**MATRIX ALGEBRA, Systems Linear Equations**

Now we change to the LINEAR ALGEBRA perspective on vectors and matrices to reformulate systems of linear equations. If you find the discussion in terms of general \( m \) and \( n \) gets lost in generality, think first of what it all means in the cases \( m = 2 \) and \( n = 2 \) or \( n = 3 \) as we often do in class at first. Many of the homework questions are also formulated for such small values of \( m, n \)! In practice, however, \( m \) and \( n \) can be huge - that's where computers come in handy!! Let's first concentrate on the algebra half of the term "Linear Algebra".

- A **Vector** is an ordered list of numbers (real or complex)

\[
\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad \mathbb{R}^n = \text{all ordered lists of } n \text{ real numbers,} \\
\mathbb{C}^n = \text{all ordered lists of } n \text{ complex numbers.}
\]

The set of right hand values \( b_1, b_2, \ldots, b_m \) and the set \( x_1, x_2, \ldots, x_n \) of variables in a system of \( m \) linear equations in \( n \) variables can be identified with vectors

\[
\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{in } \mathbb{R}^m, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{in } \mathbb{R}^n,
\]

for example, but many others occur:

<table>
<thead>
<tr>
<th>The set of sample values</th>
<th>The list of 8 roots of unity</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{v} = \begin{bmatrix} 1 \ \frac{1}{\sqrt{2}} \ 0 \ \vdots \ \frac{1}{\sqrt{2}} \ 1 \end{bmatrix} )</td>
<td>( \mathbf{v} = \begin{bmatrix} 1 \ \frac{1}{\sqrt{2}}(1+i) \ i \ \vdots \ \frac{1}{\sqrt{2}}(1-i) \end{bmatrix} )</td>
</tr>
<tr>
<td>of the function ( y = \cos x ) at ( 0, \frac{\pi}{4}, \frac{\pi}{2}, \ldots, \frac{7\pi}{4}, \pi ) as shown above is a vector in ( \mathbb{R}^8 ); a CD is a huge vector of this sort obtained by sampling an audio signal 44,000 times per second.</td>
<td>as shown above on the unit circle in the complex plane is a vector in ( \mathbb{C}^8 ). We shall meet vectors like this when we get to eigenvectors of matrices.</td>
</tr>
</tbody>
</table>

The ‘algebra’ of vectors refers to the addition, subtraction, and scalar multiplication of vectors.

- We can **add/subtract** vectors \( \mathbf{u}, \mathbf{v} \) in \( \mathbb{R}^n \) by \( \mathbf{u} \pm \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \pm \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 \pm v_1 \\ u_2 \pm v_2 \\ \vdots \\ u_n \pm v_n \end{bmatrix}, \)
- and form the **scalar multiple** of \( k \) in \( \mathbb{R} \) and \( \mathbf{v} \) in \( \mathbb{R}^n \) by \( k\mathbf{v} = k \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{bmatrix}; \)

in other words, calculations are done **component-wise**. There are similar definitions of addition and scalar multiplication for
vectors in \( \mathbb{C}^n \), except that now scalar multiplication by complex numbers is allowed. When trying to interpret concepts and results graphically it’s often convenient to identify vectors \( \mathbf{x} \) in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \) with points \( \mathbf{x} = (x_1, x_2) \) in the plane or \( \mathbf{x} = (x_1, x_2, x_3) \) in 3-space. The addition/subtraction/scalar multiplication just defined for vectors in column form then corresponds to the ones you learned for vectors in the plane and 3-space.

But what’s the point of thinking of vectors in column form?

**Example:** given vectors

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

in \( \mathbb{R}^n \), and

\[
\begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix}, \quad
\begin{bmatrix}
  a_{12} \\
  a_{22} \\
  \vdots \\
  a_{m2}
\end{bmatrix}, \ldots, \quad
\begin{bmatrix}
  a_{1n} \\
  a_{2n} \\
  \vdots \\
  a_{mn}
\end{bmatrix}
\]

in \( \mathbb{R}^m \),

then the operations of addition and scalar multiplication for vectors show that

\[
\begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix} + \begin{bmatrix}
  a_{12} \\
  a_{22} \\
  \vdots \\
  a_{m2}
\end{bmatrix} + \ldots + \begin{bmatrix}
  a_{1n} \\
  a_{2n} \\
  \vdots \\
  a_{mn}
\end{bmatrix} = \begin{bmatrix}
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
  \vdots \\
  a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{bmatrix},
\]

so the usual system of \( m \) linear equations in \( n \) variables can be rewritten as the

**Vector Equation:**

\[
x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{b} = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}.
\]

- An \( m \times n \) **Matrix** with real or complex numbers as entries can be thought of both as an **Array** of numbers and as a **row of column vectors**

\[
A = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
  a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
  \vdots & \vdots & \vdots & \cdots & \vdots \\
  a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}
\end{bmatrix} = \begin{bmatrix}
  a_1 \\
  a_2 \\
  \vdots \\
  a_n
\end{bmatrix}
\]

where

\[
\mathbf{a}_1 = \begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix}, \quad
\mathbf{a}_2 = \begin{bmatrix}
  a_{12} \\
  a_{22} \\
  \vdots \\
  a_{m2}
\end{bmatrix}, \ldots, \quad
\mathbf{a}_n = \begin{bmatrix}
  a_{1n} \\
  a_{2n} \\
  \vdots \\
  a_{mn}
\end{bmatrix}
\]

are \( n \) column vectors in \( \mathbb{R}^m \) or \( \mathbb{C}^m \). The set of all \( m \times n \) matrices with real entries will be denoted by \( \mathbb{R}^{m\times n} \) and those with complex entries by \( \mathbb{C}^{m\times n} \). To keep from drowning in notation, it’s common to write a matrix as \( A = \left[ a_{jk} \right] \) instead of writing out all the entries in \( A \). Notice that the set \( \mathbb{R}^{n\times 1} \) of all matrices consisting of one column of \( n \) real entries is just another way of thinking of \( \mathbb{R}^n \), while \( \mathbb{C}^{n\times 1} \) is just another way of thinking of \( \mathbb{C}^n \).

The term **algebra of matrices** refers to addition, subtraction, and multiplication of matrices. In a future lecture we’ll also
discuss the idea of the inverse of a matrix.

We add/subtract matrices $A$, $B$ in $\mathbb{R}^{m \times n}$ entry-by-entry:

$$A + B = [a_{jk}] + [b_{jk}] = [a_{jk} + b_{jk}].$$

In particular, $A + B$ is defined only when $A$, $B$ are $m \times n$, and then $A + B$ also is $m \times n$. For example,

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

$$\begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & -1 \end{bmatrix} - \begin{bmatrix} 4 & -1 & 1 \\ 2 & -4 & -1 \end{bmatrix} = \begin{bmatrix} -3 & 3 & 2 \\ -5 & 8 & 0 \end{bmatrix}.$$

We define the scalar multiple of $k$ in $\mathbb{R}$ and $A$ in $\mathbb{R}^{m \times n}$ by

$$kA = k[a_{jk}] = [ka_{jk}].$$

Thus each entry in $A$ is multiplied by $k$; in particular, the scalar multiple $kA$ of an $m \times n$ matrix $A$ also is $m \times n$. For example,

$$3\begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ -9 & 12 & -3 \end{bmatrix}.$$

To introduce multiplication, let’s begin with the product of a matrix and a vector:

**Matrix-vector Rule:** if $A = [a_1 \ a_2 \ldots a_n]$ with columns $a_1$, $a_2$, $\ldots$, $a_n$ in $\mathbb{R}^m$ and $x$ is a vector in $\mathbb{R}^n$, then

$$Ax = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1a_1 + x_2a_2 + \ldots + x_na_n.$$

But what’s the point of this definition? Well, if $A$ is the coefficient matrix of the usual system of $m$ linear equations in $n$ variables, and if $A$ is written in column form $A = [a_1 \ a_2 \ldots a_n]$, then

$$Ax = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1a_1 + x_2a_2 + \ldots + x_na_n = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix},$$

so the usual system of $m$ linear equations in $n$ variables can be rewritten as the

**Matrix Equation:**

$$Ax = b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Thus the algebra of vectors and matrices provides three different ways of writing a system of $m$ linear equations in $n$ unknowns: as an augmented matrix, a vector equation and a single matrix equation. In other words, by using vectors and
matrices we have made things just as compact as the equations \( a_1 x + a_2 y = b \) and \( ax = b \) studied in high school, yet we can handle systems in hundreds or thousands of variables. The next theorem says that each provides a way of interpreting and solving a system of linear equations.

**Fundamental Theorem:** if \( \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \) and \( \mathbf{b} \) are vectors in \( \mathbb{R}^m \), then each of the following

- the system of linear equations with augmented matrix \( [ \mathbf{a}_1 \ \mathbf{a}_2 \ \ldots \ \mathbf{a}_n \ \mathbf{b} ] \),
- the vector equation \( x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \ldots + x_n \mathbf{a}_n = \mathbf{b} \),
- the matrix equation \( \mathbf{A} \mathbf{x} = \mathbf{b} \) when \( \mathbf{A} = [ \mathbf{a}_1 \ \mathbf{a}_2 \ \ldots \ \mathbf{a}_n ] \),

has the same solution set.

In practice, to solve a given system of linear equations it’s probably quickest and easiest to write the associated augmented matrix in Reduced Row Echelon Form, especially if one has an electronic way of arriving at this form. Nonetheless, the other two ways of interpreting systems of linear equations will become very important conceptually.

Finally, to define products of matrices quite generally, we now simply think of a matrix in column form and then use the **Matrix-Column Rule:**

**Matrix-Column Rule:** if \( \mathbf{A} \) is an \( m \times p \) and \( \mathbf{B} \) is an \( p \times n \) matrix written in column form \( \mathbf{B} = [ \mathbf{b}_1 \ \mathbf{b}_2 \ \ldots \ \mathbf{b}_n ] \), then the product \( \mathbf{A} \mathbf{B} \) of \( \mathbf{A} \) and \( \mathbf{B} \) is the \( m \times n \) matrix \( \mathbf{A} \mathbf{B} \) defined by

\[
\mathbf{A} \mathbf{B} = \mathbf{A} [ \mathbf{b}_1 \ \mathbf{b}_2 \ \ldots \ \mathbf{b}_n ] = [ \mathbf{A} \mathbf{b}_1 \ \mathbf{A} \mathbf{b}_2 \ \ldots \ \mathbf{A} \mathbf{b}_n ].
\]

Notice that the restrictions: \( \mathbf{A} \) in \( \mathbb{R}^{m \times p} \) and \( \mathbf{B} \) in \( \mathbb{R}^{p \times n} \) are needed so the columns \( \mathbf{b}_j \) of \( \mathbf{B} \) are in \( \mathbb{R}^p \) and the matrix-vector product \( \mathbf{A} \mathbf{b}_j \) is defined as a vector in \( \mathbb{R}^m \). Since there are \( n \) columns in \( \mathbf{B} \) the product \( \mathbf{A} \mathbf{B} \) thus has \( n \) columns each in \( \mathbb{R}^m \). Hence \( \mathbf{A} \mathbf{B} \) is in \( \mathbb{R}^{m \times n} \).

**Example:** compute \( \mathbf{A} \mathbf{B} \) when

\[
\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 4 & 2 & -3 \\ -1 & 2 & 6 \end{bmatrix}.
\]

**Solution:** Write \( \mathbf{B} = [ \mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3 ] \). Then

\[
\mathbf{A} \mathbf{b}_1 = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix},
\]

while

\[
\begin{align*}
\mathbf{A} \mathbf{b}_2 &= \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ 14 \end{bmatrix}, \\
\mathbf{A} \mathbf{b}_3 &= \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -3 \\ 6 \end{bmatrix} = -3 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 6 \begin{bmatrix} -1 \\ -4 \end{bmatrix} = \begin{bmatrix} -12 \\ -15 \end{bmatrix}.
\end{align*}
\]

Thus

\[
\mathbf{A} \mathbf{B} = [ \mathbf{A} \mathbf{b}_1 \ \mathbf{A} \mathbf{b}_2 \ \mathbf{A} \mathbf{b}_3 ] = \begin{bmatrix} 9 & 2 & -12 \\ 8 & 14 & -15 \end{bmatrix}.
\]

The ‘algebra’ part of the term ‘linear algebra’ has thus been explored. But what about the ‘linear’ part. Well, one of the key
ideas underlying what we've done is the fact that matrices and vectors have a fundamental property that you met originally just for vectors in the plane, say: given vectors \( \mathbf{u}, \mathbf{v} \) in the plane and scalars \( a, b \), then the **Linear Combination** \( a\mathbf{u} + b\mathbf{v} \) is again a vector in the plane which can be defined both in terms of its \( \mathbf{i} \) and \( \mathbf{j} \) components as well as graphically in terms of the parallelogram and triangle laws for adding vectors. But you may recall that you also met linear combinations \( af(x) + bg(x) \) of functions when dealing with properties of limits, derivatives and integrals. What we've just seen is that the notion of linear combination makes good sense for vectors in \( \mathbb{R}^n \) and \( \mathbb{C}^n \) as well as for matrices in \( \mathbb{R}^{m\times n} \) and \( \mathbb{C}^{m\times n} \). **Linearity** is fundamental concept occurring everywhere in mathematics and its applications. Later we shall formalize these ideas abstractly by introducing the notion of **Vector space** (or **Linear Space** as it is also called).