FORK1003
Preparatory Course in Linear Algebra 2016/17
Lecture 2: Matrices

August 2, 2016

1 Matrices and Matrix Operations

1.1 Matrix Defined

Definition 1.1 (Matrix)**.** An $m \times n$ matrix is an array of $m \cdot n$ numbers arranged in m rows and n columns.

$m \times n$ matrix:

$$\begin{array}{cc} & (n \text{ columns}) \\ (m \text{ rows}) & \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{array}$$

The number a_{14} for example refers to the entry in 1st row, 4th column.

Notation. For brevity, when we have a $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

we abbreviate it by writing $A = (a_{ij})$. Similarly, if

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

we write $B = (b_{ij})$.

1.2 Addition and Scalar Multiplication

We can define many operations on matrices. The simplest ones are *addition* and *scalar multiplication*.

Definition 1.2 (Matrix Addition & Subtraction)**.** Let $A = (a_{ij})$ and $B = (b_{ij})$ be two matrices of the same dimension $m \times n$. Then $A + B$ is the $m \times n$ matrix

$$A + B = (a_{ij} + b_{ij}) = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Similarly, we can define matrix subtraction:

$$A - B = (a_{ij} - b_{ij}) = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2n} - b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \dots & a_{mn} - b_{mn} \end{bmatrix}$$

Note that the dimensions need to match. If A and B had a different number of columns or rows, adding them together would not make sense.

Example 1.3. Let

$$A = \begin{bmatrix} 3 & -2 & 9 \\ 1 & 2 & -4 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} -2 & 7 & -4 \\ 3 & 1 & 2 \end{bmatrix}.$$

Then,

$$A + B = \begin{bmatrix} 3-2 & -2+7 & 9-4 \\ 1+3 & 2+1 & -4+2 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 5 \\ 4 & 3 & -2 \end{bmatrix}.$$

However,

$$\begin{bmatrix} 3 & -2 & 9 \\ 1 & 2 & -4 \end{bmatrix} + \begin{bmatrix} 1 & 5 \\ 4 & 3 \end{bmatrix}$$

is not defined.

Definition 1.4 (Scalar Multiplication)**.** Let $A = (a_{ij})$ be a $m \times n$ matrix, and c a constant. Then cA is the $m \times n$ matrix

$$cA = (c \cdot a_{ij}) = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}$$

Example 1.5. Let $c = -2$ and

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

Then

$$cA = \begin{bmatrix} -6 & 4 & -2 \\ -14 & 0 & 2 \end{bmatrix}.$$

1.3 Matrix Multiplication

Matrix multiplication is slightly more complicated, but it is the most important and "natural" operation on matrices. The idea of multiplication $A \cdot B$ is that **we multiply the *rows* of $\boldsymbol{A}$ with the *columns* of $\boldsymbol{B}$.** Let us first define what we mean by multiplying a row with a column.

Definition 1.6 (Vector dot product)**.** Suppose you have a row $\begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}$ and a column

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}.$$

Then, the *vector dot product* or 'multiplication of the row and the column' is the sum

$$\begin{aligned} \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} &= a_1 b_1 + a_2 b_2 + \dots + a_p b_p \\ &= \sum_{k=1}^{p} a_k b_k. \end{aligned}$$

So when we multiply a row and a column, we multiply each of the entries in pairs, and add it all together.

Example 1.7. If we have the row vector $\begin{pmatrix} 3 & -2 & 4 \end{pmatrix}$ and the column vector

$$\begin{pmatrix} 6 \\ -1 \\ -3 \end{pmatrix},$$

we have the dot product

$$\begin{aligned} \begin{pmatrix} 3 & -2 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ -1 \\ -3 \end{pmatrix} &= 3 \cdot 6 + (-2) \cdot (-1) + 4 \cdot (-3) \\ &= 8. \end{aligned}$$

Definition 1.8 (Matrix multiplication)**.** Let $A = (a_{ij})$ be a $m \times p$ matrix and $B = (b_{ij})$ a $p \times n$ matrix. Then AB is the $m \times n$ matrix whose (i,j)-th entry[1] is the dot product of the ith row of A and the jth column of B.

$$AB = (c_{ij}) = \left(\sum_{k=1}^{p} a_{ik} b_{kj} \right).$$

Example 1.9. For example, if

$$A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & -1 \\ 4 & 7 \end{bmatrix},$$

then (denoting 1st row of A as $R1_A$ and 1st column of B as $C1_B$)

$$\begin{aligned} AB &= \begin{bmatrix} R1_A \cdot C1_B & R1_A \cdot C2_B \\ \\ R2_A \cdot C1_B & R2_A \cdot C2_B \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} & \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \end{bmatrix} \\ \\ \begin{bmatrix} -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} & \begin{bmatrix} -1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 2+4 & -2+7 \\ -1+12 & 1+21 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 5 \\ 11 & 22 \end{bmatrix} \end{aligned}$$

Remark 1.10. I find that the easiest way to carry out matrix multiplications is to use your index fingers (pekefingere): When you calculate the (i,j)th entry of AB, you place your left index finger to the left on the ith row of A, and your right index finger at the top of the jth column of B. Then you move your left finger to the right and your right finger downwards, multiplying and adding each pair of entries along the row and column.

The multiplication *AB* is defined if and only if the number of *columns* in *A* matches the number of *rows* in *B*.

[1] (i,j)th entry means entry in ith row and jth column

Example 1.11. If

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 4 & 0 & 2 \\ -3 & 2 & -2 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 3 & -1 \\ 5 & 0 & 7 \end{bmatrix},$$

then

$$\begin{aligned} AB &= \begin{bmatrix} 3\cdot 0 + 2\cdot 1 - 1\cdot 5 & 3\cdot 1 + 2\cdot 3 - 1\cdot 0 & 3\cdot(-2) + 2\cdot(-1) - 1\cdot 7 \\ 4\cdot 0 + 0\cdot 1 + 2\cdot 5 & 4\cdot 1 + 0\cdot 3 + 2\cdot 0 & 4\cdot(-2) + 0\cdot(-1) + 2\cdot 7 \\ -3\cdot 0 + 2\cdot 1 - 2\cdot 5 & -3\cdot 1 + 2\cdot 3 - 2\cdot 0 & -3\cdot(-2) + 2\cdot(-1) - 2\cdot 7 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 9 & -15 \\ 10 & 4 & 6 \\ -8 & 3 & -10 \end{bmatrix}. \end{aligned}$$

Example 1.12. If

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix},$$

then AB is not defined. BA is defined however, and we have

$$BA = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3\cdot 1 - 2\cdot 2 \\ 1\cdot 1 + 0\cdot 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

1.4 Identity and Zero Matrix

Definition 1.13 (Identity Matrix)**.** The $n \times n$ identity matrix I_n is defined as

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

I_n is called the identity matrix, because it is the **multiplicative identity**:

Proposition 1.14. *Let A be an $m \times n$ matrix. Then*

$$AI_n = I_m A = A.$$

Definition 1.15 (Zero matrix)**.** The $m \times n$ zero matrix $\underline{\mathbf{0}}_{mn}$ is defined as

$$\underline{\mathbf{0}}_{mn} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

The following result is obvious:

Proposition 1.16. *For all $m \times n$ matrixes A,*

$$A + \underline{\mathbf{0}}_{mn} = \underline{\mathbf{0}}_{mn} + A = A,$$

and

$$A\underline{\mathbf{0}}_{np} = \underline{\mathbf{0}}_{pm}A = 0.$$

1.5 Transpose

A matrix operation that is important to mention is the transpose. Given a $m \times n$ matrix A, its transpose A^T is the $n \times m$ matrix you get by reflecting the matrix across its diagonal:

Definition 1.17 (Transpose)**.** Given a matrix $A = (a_{ij})$, its transpose $A^T = (a^t_{ij}) = (a_{ji})$ is the matrix you get when flipping the rows of A into columns, and columns of A into rows. If we have

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix},$$

then

$$A^T = \begin{bmatrix} a_{1} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ a_{13} & a_{23} & \dots & a_{m3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Example 1.18.

$$A = \begin{bmatrix} 2 & 3 & -1 & 6 \\ 4 & -1 & 2 & 10 \end{bmatrix} \qquad A^T = \begin{bmatrix} 2 & 4 \\ 3 & -1 \\ -1 & 2 \\ 6 & 10 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 \\ -1 \\ 3 \\ 4 \end{bmatrix} \qquad B^T = \begin{bmatrix} 2 & -1 & 3 & 4 \end{bmatrix}$$

Properties of the Transpose

Double transpose	$(A^T)^T = A$
Additive distributivity	$(A+B)^T = A^T + B^T$
Scalar multiplication	$(cA)^T = cA^T$
Multiplication	$(AB)^T = B^T A^T$

1.6 Properties of Matrix Operations

Multiplication is not commutative	$AB \neq BA$
Multiplication is associative	$(AB)C = A(BC)$
Left-distributive law	$A(B+C) = AB + AC$
Right-distributive law	$(A+B)C = AC + BC$
Scalar	$c(AB) = (cA)B = A(cB)$

1.7 Square Matrices

Definition 1.19 (Square matrix)**.** A square matrix is any matrix with the same number of rows as columns, so a $n \times n$ matrix.

When a matrix is square, we can multiply it by itself. Thus, we can define powers of square matrices:

Definition 1.20 (Power)**.** Let A be a $n \times n$ matrix. We define A^n to be the product

$$A^n := \underbrace{A \cdot A \cdot \ldots \cdot A}_{n \text{ times}}.$$

Example 1.21. For all $k = 1, 2, 3, \ldots$,

$$(I_n)^k = I_n.$$

Example 1.22.

$$A^2 = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}^2 = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 2\cdot 2 - 1\cdot 3 & 2\cdot(-1) - 1\cdot(-2) \\ 3\cdot 2 - 2\cdot 3 & 3\cdot(-1) - 2\cdot(-2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since A^2 is the identity matrix, we know that

$$\begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}^4 = \left(\begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}^2 \right)^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

By extension, $A^{2k} = (A^2)^k = (I_2)^k = I_2$ for all $k = 1, 2, 3, \ldots$.

Diagonal matrices are also worth knowing:

Definition 1.23 (Diagonal matrix)**.** A diagonal matrix is a square matrix where the only non-zero entries are along the diagonal:

$$\begin{bmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ 0 & 0 & d_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & d_n \end{bmatrix}.$$

Example 1.24. All identity matrices are diagonal matrices. Any square zero matrix is also a diagonal matrix. The following is a 3×3 diagonal matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

2 Inverse Matrices

Inverse matrices is one of the most important concepts of Linear Algebra.

Definition 2.1 (Inverse matrix)**.** A square matrix A is invertible if there exists a matrix A^{-1} such that

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

We then call A^{-1} the inverse of A.

Proposition 2.2. *A non-square matrix is never invertible.*

Example 2.3. We saw earlier that

$$A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$$

satisfies

$$A^2 = I_2.$$

Therefore, A is invertible, and it is its own inverse:

$$A^{-1} := A.$$

Example 2.4. Not all square matrices are invertible. The matrix

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix}$$

is not invertible for example. This is because the 1st and 2nd row are the same, so when you multiply A with any matrix B, the product AB will also have equal rows, and so cannot be of the form $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

2.1 Briefly on Determinants

How do you determine if a matrix is invertible or not?

Theorem 2.5. *A square matrix is invertible, if and only if its* determinant *is nonzero.*

Determinants is the topic of the next lecture. For now we'll only give the determinant of a 2×2 matrix:

Definition 2.6 (Determinant)**.** The determinant of a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is the value

$$\det(A) := a_{11}a_{22} - a_{12}a_{21}.$$

So a quick way of checking if a 2×2 matrix is invertible is to check whether that expression is zero or nonzero.

2.2 Finding the Inverse

A reliable way of calculating the inverse of a matrix is to use *row reduction* and *elementary row operations* (see Lecture 1). As you might recall, we have three elementary row operations:

1. Scalar multiplication: Multiplying a row by a nonzero constant: $(R1 \to cR1)$.
2. Row addition: Adding a multiple of a row to another row: $(R2 \to R2 + 5R1)$.
3. Interchanging: Swapping two rows: $(R2 \leftrightarrow R3)$.

By the following process, we can find the inverse of matrix A through row reduction:

1. Write A on the left side, and the identity matrix I_n on the right side.
2. For each elementary row operations you apply on A, you apply the same operation on the right matrix.
3. Once you have row reduced A to the identity matrix I_n, the right matrix is its inverse A^{-1}.
4. If A cannot be row reduced to I_n, it means that it's not invertible.

Proposition 2.7. *A square matrix is invertible, if and only if it is row equivalent to the identity matrix.*

Example 2.8. Find the inverse to the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}.$$

We write

$$\begin{array}{ccl}
\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \\
\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 8 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} & \begin{array}{l} R2 \to R2 - R1 \\ R3 \to R3 - R1 \end{array} \\
\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix} & \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} & \begin{array}{l} R1 \to R1 - R2 \\ R3 \to R3 - 2R2 \end{array} \\
\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} & R3 \to \frac{1}{2}R3 \\
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 3 & -3 & 1 \\ -\frac{5}{2} & 4 & -\frac{3}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} & \begin{array}{l} R1 \to R1 + 2R3 \\ R2 \to R2 - 3R3 \end{array}
\end{array}$$

Now A has been row reduced to I_3, which means that

$$A^{-1} = \begin{bmatrix} 3 & -3 & 1 \\ -\frac{5}{2} & 4 & -\frac{3}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix}$$

is the inverse of A.

Fact:

1. **If the reduced echelon form of A is I_n, A is invertible.**
2. **If the reduced echelon form of A is not I_n, A is not invertible.**

3 Linear Systems as Matrix Equations

This section is dedicated to the fact that a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned}$$

can be written as a matrix equation

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}.$$

3.1 Vector

Definition 3.1 (Row vector)**.** A row vector is a matrix of dimension $1 \times m$:

$$\underline{\mathbf{v}} = \begin{bmatrix} v_1 & v_2 & \ldots & v_m \end{bmatrix}.$$

Definition 3.2 (Column vector)**.** A column vector is a matrix of dimension $n \times 1$:

$$\underline{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Matrix-vector multiplication
You can multiply matrices and vectors the same way you do matrix multiplication: Let $A = (a_{ij})$ be a $m \times n$ matrix and let

$$\underline{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

be an n-dimensional column vector. Then

$$\begin{aligned} A\underline{\mathbf{x}} &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix}. \end{aligned}$$

So $A\underline{\mathbf{x}}$ is a m-dimensional column vector, where each entry represents a linear expression

$$a_{i1}x_1 + a_{i2}x_2 + \dots a_{in}x_n.$$

Furthermore, we can write matrix equations: Let

$$\underline{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

be a m-dimensional column vector. Then

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

represents the matrix equation

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

This equation corresponds exactly to the linear system:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

Theorem 3.3. *The linear system*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

is equivalent to the matrix equation

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

where

- *A is the **coefficient matrix** of the linear system.*
- *$\underline{\mathbf{x}}$ is the column vector of the n variables $x_1, x_2, \dots, x_n$.*
- *$\underline{\mathbf{b}}$ is the column vector of the m constant terms $b_1, b_2, \dots, b_m$.*

Example 3.4. The linear system

$$\begin{cases} x_1 - 3x_2 + 2x_3 - x_4 = 4 \\ -x_1 + 3x_3 + 2x_4 = -1 \\ 4x_1 + 2x_2 - 5x_4 = 2 \end{cases}$$

is the same as the matrix equation

$$\begin{bmatrix} 1 & -3 & 2 & -1 \\ -1 & 0 & 3 & 2 \\ 4 & 2 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}.$$

3.2 Solving Linear Systems Through Matrix Equations

Consider a $n \times n$ linear system written as a matrix equation:

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}.$$

Suppose A is invertible with inverse matrix A^{-1}. Then properties of matrix multiplication tells us that

$$\begin{aligned} A^{-1}(A\underline{\mathbf{x}}) &= A^{-1}\underline{\mathbf{b}} \\ (A^{-1}A)\underline{\mathbf{x}} &= A^{-1}\underline{\mathbf{b}} \\ I_n\underline{\mathbf{x}} &= A^{-1}\underline{\mathbf{b}} \\ \underline{\mathbf{x}} &= A^{-1}\underline{\mathbf{b}}. \end{aligned}$$

So we have a solution for the above matrix equation: Setting $\underline{\mathbf{x}}$ equal to $A^{-1}\underline{\mathbf{b}}$ solves the equation

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}.$$

But this also means that

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}\underline{\mathbf{b}}$$

is a solution for $x_1, x_2, \ldots, x_n$ for the corresponding linear system. We summarize this in the following result:

Theorem 3.5. *A $(n \times n)$-linear system has a unique solution if and only if its coefficient matrix A is invertible. If so, the unique solution is given by*

$$\underline{\mathbf{x}} = A^{-1}\underline{\mathbf{b}}.$$

Example 3.6. Solve the following linear system by inverting its coefficient matrix:

$$\begin{cases} 6x_1 + 2x_2 + 6x_3 = 20 \\ 2x_1 + \ \ x_2 \qquad\ \ = 4 \\ -4x_1 - 3x_2 + 9x_3 = 3. \end{cases}$$

We can rewrite this as the matrix equation

$$\begin{bmatrix} 6 & 2 & 6 \\ 2 & 1 & 0 \\ -4 & -3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 4 \\ 3 \end{bmatrix}$$

where

$$A = \begin{bmatrix} 6 & 2 & 6 \\ 2 & 1 & 0 \\ -4 & -3 & 9 \end{bmatrix}$$

is the coefficient matrix. We find its inverse by row reducing A:

$$\begin{bmatrix} 6 & 2 & 6 \\ 2 & 1 & 0 \\ -4 & -3 & 9 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1/3 & 1 \\ 2 & 1 & 0 \\ -4 & -3 & 9 \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{6} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R1 \to \frac{1}{6}R1$$

$$\begin{bmatrix} 1 & 1/3 & 1 \\ 0 & 1/3 & -2 \\ 0 & -5/3 & 13 \end{bmatrix} \qquad \begin{bmatrix} 1/6 & 0 & 0 \\ -1/3 & 1 & 0 \\ 2/3 & 0 & 1 \end{bmatrix} \qquad \begin{matrix} R2 \to R2 - 2R1 \\ R3 \to R3 + 4R1 \end{matrix}$$

$$\begin{bmatrix} 1 & 1/3 & 1 \\ 0 & 1 & -6 \\ 0 & -5/3 & 13 \end{bmatrix} \qquad \begin{bmatrix} 1/6 & 0 & 0 \\ -1 & 3 & 0 \\ 2/3 & 0 & 1 \end{bmatrix} \qquad R2 \to 3R2$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -6 \\ 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1/2 & -1 & 0 \\ -1 & 3 & 0 \\ -1 & 5 & 1 \end{bmatrix} \qquad \begin{matrix} R1 \to R1 - \frac{1}{3}R2 \\ R3 \to R3 + \frac{5}{3}R2 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1/2 & -1 & 0 \\ -1 & 3 & 0 \\ -1/3 & 5/3 & 1/3 \end{bmatrix} R3 \to \tfrac{1}{3}R3$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3/2 & -6 & -1 \\ -3 & 13 & 2 \\ -1/3 & 5/3 & 1/3 \end{bmatrix} \begin{matrix} R1 \to R1 - 3R3 \\ R2 \to R2 + 6R3 \end{matrix}$$

So A is invertible and

$$A^{-1} = \begin{bmatrix} 3/2 & -6 & -1 \\ -3 & 13 & 2 \\ -1/3 & 5/3 & 1/3 \end{bmatrix}.$$

Therefore

$$\begin{aligned} \underline{\mathbf{x}} &= A^{-1}\underline{\mathbf{b}} \\ &= \begin{bmatrix} 3/2 & -6 & -1 \\ -3 & 13 & 2 \\ -1/3 & 5/3 & 1/3 \end{bmatrix} \begin{bmatrix} 20 \\ 4 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}. \end{aligned}$$

So the solution to the linear system

$$\begin{cases} 6x_1 + 2x_2 + 6x_3 = 20 \\ 2x_1 + x_2 \phantom{{}+6x_3} = 4 \\ -4x_1 - 3x_2 + 9x_3 = 3. \end{cases}$$

is

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \end{bmatrix}.$$

Furthermore, for any other vector of constants, $\underline{\mathbf{b}}$, the linear system $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ has a unique solution $\underline{\mathbf{x}} = A^{-1}\underline{\mathbf{b}}$.

4 Linear Systems as Linear Combinations of Columns

In the last section of this lecture, we take a look at how we can express a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n} &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n} &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn} &= b_m, \end{aligned}$$

as a linear combination of columns

$$x_1\underline{\mathbf{a}}_1 + x_2\underline{\mathbf{a}}_2 + \ldots + x_n\underline{\mathbf{a}}_n = \underline{\mathbf{b}}.$$

4.1 Matrix Columns

For the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

we denote its columns by

$$\underline{\mathbf{a}}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \qquad \underline{\mathbf{a}}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots \qquad \underline{\mathbf{a}}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix},$$

and we can write the matrix as a *row of columns*:

$$A = \begin{bmatrix} \underline{\mathbf{a}}_1 & \underline{\mathbf{a}}_2 & \dots & \underline{\mathbf{a}}_n \end{bmatrix}$$

Similarly, the linear system/matrix equation

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

can equivalently be expressed as a *linear combination* of the columns of A:

$$\begin{bmatrix} \underline{\mathbf{a}}_1 & \underline{\mathbf{a}}_2 & \dots & \underline{\mathbf{a}}_n \end{bmatrix} \underline{\mathbf{x}} = \underline{\mathbf{b}}$$
$$x_1\underline{\mathbf{a}}_1 + x_2\underline{\mathbf{a}}_2 + \dots + x_n\underline{\mathbf{a}}_n = \underline{\mathbf{b}}.$$

Example 4.1. The linear system

$$\begin{cases} 7x_1 + 2x_2 + x_3 = 1 \\ \phantom{-3x_1 + {}} 3x_2 - x_3 = -2 \\ -3x_1 + 4x_2 - 2x_3 = -1 \end{cases}$$

can be written as the matrix equation

$$\begin{bmatrix} 7 & 2 & 1 \\ 0 & 3 & -1 \\ -3 & 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$$

or the linear combination of columns

$$x_1 \begin{bmatrix} 7 \\ 0 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

4.2 Linear Combinations

Definition 4.2 (Linear Combination)**.** Let $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ be a collection of vectors. We say that a vector $\underline{\mathbf{v}}$ is a *linear combination* of $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ if you can write

$$\underline{\mathbf{v}} = c_1\underline{\mathbf{a}}_1 + c_2\underline{\mathbf{a}}_2 + \ldots + c_n\underline{\mathbf{a}}_n$$

for some constants $c_1, c_2, \ldots, c_n$.

Example 4.3. $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ since

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

However, $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is not a combination of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ since there are no numbers c_1 and c_2 such that

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 + 3c_2 \\ 0 \end{bmatrix}$$

4.3 Spanning

Definition 4.4 (Spanning Set)**.** For a list of vectors $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$, the *spanning set*

$$\mathbf{Span}\left\{\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n\right\}$$

is the set of *all possible linear combinations of* $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$.

So the spanning set is the set of all vectors that can be written as a linear combination of $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$.

Example 4.5.

$$\mathbf{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}\right\} = \left\{\begin{bmatrix} c \\ 0 \end{bmatrix} \mid \text{where } c \text{ is any number}\right\}$$

Definition 4.6 (Spanning $\mathbb{R}^n$)**.** We say that a list of vectors $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_m$ *spans* $\mathbb{R}^n$ if *every n-dimensional vector* $\underline{\mathbf{v}}$ belongs to the spanning set

$$\mathbf{Span}\left\{\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_m\right\}.$$

Remark 4.7. $\mathbb{R}^n$ is called the *n-dimensional Euclidean space* and is the set of all real-valued n-dimensional vectors.

Example 4.8. The vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ span $\mathbb{R}^2$, since for any 2-dimensional vector $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, we can write

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

4.4 Solving Linear Systems

The following result is the reason we are interested in matrix columns, linear combinations and spanning sets:

Theorem 4.9. *Consider a linear system/matrix equation* $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$*, and let* $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ *be the columns of the coefficient matrix* A*. Then, if* $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ *span* $\mathbb{R}^m$*, the linear system has a unique solution for every* m*-dimensional vector* $\underline{\mathbf{b}}$*.*

Why? Solving the linear system $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ is equivalent to solving the vector equation

$$x_1\underline{\mathbf{a}}_1 + x_2\underline{\mathbf{a}}_2 + \ldots + x_n\underline{\mathbf{a}}_n = \underline{\mathbf{b}}.$$

But if $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ span $\mathbb{R}^m$, then by definition there exists $c_1, c_2, \ldots, c_n$ so that

$$c_1\underline{\mathbf{a}}_1 + c_2\underline{\mathbf{a}}_2 + \ldots + c_n\underline{\mathbf{a}}_n = \underline{\mathbf{b}}.$$

So we have a solution to the linear system, $(x_1, x_2, \ldots, x_n) = (c_1, c_2, \ldots, c_n)$.

Definition 4.10 (Standard unit vector)**.** An *m-dimensional standard unit vector* is a vector with 1 in one position and 0 in all the other positions:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

There are m different m-dimensional standard unit vectors. The following result is also helpful:

Proposition 4.11. $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ *span* $\mathbb{R}^m$ *if and only if every* m*-dimensional standard unit vector is a linear combination of* $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$*.*

So if we want to show that

$$\underline{\mathbf{a}}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \qquad \underline{\mathbf{a}}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \qquad \underline{\mathbf{a}}_3 = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}$$

span $\mathbb{R}^m$, then all we need to show is that the unit vectors

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

belong to the spanning set.

Example 4.12. Show that the linear system

$$\begin{cases} 2x_1 + x_2 + 2x_3 = b_1 \\ x_1 + 2x_2 - x_3 = b_2 \\ x_1 + x_3 = b_3 \end{cases}$$

has a solution for every $\underline{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

We have the coefficient matrix

$$A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

and the three column vectors

$$\underline{\mathbf{a}}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \qquad \underline{\mathbf{a}}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \qquad \underline{\mathbf{a}}_3 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

We want to show that the three standard unit vectors are linear combinations of our column vectors.

1. We see that

$$\underline{\mathbf{a}}_1 - \underline{\mathbf{a}}_3 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

so

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2}\underline{\mathbf{a}}_1 - \frac{1}{2}\underline{\mathbf{a}}_3.$$

2. Next, we see that

$$\underline{\mathbf{a}}_2 - \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

so

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \underline{\mathbf{a}}_2 - 2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \underline{\mathbf{a}}_2 - 2\left(\frac{1}{2}\underline{\mathbf{a}}_1 - \frac{1}{2}\underline{\mathbf{a}}_3\right) = -\underline{\mathbf{a}}_1 + \underline{\mathbf{a}}_2 + \underline{\mathbf{a}}_3.$$

3. Lastly, we see that

$$\begin{aligned}\begin{bmatrix}0\\0\\1\end{bmatrix} &= \underline{\mathbf{a}}_1 - 2\begin{bmatrix}1\\0\\0\end{bmatrix} - \begin{bmatrix}0\\1\\0\end{bmatrix} = \underline{\mathbf{a}}_1 - 2\left(-\underline{\mathbf{a}}_1 + \underline{\mathbf{a}}_2 + \underline{\mathbf{a}}_3\right) - \left(\frac{1}{2}\underline{\mathbf{a}}_1 - \frac{1}{2}\underline{\mathbf{a}}_3\right)\\ &= \frac{5}{2}\underline{\mathbf{a}}_1 - 2\underline{\mathbf{a}}_2 - \frac{3}{2}\underline{\mathbf{a}}_3.\end{aligned}$$

4. Since every standard unit vector is a linear combination of the column vectors, we conclude that the linear system has a unique solution for every possible vector $\underline{\mathbf{b}}$.

4.5 Summary of Results

We can summarize the main results of the first two lectures accordingly:

Theorem 4.13. *Let A be a $m \times n$ matrix. The following statements are equivalent:*

1. *Every linear system with A as its coefficient matrix has a unique solution.*
2. *A has a pivot position in every row.*
3. *A is invertible.*
4. *For each $\underline{\mathbf{b}}$ in $\mathbb{R}^m$, the equation $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ has a unique solution.*
5. *The columns of A span $\mathbb{R}^m$.*