Math 33A Notes

Written by Victoria Kala
vtkala@math.ucla.edu
Last updated April 22, 2019

Systems of Linear Equations

A **linear equation** is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

where $x_1, ..., x_n$ are variables, $a_1, ..., a_n$ are constants (these are also called the **coefficients**), and b is also a constant.

Example. Consider the following equations:

(a) $4x + 3z = 1$ (b) $8\sqrt{x_1} - 2x_2 + 4x_3^{-1} = 2$ (c) $x^2 + xy + y^2 = 1$ (d) $x_1 + 2x_2 + 3x_3 - 4x_4 = 0$

The equations in (a) and (d) are linear because they can be written in the form above. The equations (b) and (c) are not linear due to the square root, exponents, and multiplication between variables. □

A **system of linear equations** is a collection of one or more linear equations. A system of linear equations is said to be **consistent** if there is a solution, or is said to be **inconsistent** if there is no solution. A consistent system is said to be **independent** if there is exactly one solution, or is said to be **dependent** if there are infinitely many solutions. To summarize:

- Inconsistent: no solution to the system
- Consistent: there is a solution to the system
  - Independent: exactly one solution to the system
  - Dependent: infinitely many solutions to the system

We can write systems of linear equation as an **augmented matrix**. We then use elementary row operations to solve the matrix/system:

- Interchange any two rows
- Multiply by a row by a nonzero constant (scaling)
- Add multiples of rows to each other and replace one of these rows (replacement)

Row Reduction and Echelon Forms

A 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 (called the **pivot**) 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 (pivot) are zeros.

I personally like to have my pivots (leading entries) be 1, but your textbook does not indicate this preference. The process of getting to echelon form is called Gaussian elimination.

If a matrix is an echelon form, it is said to be in **reduced echelon form** (or **reduced row echelon form**):

1. The leading entry (pivot) in each nonzero row is 1.

2. Each leading 1 is the only nonzero entry in its column.

The process of getting to reduced row echelon form is called Gauss-Jordan elimination.

Example. Consider the following matrices

$$\text{(a)} \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(d)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \text{(e)} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The matrices (a) and (d) are in row echelon form. The matrices (c) and (e) are in reduced row echelon form. The matrix (b) is neither in echelon nor reduced row echelon form due to the second row. □

If we are solving an augmented matrix, say for example a 3×3 system, then for row echelon our goal will be

$$\left(\begin{array}{ccc|c} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{array}\right) \quad \text{or} \quad \left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \end{array}\right)$$

and our goal for reduced row echelon will be

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right).$$

The variables that correspond with pivot columns are said to be **basic variables**. The variables that correspond with columns missing pivots are said to be **free variables**. Whenever there is a presence of a free variable the corresponding system will have infinitely many solutions. In some cases, the presence of a free variable will correspond with a row of zeros (this is especially true for square systems, like a 3×3 system). A system will be inconsistent if we get a false equation such as $0 = 2$.

Vector Equations

A matrix with one column is called a **vector**. A vector with m rows is a vector in the space $\mathbb{R}^m$. Our current operations for vectors right now only consist of addition (including subtraction) and scalar multiplication. We will learn more operations later on.

Let $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$ be vectors. We say that

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n}$$

is a **linear combination** of the vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$, where $c_1, c_2, ..., c_n$ are constants.

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

If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_n}$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the product of A and $\mathbf{x}$ is given by

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n}.$$

For $\mathbf{b}$ in $\mathbb{R}^m$, the matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution as the equation

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n} = \mathbf{b},$$

which also has the same solution as the augmented matrix

$$\begin{pmatrix} \mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} \mid \mathbf{b} \end{pmatrix}.$$

The following is a very useful theorem:

Theorem 1. *Let A be an $m \times n$ matrix. Then the following statements are equivalent (if one is true, all are true; if one is false, 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 $\mathbf{R}^m$ 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.

A is a matrix of vectors (the vectors are the columns). If we can show that A has a pivot position in every row using row operations, then the vectors that make up A span the space. We can therefore write any matrix in that space as a linear combination of these vectors.

Solution Sets of Linear Systems

The matrix equation $A\mathbf{x} = \mathbf{0}$ is said to be a **homogeneous** equation. Homogeneous is just a fancy way of saying our equation is equal to "0" (this term will show up in later math courses). $\mathbf{x} = \mathbf{0}$ is certainly a solution to this equation; this is called the **trivial solution**. Some systems have more solutions to this equation, and these are called **nontrivial solutions**. These occur if and only if the equation has at least one free variable.

If a linear system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions, the general solution can be written in the **parametric vector form** (a linear combination of vectors that satisfy the equation).

Example. Suppose a system of equations has the solution $x_1 = 4x_3 - 1, x_2 = -x_3 + 3, x_3 = x_3$ (notice here that x_3 is a free variable). We can write our solution as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4x_3 - 1 \\ -x_3 + 3 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}.$$

This right hand side is in parametric form as it is a linear combination of vectors. □

Dot Product, Length, and Orthogonality

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. The **dot product**, or **inner product**, $\mathbf{u} \cdot \mathbf{v}$ is defined to be

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

The **norm**, or **length**, of a vector $\mathbf{u} \in \mathbb{R}^n$ is defined by

$$||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}.$$

Notice that from this definition we have $||\mathbf{u}||^2 = \mathbf{u} \cdot \mathbf{u}$.

The **distance** between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is given by $||\mathbf{u} - \mathbf{v}||$. The vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthogonal** if and only if $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

The dot product of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is also given by the formula

$$\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||||\mathbf{v}|| \cos \theta$$

where θ is the angle between $\mathbf{u}, \mathbf{v}$. We can rearrange this formula to find the angle between to vectors:

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||||\mathbf{v}||}\right).$$

Properties:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (this says order doesn't matter)
- $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$
- $||c\mathbf{v}|| = |c|||\mathbf{v}||$ for any $c \in \mathbb{R}$

Linear Transformations

A **transformation** (also called a function or mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns a vector $\mathbf{x} \in \mathbb{R}^n$ to a vector $T(\mathbf{x}) \in \mathbb{R}^m$. We say a transformation T is **linear** if:

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $c \in \mathbb{R}, \mathbf{u}$ in the domain of T.

These properties imply that $T(\mathbf{0}) = \mathbf{0}$ and $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$. Linear combinations are also preserved; that is:

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n).$$

Some geometric transformation formulas:

- Projection onto a line containing the vector $\mathbf{u}$ is given by $T(\mathbf{x}) = \text{proj}_{\mathbf{u}}(\mathbf{x})$ where

$$\text{proj}_{\mathbf{u}}(\mathbf{x}) = \frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}.$$

- Reflection about a line containing the vector $\mathbf{u}$ is given by $T(\mathbf{x}) = 2\text{proj}_{\mathbf{u}}(\mathbf{x}) - \mathbf{x}$.
- Projection onto a plane $ax + by + cd = 0$ is given by $T(\mathbf{x}) = \mathbf{x} - \text{proj}_{\mathbf{n}}(\mathbf{x})$ where $\mathbf{n} = (a, b, c)$.
- Reflection about a plane $ax+by+cd = 0$ is given by $T(\mathbf{x}) = \mathbf{x} - 2\text{proj}_{\mathbf{n}}(\mathbf{x})$ where $\mathbf{n} = (a, b, c)$.

A matrix transformation from $\mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\mathbf{x}) = A\mathbf{x}$ where A is an $m \times n$ matrix. Linear transformations can be written as matrix transformations! The standard matrix representation of A is given by

$$A = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & ... & T(\mathbf{e}_n) \end{pmatrix}$$

where $\mathbf{e}_1 = (1, 0, ..., 0), \mathbf{e}_2 = (0, 1, ..., 0), ..., \mathbf{e}_n = (0, 0, ..., 1)$.

Matrix Operations

Recall that the size of a matrix is the number of rows by the number of columns. For example, a 3×100 matrix has 3 rows, 100 columns.

We can perform the following operations with matrices:

- Addition: The sum $A + B$ is defined as long as A and B are the same size, and each entry of $A + B$ is the sum of the corresponding entries in A and B.
- Scalar multiplication: If $c \in \mathbb{R}$ and A is a matrix, then cA is the matrix whose entries are all multiplied by c.

- Matrix Multiplication: The product AB is defined as long the number of columns of A is the same as the number or rows of B. It is helpful to remember the following:

$$\underbrace{A}_{m\times n} \times \underbrace{B}_{n\times p} = \underbrace{C}_{m\times p}$$

  The ijth entry (that is, the ith row, jth column) of AB is found by multiplying the ith row of A and the jth column of B.

- Transpose: The transpose of a matrix A (denoted by A^T) is found by switching its rows and columns.

Example. Let A be the 3×4 matrix is given by

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 2 \\ 5 & -7 & 0 & 0 \end{pmatrix}.$$

The transpose of A is given by

$$A^T = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 2 \\ 5 & -7 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 1 & -7 \\ 3 & -1 & 0 \\ 4 & 2 & 0 \end{pmatrix}.$$

Notice that A^T is a 4×3 matrix. (The transpose "swaps" the size of the matrix.)

Some items to remember:

- In general, $AB \neq BA$.
- The cancellation law doesn't always hold; that is, $AB = AC$ does NOT always imply $B = C$.
- $(AB)^T = B^T A^T$

The Inverse of a Matrix

An $n \times n$ matrix A is **invertible** (sometimes called **nonsingular**) if there exists an A^{-1} such that

$$AA^{-1} = A^{-1}A = I.$$

Not all matrices have inverses, however. One way we can check to see if an inverse is invertible is to calculate the determinant (see next section). If $\det A \neq 0$ then A^{-1} exists.

We can use the inverse to solve systems of equations. If A is invertible then the equation $A\mathbf{x} = \mathbf{b}$ has the solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Some nice properties of inverses:

- $(A^{-1})^{-1} = A$

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$

We can easily find the inverse of a 2×2 matrix. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

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

provided that $ad - bc \neq 0$.

For larger matrices, we use row operations to find inverses. We augment our matrix with the identity matrix:

$$\begin{pmatrix} A & I \end{pmatrix}$$

and then use row operations to reduce to the matrix to

$$\begin{pmatrix} I & A^{-1} \end{pmatrix}.$$