^m, T(\boldsymbol{b_1}) = T(\boldsymbol{b_2}) \Rightarrow \exists b_1, b_2 \in \mathbb{R}^n | \boldsymbol{b_1} = \boldsymbol{b_2}$.

Definition: A mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ is said to be **surjective** on $\mathbb{R}^m$ if each $\boldsymbol{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\boldsymbol{x}$ in $\mathbb{R}^n$.

Note 3: Some books (including ours) use the language **onto** instead of surjective. We will use them interchangeably but prefer surjective.

Note 4: Another way to think of surjectivity is that every element in the co-domain is mapped by at least one element in the domain. Or, if $\forall b \in \mathbb{R}^m \Rightarrow \exists x \in \mathbb{R}^n | T(\boldsymbol{x}) = \boldsymbol{b}$.

Examples:

Theorem 11: Condition for Injectivity

Theorem 11:
Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then T is injective (one-to-one) if, and only if, the equation $T(\boldsymbol{x}) = \boldsymbol{0}$ has only the trivial solution.

Note 1: By Theorem 11, it suffices to show $T(\boldsymbol{x}) = \boldsymbol{0}$ *has only the trivial solution* to show that T is injective.

Proof:

Theorem 12: Injectivity & Subjectivity of $T: \boldsymbol{x} \mapsto A\boldsymbol{x}$

Theorem 12:
Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation, and let A be the standard matrix for T. Then:

a. T is surjective (onto) from $\mathbb{R}^n$ to $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.

b. T is injective (one-to-one) if and only if the columns of A are linearly independent.

Proof:

Examples: