The material from this lesson is drawn from "Linear Algebra and its Application", 5th Edition, by Lay, "A First Course in Linear Model Theory" by Ravishanker and Dey, and from Dr. Steve MacEachern's Spring 2014 course "STAT 6860: Foundations of the Linear Model" at The Ohio State University.

I. Linear Independence

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is **linearly independent** if the equation

$$\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_n\mathbf{v}_n = \mathbf{0}$$

has only the trivial solution. (That is, the only way that this equation holds true is if $\beta_1 = \beta_2 = \cdots = \beta_n = 0$.)

This is equivalent to saying that each vector $\mathbf{a}_i$ **cannot** be written as a linear combination of the other vectors $\mathbf{a}_j$. (Remember: we say that W is a **linear combination** of Z if we can write W as $a + bZ$, where a and b are some real numbers.)

If two vectors are not linearly independent, then they are **linearly dependent**.

Example: Consider the following vectors.

$$\mathbf{v_1} = \begin{bmatrix}1\\2\\3\end{bmatrix} \qquad \mathbf{v_2} = \begin{bmatrix}4\\5\\6\end{bmatrix} \qquad \mathbf{v_3} = \begin{bmatrix}2\\1\\0\end{bmatrix}$$

Vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. This is because there exists no β_1 and β_2 such that $\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 = \mathbf{0}$.

Similarly, vectors $\mathbf{v}_1$ and $\mathbf{v}_3$ are linearly independent. Vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ are also linearly independent. *(Try showing this on your own.)*

However, vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly dependent. Consider the equation $\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \beta_3\mathbf{v}_3$ where $\beta_1 = -2$, $\beta_2 = 1$, and $\beta_3 = -1$:

$$\begin{aligned}
\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \beta_3\mathbf{v}_3 &:= -2\mathbf{v}_1 + 1\mathbf{v}_2 + -1\mathbf{v}_3 \\
&= -2 \times \begin{bmatrix}1\\2\\3\end{bmatrix} + 1 \times \begin{bmatrix}4\\5\\6\end{bmatrix} + -1 \times \begin{bmatrix}2\\1\\0\end{bmatrix} \\
&= \begin{bmatrix}-2\\-4\\-6\end{bmatrix} + \begin{bmatrix}4\\5\\6\end{bmatrix} + \begin{bmatrix}-2\\-1\\0\end{bmatrix} \\
&= \begin{bmatrix}-2+4+-2\\-4+5+-1\\-6+6+0\end{bmatrix} \\
&= \begin{bmatrix}0\\0\\0\end{bmatrix} \\
&= \mathbf{0}
\end{aligned}$$

Thus, by the definition of linear dependence, $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly dependent. (Equivalently, these three vectors are NOT linearly independent.)

We can append $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ into one 3×3 matrix $\mathbf{V}$:

$$V = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 5 & 1 \\ 3 & 6 & 0 \end{bmatrix}$$

Because the columns of $\mathbf{V}$ are linearly dependent, we say that $\mathbf{V}$ is not of full rank. If the columns of $\mathbf{V}$ were linearly independent, then we would say that $\mathbf{V}$ is of full rank. (The term "full rank" is widely used in resources.)

TO RECAP:

Suppose $\mathbf{A}$ is an $n \times n$ matrix with columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$. Then, the following are equivalent:

- The matrix $\mathbf{A}$ is of full column rank.
- The columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ are linearly independent.
- None of the columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ are linear combinations of one another.
- The equation $\beta_1\mathbf{a}_1 + \beta_2\mathbf{a}_2 + \cdots + \beta_n\mathbf{a}_n = \mathbf{0}$ only holds true if $\beta_1 = \beta_2 = \cdots = \beta_n = 0$.
- The matrix $\mathbf{A}$ is invertible.
- There exists some matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$.
- The determinant of matrix $\mathbf{A}$, $\det(\mathbf{A})$, is nonzero.
- For any vector $b_{n\times 1}$ and a vector of unknowns $x_{n\times 1}$, there is at least one solution to the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$.

As such, the following are also equivalent:

- The matrix $\mathbf{A}$ is NOT of full column rank.
- The columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ are NOT linearly independent.
- At least one of the columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ is a linear combination of other columns.
- The equation $\beta_1\mathbf{a}_1 + \beta_2\mathbf{a}_2 + \cdots + \beta_n\mathbf{a}_n = \mathbf{0}$ holds true for at least some $\beta_i \neq 0$.
- The matrix $\mathbf{A}$ is NOT invertible.
- There does NOT exist some matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$.
- The determinant of matrix $\mathbf{A}$, $\det(\mathbf{A})$, IS zero.
- For any vector $b_{n\times 1}$ and a vector of unknowns $x_{n\times 1}$, there is no solution to the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$.

This is largely encapsulated in this Dartmouth link about the "Invertible Matrix Theorem."

Clearly, whether or not the columns of our matrix are independent of one another is of great importance. In future lectures, we will learn about a method by which we can ensure that the columns of our matrix of independent variables $\mathbf{X}$ will always be independent. However, before that, we have to discuss linear transformations.

II. Linear Transformations

A linear transformation f is a function that preserves additivity and scalar multiplication. That is, given a linear transformation f, vectors $\mathbf{x}$ and $\mathbf{y}$, and a scalar k, then:

$$\begin{aligned} f(\mathbf{x}+\mathbf{y}) &= f(\mathbf{x}) + f(\mathbf{y}) \\ f(k\mathbf{x}) &= k \times f(\mathbf{x}) \end{aligned}$$

The particular definition of linear transformation isn't critical to remember moving forward, but suffice it to say that linear transformations are very important.

All linear transformations can be represented by a matrix $\mathbf{A}_{n\times n}$. Thus, if we want to apply some linear transformation to $\mathbf{x}_{n\times 1}$, then we simply need to evaluate $\mathbf{A}\mathbf{x}$.

Consider the vector $\mathbf{x} = [2 \;\; 1]^T$. We might think of this geometrically, as a vector starting at the origin and pointing to the coordinate point $(x_1, x_2) = (2, 1)$. Could we find some $\mathbf{A}$ that would reflect this vector over the x_1-axis? (The horizontal axis.)

The answer is yes.

Take the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then,

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \times 2 + 0 \times 1 \\ 0 \times 2 + -1 \times 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -1 \end{bmatrix} \end{aligned}$$

What if we wanted to flip the vector $\mathbf{x} = [2 \;\; 1]^T$ over the x_2-axis? (The vertical axis.) We could indeed find a matrix $\mathbf{A}$ for that as well.

Try evaluating $\mathbf{Ax}$ on your own with the matrix $\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

The natural question here might be, "How do you find that matrix $\mathbf{A}$ that gives us the transformation we want?"

Recall the $n \times n$ identity matrix, denoted $\mathbf{I}_n$. The columns of $\mathbf{I}_n$ are just vectors of length one in each x_i direction. (For example, the first column is $[1 \;\; 0 \;\; 0 \;\; \cdots \;\; 0]^T$, the second column is $[0 \;\; 1 \;\; 0 \;\; \cdots \;\; 0]^T$, and so on.) We denote these vectors of length one, called unit vectors, as e_1, e_2, $\ldots$, where e_1 is the unit vector in the x_1 direction, e_2 is the unit vector in the x_2 direction, and so on.

In order to construct the matrix $\mathbf{A}$ that transforms $\mathbf{x}$ in the way that we want, we should begin with the matrix $\mathbf{I}_n$. For each column, replace the vector e_i with the transformed value of e_i, denoted $f(e_i)$.

For example, in the case where we want to flip our vector over the x_1-axis, e_1 would stay the same because reflecting e_1 over the x_1-axis will have no effect. However, e_2 will flip from $[1 \;\; 0]^T$ to $[-1 \;\; 0]^T$, so we replace the second column of $\mathbf{I}_n$ with $[-1 \;\; 0]^T$. We repeat this for each unit vector e_i.

To summarize, the columns of our matrix $\mathbf{A}$ will be the transformed values of each column of $\mathbf{I}_n$:

$$\begin{aligned} \mathbf{I}_n &= [e_1 \;\; e_2 \;\; \cdots \;\; e_n] \\ \Rightarrow \mathbf{A} &= [f(e_1) \;\; f(e_2) \;\; \cdots \;\; f(e_n)] \end{aligned}$$

To wrap up, there are a certain class of linear transformations that will be of particular interest to us. For each matrix $\mathbf{A}$, there may be vectors $\mathbf{x}$ such that $\mathbf{Ax}$ is just a scalar multiple of $\mathbf{x}$. (We most frequently write this $\mathbf{Ax} = \lambda\mathbf{x}$.) The vectors $\mathbf{x}$ and the scalars λ for which this equation holds true are called **eigenvectors** and **eigenvalues**, and their applications are very important to the method by which we will ensure that the columns of our matrix of independent variables will always be independent.