3. Lecture 3: (1.2) 1.3 Vector Equations and Matrix multiplication

In linear algebra we think of vectors in $\mathbb{R}^n$ as column vectors or $n \times 1$ matrices

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

**Addition** and **scalar multiplication** are defined by

$$u + v = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}, \quad \lambda u = \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \\ \vdots \\ \lambda u_n \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

Given vectors $v_1, \ldots, v_k$ and scalars $\lambda_1, \ldots, \lambda_k$, the vector

$$w = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

is called a **linear combination** of the vectors $v_1, \ldots, v_k$, with weights $\lambda_1, \ldots, \lambda_k$.

The first question we will ask today is: Given a vector $w$ and vectors $v_1, \ldots, v_k$, can we find scalars $\lambda_1, \ldots, \lambda_k$, such that $w$ is a linear combination of $v_1, \ldots, v_k$?

In $\mathbb{R}^2$ and $\mathbb{R}^3$ we have a geometric notion of vector addition and scalar multiplication. We think of vectors as arrows with a length and a direction.

The parallelogram law says that the sum $u + v$ is given by placing the start of $v$ where $u$ ends. Check this by drawing $u = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $u + v = \begin{bmatrix} 1 + 2 \\ 3 + 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

If $\lambda > 0$ then scalar multiplication $\lambda u$ is the vector in the same direction as $u$ with length $\lambda$ times the length of $u$. Check this by drawing $u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $2u = \begin{bmatrix} 2 \cdot 1 \\ 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

**Ex** Let $v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Express $b = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ as linear combinations of $v_1$ and $v_2$.

**Sol.** We start by drawing a net of parallelograms with sides $v_1$ and $v_2$. Then we see how far we should go first in the $v_1$ and then in the $v_2$ direction to reach $b$. We see that $b = v_1 + 2v_2$.

**Ex** Let $v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$. Express $b = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$ as linear combinations of $v_1$ and $v_2$.

**Sol.** Since $v_2 = 2v_1$ we can only reach vectors $b$ which are on the line $tv_1$ for some $t$, but this can not be equal to be for any $t$ since one of the components of $b$ vanishes.
Ex Let $a_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $a_2 = \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix}$, $a_3 = \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix}$, and $b = \begin{bmatrix} -1 \\ 8 \\ -5 \end{bmatrix}$.

Determine if $b$ is a linear combination of $a_1$, $a_2$, $a_3$.

**Sol** $b$ is a linear combination of $a_1$, $a_2$, and $a_3$ if we can find scalars $x_1, x_2, x_3$ so

$$x_1 a_1 + x_2 a_2 + x_3 a_3 = b.$$ 

If we write it out we get the vector equation

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ -5 \end{bmatrix}.$$ 

If we add up the vectors to the left we get

$$\begin{bmatrix} x_1 + 4x_2 + 3x_3 \\ 2x_2 + 6x_3 \\ 3x_1 + 14x_2 + 10x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ -5 \end{bmatrix}.$$ 

i.e. we get a linear system of equations

$$x_1 + 4x_2 + 3x_3 = -1,$$
$$2x_2 + 6x_3 = 8,$$
$$3x_1 + 14x_2 + 10x_3 = -5.$$ 

The corresponding augmented matrix is

$$\begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 2 & 6 & 8 \\ 3 & 14 & 10 & -5 \end{bmatrix} \iff \begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 2 & 6 & 8 \\ 0 & 1 & 3 & 1 -2 \end{bmatrix} \iff \begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\iff \begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\iff \begin{bmatrix} 1 & 0 & -9 & -17 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\iff \begin{bmatrix} 1 & 0 & -9 & -17 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

i.e. we get the system

$$x_1 = 1,$$
$$x_2 = -2,$$
$$x_3 = 2$$

and hence

$$b = a_1 - 2a_2 + 2a_3.$$ 

Note that $a_1$, $a_2$, $a_3$ and $b$ are columns of the augmented matrix $[a_1 \ a_2 \ a_3 \ b]$. Hence solving the vector equation $b = x_1 a_1 + x_2 a_2 + x_3 a_3$ is the same as solving the linear system with augmented matrix $[a_1 \ a_2 \ a_3 \ b]$.

In general the vector equation

$$b = x_1 a_1 + \cdots + x_k a_k$$

has the same solution set as the linear system with augmented matrix

$$[a_1 \ \cdots \ a_k \ b]$$

i.e. $b$ can be generated as a linear combination of $a_1, \cdots, a_k$ if and only if there is a solution to the linear system with the corresponding augmented matrix.
Matrix Multiplication

Recall that the dot product of two vectors \( \mathbf{w} = [w_1 \ w_2 \ \ldots \ w_n] \) and \( \mathbf{v} = [v_1 \ v_2 \ \ldots \ v_n] \) is

\[
\mathbf{w} \cdot \mathbf{v} = w_1v_1 + w_2v_2 + \cdots + w_nv_n
\]

One can think of linear system

\[
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\]

(3.1)

as a single vector equation in matrix form

\[
A\mathbf{x} = \mathbf{b},
\]

(3.2)

where

\[
A = \begin{bmatrix}
  a_{11} & a_{12} & \ldots & a_{1n} \\
  a_{21} & a_{22} & \ldots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \ldots & a_{mn}
\end{bmatrix}, \quad
\mathbf{x} = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}, \quad
\mathbf{b} = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}
\]

and matrix multiplication \( A\mathbf{x} \) of the \( m \times n \) matrix \( A \) and the \( n \times 1 \) column vector \( \mathbf{x} \) is defined to be the \( m \times 1 \) column vector formed from the dot product of the row vectors \( \mathbf{w}_i = [a_{i1} \ a_{i2} \ \ldots \ a_{in}] \) of \( A \) with the column vector \( \mathbf{x} \):

\[
A\mathbf{x} = \begin{bmatrix}
  w_1 \cdot \mathbf{x} \\
  w_2 \cdot \mathbf{x} \\
  \vdots \\
  w_m \cdot \mathbf{x}
\end{bmatrix}, \quad \text{if} \quad A = \begin{bmatrix}
  -w_1 & -w_2 & \ldots & -w_m
\end{bmatrix}
\]

(3.3)

Then (3.2) says that the column vector \( A\mathbf{x} \) is equal to the column vector \( \mathbf{b} \). (3.1) just says that the components of these column vectors are equal.

The product (3.3) can be written as linear combination of the column vectors \( \mathbf{v}_i \) of \( A \)

\[
A\mathbf{x} = x_1\begin{bmatrix} a_{11} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \end{bmatrix} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_m, \quad \text{if} \quad A = \begin{bmatrix} \mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_m \end{bmatrix}
\]

(3.4)

(3.4) is multiplication by columns. (3.3) is multiplication by rows.

Multiplying by the \( m \times n \) matrix \( A \) hence defines a map \( \mathbf{f} : \mathbb{R}^n \ni \mathbf{x} \rightarrow A\mathbf{x} \in \mathbb{R}^m \), for each \( n \times 1 \) column vector \( \mathbf{x} \) we get an \( m \times 1 \) column vector \( A\mathbf{x} \) defined by (3.3). The map is linear; \( A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}, \quad A(\lambda \mathbf{x}) = \lambda A\mathbf{x} \). All linear maps are of this form.
1.5 in Lay et al Solution sets of linear systems

A homogeneous system is a system of the form

\[ Ax = 0 \]

where \( A \) is an \( m \times n \) matrix and \( 0 \) is the zero vector in \( \mathbb{R}^m \).

A homogeneous system always has the trivial solution \( x = 0 \) so it’s consistent. Consistent systems with a free variable have infinitely many solutions.

A homogeneous system has a nontrivial solution \( x \neq 0 \) if and only if it has free variables.

**Ex 1** Describe the solution set of the system

\[
\begin{align*}
  x_1 + 2x_2 - 3x_3 &= 0 \\
  4x_1 + 8x_2 - 11x_3 &= 0
\end{align*}
\]

**Sol** There is at least one free variable since there are 2 equations in 3 variables.

\[
\begin{bmatrix}
  1 & 2 & -3 & 0 \\
  4 & 8 & -11 & 0
\end{bmatrix} \sim -4(1) \begin{bmatrix}
  1 & 2 & -3 & 0 \\
  0 & 0 & 1 & 0
\end{bmatrix} \sim +3(1) \begin{bmatrix}
  1 & 2 & 0 & 0 \\
  0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
x_1 + 2x_2 = 0 \\
x_3 = 0
\]

\[
x = \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix} = \begin{bmatrix}
  -2x_2 \\
  x_2 \\
  0
\end{bmatrix} = \begin{bmatrix}
  -2 \\
  1 \\
  0
\end{bmatrix} = x_2 v
\]

This is the parametric equation of a line through 0 in the direction of \( v \).

**Ex 2** Determine the solution set of

\[
\begin{align*}
  x_1 + 2x_2 - 3x_3 &= 0 \\
  4x_1 + 8x_2 - 11x_3 &= 2
\end{align*}
\]

**Sol**

\[
\begin{bmatrix}
  1 & 2 & -3 & 0 \\
  4 & 8 & -11 & 2
\end{bmatrix} \sim -4(1) \begin{bmatrix}
  1 & 2 & -3 & 0 \\
  0 & 0 & 1 & 2
\end{bmatrix} \sim +3(1) \begin{bmatrix}
  1 & 2 & 0 & 6 \\
  0 & 0 & 1 & 2
\end{bmatrix}
\]

\[
x = \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix} = \begin{bmatrix}
  6 - 2x_2 \\
  x_2 \\
  2
\end{bmatrix} = \begin{bmatrix}
  6 \\
  0 \\
  2
\end{bmatrix} + \begin{bmatrix}
  -2 \\
  1 \\
  0
\end{bmatrix} = p + x_2 v, \quad x_2 \text{ is a free parameter.}
\]

This is the parametric equation of a line through \( p \) in the direction of \( v \), parallel to Ex 1.

If the nonhomogeneous equation \( Ax = b \) is consistent its solution set is parallel to the solution set to the homogeneous equation \( Ax = 0 \).

**Th** Suppose \( Ax = b \) is consistent and let \( p \) be a solution. Then any other solution \( x = p + v_h \), where \( v_h \) is a solution to \( Av_h = 0 \).

**Pf** Since matrix multiplication is linear \( A(p + v_h) = Ap + Av_h = b \).

**Ex** Describe the solution set to \( x_1 - 2x_2 - 2x_3 = b \) for \( b = 0, 1 \).

**Sol** \( x_2 \) and \( x_3 \) are free variables

\[
x = \begin{bmatrix}
  b + 2x_2 + 2x_3 \\
  x_2 \\
  x_3
\end{bmatrix} = \begin{bmatrix}
  b \\
  0 \\
  0
\end{bmatrix} + \begin{bmatrix}
  2 \\
  1 \\
  0
\end{bmatrix} + \begin{bmatrix}
  2 \\
  0 \\
  1
\end{bmatrix} = p + x_2 v_1 + x_3 v_2
\]

Since \( x_2 \) and \( x_3 \) are free parameters this is the parametric vector equation of a plane through \( p \) and parallel to the vectors \( v_1 \) and \( v_2 \).

If \( b = 0 \) its the plane spanned by \( v_1, v_2 \) and if \( b \neq 0 \) it is a plane parallel to this plane.
Summary and Conceptual Questions

We can now write a linear system with augmented matrix
\[
\begin{bmatrix}
2 & 3 & 4 & 9 \\
-3 & 1 & 0 & -2 \\
\end{bmatrix},
\] (3.5)
as a System of Linear Equations

\[
\begin{align*}
2x_1 + 3x_2 + 4x_3 &= 9 \\
-3x_1 + x_2 &= -2
\end{align*}
\] (3.6)
as a Vector Equation

\[
x_1 \begin{bmatrix} 2 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 9 \\ -2 \end{bmatrix}
\] (3.7)
or as a Matrix Equation

\[
\begin{bmatrix}
2 & 3 & 4 \\
-3 & 1 & 0 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
\end{bmatrix} = \begin{bmatrix}
9 \\
-2 \\
\end{bmatrix}
\]

Viewing the system as the intersection of planes (3.6) is called the row picture since each equation corresponds to a row of the augmented matrix (3.5).

Viewing the system as a linear combination of vectors (3.7) is called the column picture since each vector corresponds to a column of the augmented matrix (3.5).

Matrix multiplication can be calculated in two ways corresponding to the row picture:

\[
\begin{bmatrix}
1 & -4 \\
3 & 2 \\
0 & 5 \\
\end{bmatrix}
\begin{bmatrix}
7 \\
-6 \\
\end{bmatrix} = \begin{bmatrix}
1 \cdot 7 + (-4) \cdot (-6) \\
3 \cdot 7 + 2 \cdot (-6) \\
0 \cdot 7 + 5 \cdot (-6) \\
\end{bmatrix} = \begin{bmatrix}
31 \\
9 \\
-30 \\
\end{bmatrix}.
\]

respectively the column picture:

\[
\begin{bmatrix}
1 & -4 \\
3 & 2 \\
0 & 5 \\
\end{bmatrix}
\begin{bmatrix}
7 \\
-6 \\
\end{bmatrix} = 7 \begin{bmatrix}
1 \\
3 \\
0 \\
\end{bmatrix} - 6 \begin{bmatrix}
-4 \\
2 \\
5 \\
\end{bmatrix} = \begin{bmatrix}
7 \\
21 \\
0 \\
\end{bmatrix} + \begin{bmatrix}
-4 \\
24 \\
-12 \\
\end{bmatrix} = \begin{bmatrix}
31 \\
9 \\
-30 \\
\end{bmatrix}.
\]

An \(n \times n\) matrix \(A\) gives a linear map \(\mathbb{R}^n \ni x \rightarrow Ax \in \mathbb{R}^n\) i.e. \(A(x+y) = Ax+Ay, A(\lambda x) = \lambda A(x)\).

If \(A\) is an \(n \times n\) matrix, we have learned that the linear system

\[
Ax = b
\]

can be solved uniquely for every \(b\) iff the reduced row echelon form has only 1’s in the diagonal.

An \(n \times n\) matrix \(B\) is called an inverse of an \(n \times n\) \(A\) if the solution to \(Ax = b\) can be give as \(x = Bb\)

**Question** When does an \(n \times n\) matrix \(A\) have an inverse? Hint: Try the \(2 \times 2\) case.

The system \(Ax = 0\) is called the homogeneous equation:

\[
Ax = 0
\]

It always has the trivial solution \(x = 0\). A solution \(x \neq 0\) is called nontrivial.

**Question** Let \(A\) be an \(n \times n\) matrix. When does the system \(Ax = 0\) have a nontrivial solution?