Unless otherwise cited, 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. Framing

- In addition to familiarity with basic linear algebra and matrices concepts/operations, this sub-curriculum is designed to enhance your understanding of four topics:
  A. Assumptions of the Linear Model
  B. Computational Challenges with Modeling
  C. Principal Component Analysis
  D. Markov Chains
- There are six lessons within this sub-curriculum:
  A. Matrices
  B. From Matrices to Linear Modeling
  C. From Modeling by Hand to Computational Modeling
  D. From Modeling to Assumptions
  E. From Assumptions to Principal Component Analysis
  F. From Matrices to Markov Chains

II. Introduction to Linear Algebra

- "Linear algebra is the branch of mathematics concerning vector spaces and linear mappings between such spaces." - Wikipedia
- "Linear algebra is where we use matrices to represent, manipulate, and solve systems of linear equations." - Matt Brems

A. Introduction to Matrices

Our goal is to use matrices to solve systems of linear equations. To begin, we'll need to define what exactly a matrix is and then dive into how to work with matrices to arrive at a solution.

matrix - a rectangular array with m rows, n columns, and $m \times n$ entries.

$$\mathbf{A}_{2\times 3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \{a_{ij}\}$$

$$\mathbf{B}_{2\times 3} = \begin{bmatrix} -2 & 4.1 & 0 \\ 3 & \pi & -7 \end{bmatrix}$$

Most commonly, these will be real-valued matrices, like above. *(Real-valued just means that every entry in the matrix is a real number.)*

Sometimes, however, we can represent random variables in this fashion. Suppose that $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$, and the covariance (unscaled correlation) of X_1 and X_2 is ρ. Then we can represent $\mathbf{X}_{2\times 1} = [X_1 \;\; X_2]^T$ as follows:

$$\mathbf{X}_{2\times 1} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1 & \rho \\ \rho & \sigma_2 \end{bmatrix}\right)$$

.

B. Matrix Operations

Here we detail several important operations one can perform on matrices. For each operation, we will show a specific example and a general example.

i. Matrix Addition

Let's add two matrices **A** and **B**. The matrices must be of the same dimensions. (That is, the number of rows and the number of columns must be the same.)

$$\begin{aligned}\mathbf{A}+\mathbf{B} &= \{a_{ij}\}+\{b_{ij}\} \\ &= \{a_{ij}+b_{ij}\}\end{aligned}$$

$$\begin{aligned}\mathbf{A}+\mathbf{B} &= \begin{bmatrix} 1 & 0 \\ -5 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3 \\ -2 & -4 \end{bmatrix}\end{aligned}$$

ii. Scalar Multiplication

Let's multiply a matrix **A** by some scalar (constant number) k.

$$\begin{aligned}k \times \mathbf{A} &= k \times \{a_{ij}\} \\ &= \{k \times a_{ij}\}\end{aligned}$$

$$\begin{aligned}k \times \mathbf{A} &= 2 \times \begin{bmatrix} 1 & 0 \\ -5 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ -10 & 0 \end{bmatrix}\end{aligned}$$

iii. Matrix Multiplication

Let's multiply two matrices **A** and **B**. The number of columns in matrix **A** must equal the number of rows in matrix **B**.

$$\begin{aligned}\mathbf{A}\times\mathbf{B} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix}\end{aligned}$$

$$\begin{aligned}\mathbf{A}\times\mathbf{B} &= \begin{bmatrix} 1 & 1 \\ -5 & 0 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \\ &= \begin{bmatrix} (1\times 2)+(1\times 3) & (1\times 3)+(1\times -4) \\ (-5\times 2)+(0\times 3) & (-5\times 3)+(0\times -4) \end{bmatrix} \\ &= \begin{bmatrix} 5 & -1 \\ -10 & -15 \end{bmatrix}\end{aligned}$$

iv. Matrix Transposition

Let's take the transpose of a matrix $\mathbf{A}$. We essentially "reflect" the matrix so that its columns become its rows and its rows become its columns.

$$\mathbf{A}^T = \{a_{ji}\}$$

$$\begin{aligned}\mathbf{A}^T &= \begin{bmatrix}1 & 0\\ -5 & 0\end{bmatrix}^T\\ &= \begin{bmatrix}1 & -5\\ 0 & 0\end{bmatrix}\end{aligned}$$

v. Inverse of Matrix

For <u>some</u> matrices, there exists an <u>inverse</u> matrix. (We'll get into matrices for which an inverse does and does not exist later.) Inverse matrices will probably always be calculated using a computer.

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix}1 & 1\\ -5 & 0\end{bmatrix}\\ \Rightarrow \mathbf{A}^{-1} &= \begin{bmatrix}0 & 0.2\\ -1 & 0.2\end{bmatrix}\end{aligned}$$

vi. Determinant of Matrix

For all square matrices, we can calculate the determinant of that matrix. This will probably always be done by computer.

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix}1 & 1\\ -5 & 0\end{bmatrix}\\ \Rightarrow \det(\mathbf{A}) &= |\mathbf{A}| = 5\end{aligned}$$

C. Classes of Matrices

There are many different classes of matrices that are important to know.

- * Row vector: $\mathbf{A}_{1\times n}$
- * Column vector: $\mathbf{A}_{m\times 1}$
- * Square matrix: $\mathbf{A}_{n\times n}$
- * Symmetric matrix: $\mathbf{A}^T = \mathbf{A}$
- Upper triangular matrix: $\{a_{ij}\} = 0$ for all $i > j$
- Lower triangular matrix: $\{a_{ij}\} = 0$ for all $i < j$
- Diagonal matrix: $\{a_{ij}\} = 0$ for all $i \neq j$
- Null matrix: $\{a_{ij}\} = 0$ for all i, j
- $\mathbf{J}$ matrix: $\{a_{ij}\} = 1$ for all i, j
- * Identity matrix ($\mathbf{I}_n$): $\{a_{ij}\} = 1$ for all $i = j$ and $\{a_{ij}\} = 0$ otherwise
- * $\mathbf{1}$: a vector with all entries equal to 1
- $\mathbf{0}$: a vector with all entries equal to 0
- * Orthogonal matrix: $\mathbf{A}^T = \mathbf{A}^{-1}$

D. Systems of Equations

Since linear algebra has to do with systems of linear equations, let's check out such a system. (Example pulled from https://en.wikipedia.org/wiki/System_of_linear_equations.)

$$\begin{aligned} 2x + 3y &= 6 \\ 4x + 9y &= 15 \end{aligned}$$

Our goal is to identify values of x and y that satisfy the above two equations simultaneously. There are three possible outcomes.

- There is no solution.
- There is exactly one solution.
- There are infinitely many solutions.

Generally, we could use either elimination or substitution to find the answer.

- Elimination: Subtract a multiple of one equation from the other in order to isolate one variable, then find the value of that variable, then plug in to solve for the other variable.
- Substitution: Within one equation, solve for one variable. (Say x.) In the other equation, replace x with the quantity you found earlier so that the equation is in terms of one variable. Simplify and solve for the remaining variable.

Think about a real-world scenario involving linear equations. Why might this be a problem?

E. Representing Equations as Matrices

$$\begin{bmatrix} 2 & 3 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix}$$

These are often represented as $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the matrix of coefficients, $\mathbf{x}$ is the vector of unknowns, and $\mathbf{b}$ is the vector of answers.

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{A}^{-1}\mathbf{Ax} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{Ix} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

Thus, if we are able to invert $\mathbf{A}$, then the vector of unknowns $\mathbf{x}$ will be equal to the vector given by $\mathbf{A}^{-1}\mathbf{b}$.

We can do this by either finding $\mathbf{A}^{-1}$ directly or by executing an algorithm called "Gaussian elimination" designed to take our matrix $\mathbf{A}$, append $\mathbf{b}$, and turn this into "row-reduced echelon form" which will ultimately yield the same answers. Gaussian elimination is the more preferred method, as inverting a matrix is very computationally expensive and relies on a few other conditions.