Math Review
Summer 2016

Topic 6

6. Matrices and linear algebra

6.1. Introduction to matrices and matrix operations

Most of you have had some basic linear algebra at this point. This should be smooth sailing. The goal is to give you the fundamental notions and instruments in linear algebra that may be helpful as you progress through the program. Matrices are very valuable for economists. Just think of how much we've used matrices already so far… when totally differentiating, for the Hessian, for the gradient, etc. The usefulness of matrices pans higher and further.

A matrix is a rectangular array of real numbers. It is usually written as:

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

A m x n matrix will have m rows and n columns. Recall that a m x 1 matrix is called a column vector. A 1 x n matrix is called a row vector. A matrix is generally denoted with a bold letter (*forgive me if some are not bolded by mistake- for every time someone finds a matrix letter not bolded, I will give out a candy*!)

You may sometime need to refer to an element in a matrix. For example, if you have the matrix given by:

$$\boldsymbol{A} = \begin{bmatrix} 4 & 9 \\ 5 & 8 \\ 8 & 13 \end{bmatrix}$$

Then, $a_{11} = 4$ and $a_{32} = 13$

Q: What is a_{21}, $a_{22,}$and a_{12}?

$\mathcal{A}$: 5, 8, and 9

A n x n matrix is called a square matrix of order n. There are some special types of square matrices.

A matrix where all the entries above the main diagonal are zero is called *lower triangular*:

$$\boldsymbol{L} = \begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

A matrix where all the entries above the main diagonal are zero is called *upper triangular*:

$$\boldsymbol{U} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

A *diagonal matrix* is one in which all the entries off the main diagonal are zero:

$$\boldsymbol{D} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

We also sometimes need to use the following ‘special’ types of matrices. They are the:

(1) **Zero matrix**, we denote by **0** the null matrix which contains zeros only.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \; ; \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}; [0], etc \ldots$$

$\mathcal{Q}$: What are these dimensions?

$\mathcal{A}$: 2 x 2 ; 3 x 1 ; 3 x 2 ; 1 x 1

(2) The other one is the **identity matrix.** The identity matrix is a matrix $I = I_n$ of size n x n whose elements are all units on the diagonal and zeroes on the other places.

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

One more matrix to recall:

(3) The matrix $\boldsymbol{A}$ is an idempotent matrix if it is equal to its square:

$$\boldsymbol{A}^2 = \boldsymbol{AA} = \boldsymbol{A}$$

6.1.1. Basic operations.

If A and B are both m x n matrices then $\boldsymbol{A} + \boldsymbol{B}$ is obtained by adding the corresponding elements of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. The matrix $\boldsymbol{A} - \boldsymbol{B}$ is obtained by subtracting the corresponding elements of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. In matrix notation we have:

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

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

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

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

Given the following matrices, find:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 8 \\ 0 & 7 \\ 8 & 5 \end{bmatrix} \;;\boldsymbol{B} = \begin{bmatrix} 11 & -2 \\ 4 & 9 \\ 0 & -1 \end{bmatrix};\quad \boldsymbol{C} = \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}\;;\ \boldsymbol{D} = \begin{bmatrix} 3 & 9 & 9 \\ 6 & 7 & 1 \end{bmatrix};\ \boldsymbol{E} = \begin{bmatrix} 0 & 2 & 0 \\ 10 & 9 & 3 \end{bmatrix}$$

Find:

(1) $\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} 13 & 6 \\ 4 & 16 \\ 8 & 4 \end{bmatrix}$

(2) $\boldsymbol{B} + \boldsymbol{C}$= undefined

(3) $\boldsymbol{B} - \boldsymbol{A} = \begin{bmatrix} 9 & -10 \\ 4 & 2 \\ -8 & -6 \end{bmatrix}$

(4) $\boldsymbol{C} - \boldsymbol{E}$= undefined

(5) $\boldsymbol{E} + \boldsymbol{D} = \begin{bmatrix} 3 & 11 & 9 \\ 16 & 16 & 4 \end{bmatrix}$

Other key properties that may be worth remembering:

(i) $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$
(ii) $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$
(iii) $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$, where $-\mathbf{A} = (\mathbf{1})\,\mathbf{A}$
(iv) $\mathbf{A} + \mathbf{0} = \mathbf{A}$
(v) $\mathbf{IA} = \mathbf{A}$

For $\lambda, \mu \in \mathbb{R}$

(vi) $\lambda\,(\mu\boldsymbol{A}) = (\mu\lambda\,)\boldsymbol{A}$
(vii) $(\lambda + \mu)\,\boldsymbol{A} = \lambda\boldsymbol{A} + \mu\boldsymbol{A}$
(viii) $\boldsymbol{\lambda}(\boldsymbol{A} + \boldsymbol{B}) = \lambda\boldsymbol{A} + \lambda\boldsymbol{B}$

Multiplication by a scalar:

If $\boldsymbol{A}$ is a m x n matrix and c is any scalar (i.e. a number), $c\boldsymbol{A}$ is obtained by multiplying each element of **A** by c. In matrix notation:

$$c\boldsymbol{A} = c.\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} c \text{ x } a_{11} & c \text{ x } a_{12} & \cdots & c \text{ x } a_{1n} \\ c \text{ x } a_{21} & c \text{ x } a_{22} & \cdots & c \text{ x } a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c \text{ x } a_{m1} & c \text{ x } a_{m2} & \cdots & c \text{ x } a_{mn} \end{bmatrix}$$

Matrix Multiplication:

If $\boldsymbol{A}$ is a m x p matrix and $\boldsymbol{B}$ is a p x n matrix, then $\boldsymbol{AB}$ is then m x n matrix given by:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \text{ x } \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{p} a_{1i}b_{i1} & \sum_{i=1}^{p} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{p} a_{1i}b_{in} \\ \sum_{i=1}^{p} a_{2i}b_{i1} & \boxed{\sum_{i=1}^{p} a_{2i}b_{i2}} & \cdots & \sum_{i=1}^{p} a_{2i}b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{p} a_{mi}b_{i1} & \sum_{i=1}^{p} a_{mi}b_{i2} & \cdots & \sum_{i=1}^{p} a_{mi}b_{in} \end{bmatrix}$$

For example, as highlighted:

$$\sum_{i=1}^{n} a_{2i}b_{i2} = a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2p}b_{p2}$$

Recall that we need that the 'inside dimension' of the matrices being multiplied to *equate* or *match*. That is we can multiply $\boldsymbol{AB}$ only if we have the form $\boldsymbol{A_{m\,x\,p}}$ **and** $\boldsymbol{B_{p\,x\,n}}$. Denote the subscript to be the dimension of the matrices here.

Q: Find $\boldsymbol{AB}$, where $\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 3 & 6 & 4 \\ 2 & 5 & 8 \\ 7 & 1 & 9 \end{bmatrix}$

$\mathcal{A}$: $\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 6 & 4 \\ 2 & 5 & 8 \\ 7 & 1 & 9 \end{bmatrix} = \begin{bmatrix} 16 & 7 & 26 \\ 43 & 22 & 61 \end{bmatrix}$

Quick Q: Can you also find $\boldsymbol{BA}$?

In general, it is safe to assume that $\boldsymbol{AB} \neq \boldsymbol{BA}$

Q: Given the dimension below, which matrices can you multiply? Write out the matrices with their dimensions.

$$\boldsymbol{A}_{3x4},\ \ \boldsymbol{B}_{4x6},\ \ \boldsymbol{C}_{6x3}$$

$\mathcal{A}$: $\boldsymbol{AB}_{3x6}$; $\boldsymbol{CA}_{6x4}$; $\boldsymbol{BC}_{4x3}$.

Try at home:

Now that you remember your basic operations, which of the following matrices are idempotent?

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

Ans: $\boldsymbol{A}$ and $\boldsymbol{C}$

Let's remember two more properties of matrices known as the:

(1) Distributive property: $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$ and $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$
(2) Associative Addition: $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$.

6.2. Transpose and trace

Definition. The transpose of matrix the n x m matrix $\boldsymbol{A}$ is the m x n matrix that results from interchanging the rows and columns of $\boldsymbol{A}$:

Formally, for $\boldsymbol{A} = [a_{ij}]$ the tranpose is $\boldsymbol{A}' = \boldsymbol{A}^{\boldsymbol{T}} = [a_{ji}]$.

Consider the following matrices, what are the respective transpose?

$$\boldsymbol{A} = \begin{bmatrix} 2 & 7 \\ 3 & 0 \end{bmatrix} \qquad \boldsymbol{A}' = \boldsymbol{A}^{\boldsymbol{T}} = \begin{bmatrix} 2 & 3 \\ 7 & 0 \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \qquad \boldsymbol{B}' = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \boldsymbol{C}' = [x_1 \quad x_2 \quad x_3]$$

Properties of transposed matrices:

(i) $\boldsymbol{A}'(\boldsymbol{A}') = \boldsymbol{A}$
(ii) $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$
(iii) $(c\boldsymbol{A})' = c(\boldsymbol{A}')$, where $c \in \mathbb{R}$
(iv) $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$

Definition. A square matrix $\boldsymbol{A}$ is symmetric if $\boldsymbol{A}' = \boldsymbol{A}$. For example, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Definition. The trace of a square matrix $\boldsymbol{A}$ is the sum of the entries on the main diagonal of $\boldsymbol{A}$, say given by $\sum a_{ii}$. If $\boldsymbol{A}$ is not a square matrix then the trace of $\boldsymbol{A}$ is undefined.

Q:For example, looking at the definition, what is the trace of this matrix?

$$Tr \begin{bmatrix} 1 & 2 & 3 \\ 10 & 20 & 30 \\ 100 & 200 & 300 \end{bmatrix}$$

$\mathcal{A}$: 321

6.3. Determinants, inverses and related properties

6.3.1 Determinants

When we covered concavity in n-space, we touched on determinants. We will review quickly if you think it is helpful and move forward.

For a 2 x 2 matrix: $det\ \boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$

Q: Use the definition to find the determinant of the following:

$$\boldsymbol{C} = \begin{bmatrix} 4 & 6 \\ 3 & 8 \end{bmatrix}$$

$\mathcal{A}$:

$$\begin{aligned} |\boldsymbol{C}| &= 4 \times 8 - 6 \times 3 \\ &= 32 - 18 \\ &= 14 \end{aligned}$$

For a 3 x 3 matrix: $\det \boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}2_{32} + a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

Q: Let $\boldsymbol{A}$=$\begin{bmatrix} 6 & 1 & 1 \\ 4 & -2 & 5 \\ 2 & 8 & 7 \end{bmatrix}$. Find the determinant of this matrix $\boldsymbol{A}$.

$\mathcal{A}$:

$$\begin{aligned} |\boldsymbol{A}| &= 6 \times (-2 \times 7 - 5 \times 8) - 1 \times (4 \times 7 - 5 \times 2) + 1 \times (4 \times 8 - -2 \times 2) \\ &= 6 \times (-54) - 1 \times (18) + 1 \times (36) \\ &= -306 \end{aligned}$$

Try at home:

Find the determinant of the following matrices:

$$\boldsymbol{A} = \begin{bmatrix} 3 & 5 \\ 1 & 14 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} -7 & 5 \\ -14 & 4 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 3 & -7 \\ 1 & -14 \end{bmatrix}$$

$$det\ \boldsymbol{A} = 7, det\ \boldsymbol{B} = 42, det\ \boldsymbol{C} = \boldsymbol{-35}$$

$$\boldsymbol{D} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 1 \\ 0 & 3 & -1 \end{bmatrix},\ \boldsymbol{E} = \begin{bmatrix} 5 & -2 & 1 \\ 0 & 3 & -1 \\ 2 & 0 & 7 \end{bmatrix}$$

$$det\ \boldsymbol{D} = 1, \det \boldsymbol{E} = 103$$

Like I mentioned before, I think is it quasi impossible that you will be asked to compute the determinants for matrices bigger than 3 x 3. If you are, probably a TA is feeling like getting you thinking about it a little more. I doubt you'll see it on an exam, ever. Now if it comes up, … uhm, good luck?

Properties of determinants:

Note that we will not be proving any of these. Proofs are easily available is so desired.

Let A and B be square n x n matrices. Then:

(i) If $\boldsymbol{A}$ has rows and columns of 0's then, the $det\ \boldsymbol{A} = \boldsymbol{0}$

(ii) $det\ \boldsymbol{A} = det\ \boldsymbol{A}'$

(iii) If $\boldsymbol{A}$ is upper, lower triangular or a diagonal matrix, then, $det\ \boldsymbol{A} = \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{a_{ii}} = a_{11}a_{22}a_{23} \dots a_{nn}$.

(iv) Let $\boldsymbol{B}$ be the square matrix obtained from $\boldsymbol{A}$ by multiplying a single row by the scalar α, or by multiplying a single column by the scalar α. Then $\det\ \boldsymbol{B} = \alpha \det\ \boldsymbol{A}$.

(v) If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by interchanging two rows (or two columns), then $det\ \boldsymbol{B} = -\ det\ \boldsymbol{A}$.

(vi) If two rows (or two columns) of $\boldsymbol{A}$ are identical, then $\det \boldsymbol{A} = 0$

(vii) If $\boldsymbol{B}$ is obtained by adding a constant multiple of one row (or column) of $\boldsymbol{A}$ to another row (or column) of $\boldsymbol{A}$, then $det\ \boldsymbol{B} = det\ \boldsymbol{A}$.

(viii) $det\ \boldsymbol{AB} = det\ \boldsymbol{A}\ det\ \boldsymbol{B}$.

6.3.2. Inverses

Intuition: What is the reciprocal of a number, say 8? It is really $\frac{1}{8}$.
The inverse of a matrix is the same essence but we write it out as Inverse of $\boldsymbol{A}$ given by $\boldsymbol{A}^{-1}$. Remember, we cannot divide by a matrix! Thus, we do not use, for instance, $\frac{\mathbf{1}}{\boldsymbol{A}}$.

Definition. If $\boldsymbol{A}$ is a square n x n matrix, and if there exists an n x n matrix $\boldsymbol{B}$ such that:

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I},$$

then $\boldsymbol{A}$ is invertible and $\boldsymbol{B}$ is the inverse of $\boldsymbol{A}$. Matrix $\boldsymbol{B}$ can also be denoted $\boldsymbol{A}^{-\mathbf{1}}$.
If matrix B does not exist then A is singular.

(Uniqueness of Inverse) Theorem. If both $\boldsymbol{B}$ and $\boldsymbol{C}$ are inverses of the matrix $\boldsymbol{A}$, then $\boldsymbol{B} = \boldsymbol{C}$.

Properties of Inverses

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are n x n matrix then:

(i) $\boldsymbol{A}^{-1}$ is invertible and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$
(ii) For $c \in \mathbb{N}$ the matrix $\boldsymbol{A}^{c}$ is invertible and $(\boldsymbol{A}^{c})^{-1} = (\boldsymbol{A}^{-1})^{c}$
(iii) For any nonzero scalar c, the matrix $c\boldsymbol{A}$ is invertible and $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$
(iv) $\boldsymbol{AB}$ is invertible and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
(v) $\boldsymbol{A}^{-\mathbf{1}}$ is invertible and $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$

Before jumping into finding inverses, I wanted to make a quick note. Matrices have a lot of value. Inverses and everything we are learning do too. This is an example that always helps me with inverses. Imagine you have someone ask you, "how do I share these 10 apples between 2 people?". You say, well,
$\frac{10}{2} = 10 \text{ x } 2^{-1} = 5$. Each gets 5!

In matrices, since you cannot divide, if you have to find an alternate way and inverses come to rescue.
If you have $\boldsymbol{AX} = \boldsymbol{B}$. It may be nice to divide $\boldsymbol{X} = \frac{\boldsymbol{B}}{\boldsymbol{A}}$, but once again **we cannot**.

Thus, think of this:

$$\boldsymbol{A}^{-\mathbf{1}}\boldsymbol{AX} = \boldsymbol{A}^{-\mathbf{1}}\boldsymbol{B}$$
$$\boldsymbol{IX} = \boldsymbol{A}^{-\mathbf{1}}\boldsymbol{B}$$

$$\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B}$$

This is why we care for the inverse.

Finding inverses

If the 2 x 2 square matrix given by $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then the inverse given by:

$$\boldsymbol{A}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Basically what we did here is swap the positions of a and d, put negatives in front of b and c, and divide everything by the determinant (ad-bc).

Q:Find the inverse of matrix $\boldsymbol{D} = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}$

$\mathcal{A}$: $\boldsymbol{D}^{-1} = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}^{-1} = \frac{1}{4x6-7x2}\begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix}$

$$= \frac{1}{10}\begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.6 & -0.7 \\ -0.2 & -0.4 \end{bmatrix}$$

Recall one of the properties of inverses, that is $\boldsymbol{A}$ x $\boldsymbol{A}^{-1} = I$.

Q: Is that true for the case above? Can you check?

$\mathcal{A}$: $\begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} 0.6 & -0.7 \\ -0.2 & -0.4 \end{bmatrix}$

$$= \begin{bmatrix} 2.4 - 1.4 & -2.8 + 2.8 \\ 1.2 - 1.2 & -1.4 + 2.4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Inverses may not always exist! This happens with any instance where the determinant is equal to 0 (and we cannot divide by 0). Think of a matrix given by: $\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$. This is called a singular matrix (*look back at definition of an inverse*).

<u>Definition</u>. An n x n matrix $\boldsymbol{E}$ is called an **elementary matrix** if it can be obtained from the n x n identity matrix $\boldsymbol{I_n}$ by performing a single elementary row operation. An **elementary row operation** is:

(i) Multiplying a row through by a nonzero constant,
(ii) Interchanging two rows, or
(iii) Adding a multiple of one row to another row.

In other words, to solve for $\boldsymbol{A^{-1}}$, we follow the following steps:

1) Find a sequence of elementary row operations that reduce $\boldsymbol{A}$ to the identity matrix.
2) Perform the same sequence of operations on $\boldsymbol{I_n}$to obtain $\boldsymbol{A}.$

I'll work through an example and you can do the next.

Example. Let $\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$, find $\boldsymbol{B^{-1}}$.

To solve for $\boldsymbol{B^{-1}}$, remember that we want to reduce the given $\boldsymbol{B}$, to the 3 x 3 identity matrix using row operations while doing these operations simultaneously to I, to produce $\boldsymbol{B^{-1}}$.

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 4 & 5 & 0 & 1 & 0 \\ 3 & 5 & 6 & 0 & 0 & 1 \end{array}\right]$$ Row Operation:

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 4 & 5 & 0 & 1 & 0 \\ 0 & -1 & -3 & -3 & 0 & 1 \end{array}\right]$$ -3 x 1st row added to 3rd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 0 & -1 & -2 & 1 & 0 \\ 0 & -1 & -3 & -3 & 0 & 1 \end{array}\right]$$ -2 x 1st row added to 2nd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -3 & -3 & 0 & 1 \\ 0 & 0 & -1 & -2 & 1 & 0 \end{array}\right]$$

Interchange 2nd and 3rd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 0 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

Multiply 2nd and 3rd row with -1

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

-3 x 3rd row added to 2nd row

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 7 & -6 & 2 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

-2 x 2nd row added to 1st row

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -3 & 2 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

-3 x 3rd row added to 1st row

Thus, $\boldsymbol{B}^{-1} = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 3 & -1 \\ 2 & -1 & 0 \end{bmatrix}$

Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$, find $\boldsymbol{A}^{-1}$.

To solve for $\boldsymbol{A}^{-1}$, remember that we want to reduce the given $\boldsymbol{A}$, to the 3 x 3 identity matrix using row operations while doing these operations simultaneously to I, to produce $\boldsymbol{A}^{-1}$.

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$ Row Operation:

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$ -2 x 1st row added to 2nd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right]$$ -1 x 1st row added to 3rd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right]$$ 2 x 2nd row added to 3rd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$ 3rd row times -1

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$ 3 x 3rd row added to 2nd row

$$= \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$ -3 x 3rd row added to 1st row

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$ -2 times 2nd row added to 1st

We have the Identity matrix I on the left hand side and the inverse A^{-1} on the right. If you cannot transform A to the identity matrix using elementary row operations, then A is not invertible. Recall, a square matrix $\boldsymbol{A}$ is invertible if and only if $\det A \neq 0$.

Thus: $A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$

To try at home:

Find the inverse of the following matrices:

$$\boldsymbol{E} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}$$

$$Ans: \boldsymbol{E}^{-1} = \begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix}$$

$$\boldsymbol{G} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$Ans: \boldsymbol{G}^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

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

$$\boldsymbol{K^{-1}} = \frac{1}{11}\begin{bmatrix} 12 & -6 & -1 \\ \frac{5}{2} & \frac{3}{2} & -\frac{5}{2} \\ -2 & 1 & 2 \end{bmatrix}$$

$$\boldsymbol{L} = \begin{bmatrix} 7 & 2 & 1 \\ 0 & 3 & -1 \\ -3 & 4 & -2 \end{bmatrix}$$

$$\boldsymbol{L^{-1}} = \begin{bmatrix} -2 & 8 & -5 \\ 3 & -11 & 7 \\ 9 & -34 & 21 \end{bmatrix}$$

6.4. Systems of linear equations, independence and rank

6.4.1. System of linear equations

You will find system of linear equations widely used in economics. To begin, consider a simple arbitrary system of two linear equations given by:

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

From the above, we can let $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. This can be denoted as the matrix of *known* coefficients.

The vector of unknown variables is given as $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

and the vector of constants is $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. In matrix form, this is the familiar **Ax=b**

Suppose you have a system of m linear equations and n unknowns, it can be written as:

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_3 \end{gathered}$$

Here, $x_1, x_2, \dots, x_n$ are the unknowns, and a_{ji} and b_i are constants. This system can be in the following form:

$$\boldsymbol{Ax} = \boldsymbol{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

The solution to this system are values for x_1 through x_n that make the above equality true. The augmented matrix for this system is:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right]$$

Our goal is often to solve this system of linear equations and the most basic method is the **Guassian elimination**.

Gauss-Jordan Elimination

To solve a system of linear equations we can reduce the systems associated augmented matrix to reduced row-echelon form:

Definition. A matrix is in *reduced row-echelon* form if it satisfies the following four properties:

(i) If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a leading 1.
(ii) If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
(iii) If any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right then the leading 1 in the higher row.
(iv) Each column that contains a leading 1 has zeros everywhere else.

$\mathcal{Q}$: Which of these are in reduced row echelon form?

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{E} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\mathcal{A}$: A – yes, B – yes, C- no, D – yes, E – no

To solve a system of linear equations using Gauss-Jordan elimination we apply the following algorithm to the augmented matrix, $\boldsymbol{A}$:

1. Write the augmented matrix of the system.

2. Use row operations to transform the augmented matrix in the form described below, which is called the *reduced row echelon form (RREF).*
   (a) The rows (if any) consisting entirely of zeros are grouped together at the bottom of the matrix.
   (b) In each row that does not consist entirely of zeros, the leftmost nonzero element is a 1 (called a leading 1 or a pivot).
   (c) Each column that contains a leading 1 has zeros in all other entries.
   (d) The leading 1 in any row is to the left of any leading 1's in the rows below it.

3. Stop process in step 2 if you obtain a row whose elements are all zeros except the last one on the right. In that case, the system is inconsistent and has no solutions. Otherwise, finish step 2 and read the solutions of the system from the final matrix.

Note: When doing step 2, row operations can be performed in any order. Try to choose row operations so that as few fractions as possible are carried through the computation. This makes calculation easier when working by hand.

Example. Solve the following system by Gauss –Jordan elimination.

$$x_1 - 2x_2 + 3x_3 + 2x_4 + x_5 = 10$$
$$2x_1 - 4x_2 + 8x_3 + 3x_4 + 10x_5 = 7$$
$$3x_1 - 6x_2 + 10x_3 + 6x_4 + 5x_5 = 27$$

The augmented matrix for this system is given by:

$$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 2 & -4 & 8 & 3 & 10 & 7 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{array}\right]$$

We want to follow the above steps to transform this augmented matrix into reduced row echelon form.

$$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{array}\right]$$ -2 x 1st row added to 2nd row

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right]$ -3 x 1st row added to 3rd row

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 1 & -\frac{1}{2} & 4 & -\frac{13}{2} \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right]$ x 2nd row by ½

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 1 & -\frac{1}{2} & 4 & -\frac{13}{2} \\ 0 & 0 & 0 & \frac{1}{2} & -2 & \frac{7}{2} \end{array}\right]$ -1 x 2nd row added to 3rd row

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 1 & -\frac{1}{2} & 4 & -\frac{13}{2} \\ 0 & 0 & 0 & 1 & -4 & 7 \end{array}\right]$ x 3rd row by 2

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 1 & -4 & 7 \end{array}\right]$ ½ x 3rd row added to 2nd row

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 0 & 9 & -4 \\ 0 & 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 1 & -4 & 7 \end{array}\right]$ -2 x 3rd row added to 1st row

$\left[\begin{array}{ccccc|c} 1 & -2 & 3 & 0 & 9 & -4 \\ 0 & 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 1 & -4 & 7 \end{array}\right]$ -2 x 3rd row added to 1st row

$\left[\begin{array}{ccccc|c} 1 & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 1 & -4 & 7 \end{array}\right]$ -3 x 2nd row added to 1st row

Note that this meets all of the requirements above. The system has been reduced to:

$x_1 - 2x_2 + 3x_5 = 5$
$x_3 + 2x_5 = -3$
$x_4 - 4x_5 = 7$

Since there is no specific value for x_2 and x_5, those can be chosen arbitrarily. This means that there are *infinitely many solutions* for this system. We can represent all the solutions by using parameter r, s as follows.

We call this assigning 'free' variables, say $x_2 = r$ and $x_5 = s$, then the solution is given by:

$$x_1 = 5 + 2r - 3s, x_2 = r, x_3 = -3 - 2s, x_4 = 7 + 4s \text{ and } x_5 = s$$

Any values of the parameter r, s gives us a solution of the system.

Every augmented matrix $\boldsymbol{A}$ has a unique reduced row-echelon form.

Q: Try this example (*slightly simpler actually – maybe, I don't know, you tell me ☺!*)

Solve the following system by using the Gauss-Jordan elimination method.

$$\begin{aligned} 4y + z &= 2 \\ 2x + 6y - 2z &= 3 \\ 4x + 8y - 5z &= 4 \end{aligned}$$

This can be represented in an augmented matrix as follows:

$$\left[\begin{array}{ccc|c} 0 & 4 & 1 & 2 \\ 2 & 6 & -2 & 3 \\ 4 & 8 & -5 & 4 \end{array}\right]$$

We want to follow the above steps to transform this augmented matrix into reduced row echelon form.

$$\left[\begin{array}{ccc|c} 2 & 6 & -2 & 3 \\ 0 & 4 & 1 & 2 \\ 4 & 8 & -5 & 4 \end{array}\right]$$ Switch 1st row with 2nd row

$$\left[\begin{array}{ccc|c} 2 & 6 & -2 & 3 \\ 0 & 4 & 1 & 2 \\ 0 & -4 & -1 & -2 \end{array}\right]$$ -2 x 1st row added to 3rd row

$$\left[\begin{array}{ccc|c} 2 & 6 & -2 & 3 \\ 0 & 4 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$ 2nd row added to 3rd row

$$\left[\begin{array}{ccc|c} 2 & 6 & -2 & 3 \\ 0 & 1 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$ ¼ 2^{nd} row

$$\left[\begin{array}{ccc|c} 2 & 0 & -\frac{7}{2} & 0 \\ 0 & 1 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$ -6 x 2^{nd} row added to 1^{st} row

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{7}{4} & 0 \\ 0 & 1 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$ ½ of 1^{st} row

We have gone through all the necessary operations. We can express the solutions as:

$$x - \frac{7}{4}z = 0 \to x = \frac{7}{4}z$$
$$y + \frac{1}{4}z = \frac{1}{2}$$

Since there is no specific value for z, it can be chosen arbitrarily. This means that there *are infinitely many solutions* for this system. We can represent all the solutions by using a parameter t as follows. Any value of the parameter t gives us a solution of the system.

To try at home:

Solve the following systems by using the Gauss-Jordan elimination method.

$$\begin{cases} x+y+z=5 \\ 2x+3y+5z=8 \\ 4x+5z=2 \end{cases}$$

$$Ans: x = 3, y = 4, z = -2$$

$$\begin{cases} x+2y-3z=2 \\ 6x+3y-9z=6 \\ 7x+14y-21z=13 \end{cases}$$

$$Ans: No\ solutions, inconsistent\ system$$

$$\begin{cases} A+B+2C=1 \\ 2A-B+D=-2 \\ A-B-C-2D=4 \\ 2A-B+2C-D=0 \end{cases}$$

$$Ans: A = 1, B = 2, C = -1, D = -2$$

6.4.2. Cramer's Rule

Now, there is a second way that we can solve a system of linear equations where $\boldsymbol{A}$ is a square matrix is via Cramer's Rule:

(Cramer's Rule) Theorem. If $\boldsymbol{Ax} = \boldsymbol{b}$ is a system of m linear equations in n unknowns such that $det\ \boldsymbol{A} \neq 0$ then the system has a unique solution. The solution is:

$$x_1 = \frac{det\boldsymbol{A}_1}{det\boldsymbol{A}}; x_2 = \frac{det\boldsymbol{A}_2}{det\boldsymbol{A}}, x_3 = \frac{det\boldsymbol{A_3}}{det\boldsymbol{A}}, \dots, x_n = \frac{det\boldsymbol{A_n}}{det\boldsymbol{A}}$$

where $\boldsymbol{A}_j$ is the matrix obtained by replacing the entries in the jth column of $\boldsymbol{A}$ by $\boldsymbol{b}$

Example. Solve this system using Cramer's rule:

$$7x_1 - 2x_2 = 3$$
$$3x_1 + x_2 = 5$$

We denote: coefficient matrix $\boldsymbol{A} = \begin{bmatrix} 7 & -2 \\ 3 & 1 \end{bmatrix}$; $\boldsymbol{A}_{x_1} = \begin{bmatrix} 3 & -2 \\ 5 & 1 \end{bmatrix}$; $\boldsymbol{A}_{x_2} = \begin{bmatrix} 7 & 3 \\ 3 & 5 \end{bmatrix}$

$$\det \boldsymbol{A} = (7)(1) - (3)(-2) = 7 + 6 = 13$$
$$\det \boldsymbol{A}_{x_1} = (3)(1) - (5)(-2) = 3 + 10 = 13$$
$$\det \boldsymbol{A}_{x_2} = (7)(5) - (3)(3) = 35 - 9 = 26$$

$$x_1 = \frac{det\boldsymbol{A}_{x_1}}{det\boldsymbol{A}} = \frac{13}{13} = 1$$
$$x_2 = \frac{det\boldsymbol{A}_{x_2}}{det\boldsymbol{A}} = \frac{26}{13} = \frac{1}{2}$$

$\mathcal{Q}$: Use Cramer's Rule to solve this system:

$$3x - y = 7$$
$$-5x + 4y = -2$$

$\mathcal{A}$:

We denote: $\boldsymbol{B} = \begin{bmatrix} 3 & -1 \\ -5 & 4 \end{bmatrix}$; $\boldsymbol{B}_x = \begin{bmatrix} 7 & -1 \\ -2 & 4 \end{bmatrix}$; $\boldsymbol{B}_y = \begin{bmatrix} 3 & 7 \\ -5 & -2 \end{bmatrix}$

$$\det \boldsymbol{B} = (3)(4) - (-1)(-5) = 12 - 5 = 7$$
$$\det \boldsymbol{B}_x = (7)(4) - (-1)(-2) = 28 - 2 = 26$$
$$\det \boldsymbol{B}_y = (3)(-2) - (7)(-5) = -6 + 35 = 29$$

$$x = \frac{det\boldsymbol{B}_x}{det\boldsymbol{B}} = \frac{26}{7}$$
$$y = \frac{det\boldsymbol{B}_y}{det\boldsymbol{B}} = \frac{29}{7} = \frac{1}{2}$$

I will do an example of Cramer's rule with a 3 x 3 matrix. The steps are very similar, so if you want, you can try a couple at home as well. I think it might be a little repetitive to have you guys do a 3 x 3 matrix exercise in class today though.

Example. Consider the following system of equations. Use Cramer's rule to solve for x, y and z.

$$2x + y + z = 3$$
$$x - y - z = 0$$
$$x + 2y + z = 0$$

The coefficient matrix is given by:

$$Let\ \boldsymbol{F} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{bmatrix}$$

Denote, $\boldsymbol{F_x} = \begin{bmatrix} 3 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 1 \end{bmatrix}$, $\boldsymbol{F_y} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix}$, $\boldsymbol{F_z} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 1 & 2 & 0 \end{bmatrix}$

$|\boldsymbol{F}| = 3,\ |\boldsymbol{F_x}| = 3, |\boldsymbol{F_y}| = -6, |\boldsymbol{F_z}| = 9$

Thus

$$x = \frac{|\boldsymbol{F_x}|}{|F|} = \frac{3}{3} = 1$$
$$y = \frac{|\boldsymbol{F_y}|}{|F|} = \frac{-6}{3} = -2$$
$$z = \frac{|\boldsymbol{F_z}|}{|F|} = \frac{9}{3} = 3$$

That's really all there is to Cramer's Rule. It is very simple and straightforward to use.

Try at home:

Use Cramer's Rule to solve these system:

$$\begin{cases} 4x - 3y = 11 \\ 6x + 5y = 7 \end{cases}$$

$$Ans: x = 2, y = -1$$

$$\begin{cases} 3x + 5y = -7 \\ x + 4y = -14 \end{cases}$$

$$Ans: x = 6, y = -5$$

Solve for z using Cramer's Rule:

$$\begin{cases} 2x + y + z = 1 \\ x - y + 4z = 0 \\ x + 2y - 2z = 3 \end{cases}$$

$$Ans: z = 2$$

Every system of linear equations has either

(a) a unique solution,
(b) no solution, or
(c) infinitely many solutions.

If a system of equations has *no solutions it is said to be inconsistent*, and if a system of equations has *at least one solution it is said to be consistent.* A very important type of linear systems is that of a *homogenous linear system*:

Definition. A system of linear equations is said to be homogenous if the constant terms are all zero. That is $b_i = 0$ for all i and the *augmented matrix* is of the form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 \end{bmatrix}$$

As you can deduce from looking at the matrix, every homogenous system has the trivial solution $x_1 = x_2 = \cdots = x_n = 0$. A nontrivial solution is any other solution.

Theorem. A homogenous system of linear equations with more unknowns than equations has infinitely many solutions.

6.4.3 Linear Dependence/Independence

Definition: If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ is a nonempty set of vectors, then the vector equation:

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$$

has at least one solution, namely $c_1 = c_2 = \cdots = c_n = 0$. If this is the only solution, then S is called a linearly independent set. If there are other solutions, then S is called a linearly dependent set. We say that the column vectors of the $m \times n$ matrix A are linearly independent if the n column vectors are linearly independent. The row vectors of A are linearly independent if the m row vectors are linearly independent.

We have the following vectors:

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$$

Q: Consider the definition, are these vectors linearly dependent or linearly independent?
Hint: Try to check if you can express these vectors in an equation that leads to $c_1x_1 + c_2x_2 + c_3x_3 = 0$, *where* $c_1, c_2, c_3 \neq 0$.

$\mathcal{A}$: Since we can write these vectors as $2x_1 + x_2 - x_3 = 0$, they are linearly dependent.

6.4.5. Rank of a matrix

Definition. The rank of matrix $\boldsymbol{A}$ is the equivalent to the number of non-zero rows when $\boldsymbol{A}$ is in its reduced row-echelon form. If the reduced row-echelon form has no zeros, then we say that **matrix $\boldsymbol{A}$ has full rank**.

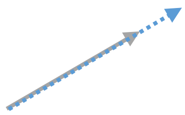
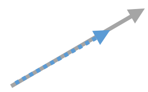
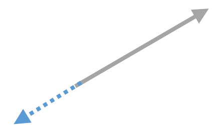
Theorem: If $\boldsymbol{A}$ is an $n \times n$ matrix, then the following are equivalent:

(a) $\boldsymbol{A}$ is invertible. (i.e. $\boldsymbol{A}$ is nonsingular.)
(b) $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution.
(c) The reduced row-echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}_n$.
(e) $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for every $n \times 1$ matrix $\boldsymbol{b}$. (i.e. The system has at least one solution.)
(f) $\boldsymbol{Ax} = \boldsymbol{b}$ has exactly one solution for every $n \times 1$ matrix $\boldsymbol{b}$.
(g) $det\ \boldsymbol{A} \neq 0$.
(h) The column vectors of $\boldsymbol{A}$ are linearly independent.
(i) The row vectors of $\boldsymbol{A}$ are linearly independent.
(j) $\boldsymbol{A}$ has rank n. (i.e. $\boldsymbol{A}$ has full rank)

6.6. Eigenvalues, eigenvectors and Markov processes

Let $\boldsymbol{A}$ be an $n \times n$ matrix. The number λ is an eigenvalue of $\boldsymbol{A}$ if there exists a non-zero vector v such that $\boldsymbol{Av} = \lambda\boldsymbol{v}$. In this case, vector $\boldsymbol{v}$ is called an eigenvector of $\boldsymbol{A}$ corresponding to λ.

One intuitive way to think about it (*I hope this helps – it helps me!*)

 $\boldsymbol{Av} = \lambda\boldsymbol{v}, \lambda > 1$	 $\boldsymbol{Av} = \lambda\boldsymbol{v}, 0 < \lambda < 1$	 $\boldsymbol{Av} = \lambda\boldsymbol{v}, \lambda < 0$

<u>Definition</u>. The characteristic equation of the $n \times n$ matrix $\boldsymbol{A}$ is:

$$\det(\lambda \boldsymbol{I_n} - \boldsymbol{A}) = 0$$

$\lambda \in \mathbb{R}$. The eigenvalue λ solves the characteristic equation given above. The eigenvectors are the nontrivial solutions to:

$$(\lambda \boldsymbol{I_n} - \boldsymbol{A})\boldsymbol{x} = 0$$

where λ is an eigenvector.

An example can clear things up as eigenvalues and eigenvectors are usually confusing for many.[1]

Example. What is the characteristic equation, eigenvectors, and eigenvalues for the following linear system:

$$\begin{aligned} 2x_1 + 3x_2 &= \lambda x_1 \\ 4x_1 + 3x_2 &= \lambda x_2. \end{aligned}$$

The characteristic equation for this system is:

$$\det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}\right) = \det\begin{bmatrix} \lambda - 2 & -3 \\ -4 & \lambda - 3 \end{bmatrix} = 0$$

Then, we have: $\lambda^2 - 5\lambda - 6 = 0$, $\lambda_1 = 1$ and $\lambda_2 = 6$.

From the definition, the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is an eigenvector if and only if it is the nontrivial solution to:

[1] *I would refer you here for a deeper explanation (with visuals) of the mechanics of eigenvalues*: http://setosa.io/ev/eigenvectors-and-eigenvalues/

$\begin{bmatrix} \lambda - 2 & -3 \\ -4 & \lambda - 3 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

For $\lambda_1 = 1$: $\begin{bmatrix} -3 & -3 \\ -4 & -4 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Do a quick augmented matrix

$\left[\begin{array}{cc|c} -3 & -3 & 0 \\ -4 & -4 & 0 \end{array}\right]$

$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right]$ -1/3 x 1st row AND -1/4 x 2nd row

$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$ -1 x 2nd row added to 2nd row

Thus $x + y = 0, or\ x = -y$, more generally, $x = t, y = -t$

For $\lambda_1 = 6$: $\begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

We ca go through the same steps and find that $y_1 = \frac{3}{4}t, y_2 = t$

Thus, eigenvectors $\boldsymbol{x} = \begin{bmatrix} t \\ -t \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} \frac{3}{4}t \\ t \end{bmatrix}$.

Q: What is the characteristic equation, eigenvectors, and eigenvalues for the following linear system:

$$2x_1 + 2x_2 = \lambda x_1$$
$$5x_1 - x_2 = \lambda x_2.$$

$\mathcal{A}$: The characteristic equation for this system is:

$$\det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix}\right) = \det\begin{bmatrix} \lambda - 2 & -2 \\ -5 & \lambda + 1 \end{bmatrix} = 0$$

Then, we have$\lambda^2 - \lambda - 12$, $\lambda_1 = -3$ and $\lambda_2 = 4$.

From the definition, the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is an eigenvector if and only if it is the nontrivial solution to:

$\begin{bmatrix} \lambda - 2 & -2 \\ -5 & \lambda + 1 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

For $\lambda_1 = -3$: $\begin{bmatrix} -5 & -2 \\ -5 & -2 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Do a quick augmented matrix

$\left[\begin{array}{cc|c} -5 & -2 & 0 \\ -5 & -2 & 0 \end{array}\right]$

$\left[\begin{array}{cc|c} 5 & 2 & 0 \\ 5 & 2 & 0 \end{array}\right]$ -1 x 1st row and 2nd

$\left[\begin{array}{cc|c} 5 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$ -1 x 2nd row added with 1st row

Thus $5x + 2y = 0, or\ x = -\frac{2}{5}y$, more generally, $x = t, y = -\frac{2}{5}t$

For $\lambda_1 = 4$: $\begin{bmatrix} 2 & -2 \\ -5 & 5 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Do a quick augmented matrix

$\left[\begin{array}{cc|c} 2 & -2 & 0 \\ -5 & 5 & 0 \end{array}\right]$

$\left[\begin{array}{cc|c} 2 & -2 & 0 \\ -1 & 1 & 0 \end{array}\right]$ 1/5 x 2nd row

$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$ 1/2 x 1st row added to 2nd row

Thus $y_1 = y_2 = t$

Thus, eigenvectors $\boldsymbol{x} = \begin{bmatrix} t \\ -\frac{2}{5}t \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} t \\ t \end{bmatrix}$.

6.5. (*Some more*) Matrix Calculus

So far, we've already done plenty of calculus with matrices when we covered real valued functions. In this section, the goal is to summarize main results of what is happening when we are undertaking derivatives of matrices. We won't spend too much time on this unless the class wants to. Let's go through it and work through an example to see what the **main results** are saying. Some stuff here is more of **nice to know** but still intuitive.

Example. Let $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Given:

$$y_1 = x_1^2 - x_2$$
$$y_2 = x_3^2 + 3x_2$$

Find $\frac{dy}{dx}$ which is the partial derivative matrix can be computed as follows:

Recall, $\frac{dy}{dx} = \begin{bmatrix} \frac{dy_1}{dx_1} & \frac{dy_1}{dx_2} & \frac{dy_1}{dx_3} \\ \frac{dy_2}{dx_1} & \frac{dy_2}{dx_2} & \frac{dy_2}{dx_3} \end{bmatrix} = \begin{bmatrix} 2x & 1 & 0 \\ 0 & 3 & 2x_3 \end{bmatrix}$

Generally, if y is an m-element vector, and x is an n-element vector, then the first-order partial derivatives is given by a m x n matrix.

Please note that <u>many</u> resources (especially in the discipline of engineering) report this matrix of first order derivatives as the transpose of what we have… but in economics and statistics, we most usually use what I am presenting. Just something to keep in mind.

Main results:

$\boldsymbol{y}$	$\frac{d\boldsymbol{y}}{d\boldsymbol{x}}$
$\boldsymbol{Ax}$	$\boldsymbol{A}$
$\boldsymbol{x^T A}$	$\boldsymbol{A^T}$
$\boldsymbol{x^T x}$	$2\boldsymbol{x^T}$

Let's apply these results to a simple example to see what is happening:

$$Let\ \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, x^T = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_1}} & \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_2}} \\ \frac{\partial \boldsymbol{y_2}}{\partial \boldsymbol{x_1}} & \frac{\partial \boldsymbol{y_2}}{\partial \boldsymbol{x_2}} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \boldsymbol{A}$$

$$\boldsymbol{x}^T \boldsymbol{A} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} . \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 & 2x_1 + 4x_2 \end{bmatrix}$$

$$\boldsymbol{y} = \boldsymbol{x}^T \boldsymbol{A} = \begin{bmatrix} x_1 + 3x_2 & 2x_1 + 4x_2 \end{bmatrix}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_1}} & \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_2}} \\ \frac{\partial \boldsymbol{y_2}}{\partial \boldsymbol{x_1}} & \frac{\partial \boldsymbol{y_2}}{\partial \boldsymbol{x_2}} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \boldsymbol{A}^T$$

$$\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 x_1 + x_2 x_2 \end{bmatrix}$$

$$\boldsymbol{y} = \boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{x} = \begin{bmatrix} x_1 x_1 + x_2 x_2 \end{bmatrix}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_1}} & \frac{\partial \boldsymbol{y_1}}{\partial \boldsymbol{x_2}} \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix} = 2\boldsymbol{x}^T$$

The Chain Rule for Vector Functions:

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{bmatrix}$ $z = \ y = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix}$, where z is a function of y which is a function of x.

We can write $\frac{\partial z}{\partial x} = \begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} & \cdots & \frac{\partial z_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_m}{\partial x_1} & \frac{\partial z_m}{\partial x_2} & \cdots & \frac{\partial z_m}{\partial x_n} \end{bmatrix}$,

where $\frac{\partial z_i}{\partial x_j} = \sum_{q=1}^{r} \frac{\partial z_i}{\partial y_q} . \frac{\partial y_q}{\partial x_j}$ $\quad \begin{cases} i = 1,2,\ldots,m \\ j = 1,2,\ldots,n \end{cases}$

$$\frac{\partial z}{\partial x} = \begin{bmatrix} \sum_{q=1}^{r} \frac{\partial z_1}{\partial y_q} . \frac{\partial y_q}{\partial x_1} & \sum_{q=1}^{r} \frac{\partial z_1}{\partial y_q} . \frac{\partial y_q}{\partial x_2} & \cdots & \sum_{q=1}^{r} \frac{\partial z_1}{\partial y_q} . \frac{\partial y_q}{\partial x_n} \\ \sum_{q=1}^{r} \frac{\partial z_2}{\partial y_q} . \frac{\partial y_q}{\partial x_1} & \sum_{q=1}^{r} \frac{\partial z_2}{\partial y_q} . \frac{\partial y_q}{\partial x_2} & \cdots & \sum_{q=1}^{r} \frac{\partial z_2}{\partial y_q} . \frac{\partial y_q}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{q=1}^{r} \frac{\partial z_m}{\partial y_q} . \frac{\partial y_q}{\partial x_1} & \sum_{q=1}^{r} \frac{\partial z_m}{\partial y_q} . \frac{\partial y_q}{\partial x_2} & \cdots & \sum_{q=1}^{r} \frac{\partial z_m}{\partial y_q} . \frac{\partial y_q}{\partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_1}{\partial y_2} & \cdots & \frac{\partial z_1}{\partial y_n} \\ \frac{\partial z_2}{\partial y_1} & \frac{\partial z_2}{\partial y_2} & \cdots & \frac{\partial z_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_m}{\partial y_1} & \frac{\partial z_m}{\partial y_2} & \cdots & \frac{\partial z_m}{\partial y_n} \end{bmatrix} . \begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} & \cdots & \frac{\partial z_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_m}{\partial x_1} & \frac{\partial z_m}{\partial x_2} & \cdots & \frac{\partial z_m}{\partial x_n} \end{bmatrix}$$

$$= \frac{\partial z}{\partial y} . \frac{\partial y}{\partial x} = \frac{\partial z}{\partial y} \frac{\partial y}{\partial x}$$

This is the chain rule for all vectors:

$$\frac{\partial z}{\partial x} = \frac{\partial y}{\partial x}\frac{\partial z}{\partial y} = \frac{\partial z}{\partial y}\frac{\partial y}{\partial x}$$

One last note: This is a conventional chain rule of calculus. Note, however, we are dealing with vectors, and the *chain* builds towards the left. For example, if w is a function of z, which is a function of y, which is a function of x, we have

$$\frac{\partial w}{\partial x} = \frac{\partial y}{\partial x}\frac{\partial z}{\partial y}\frac{\partial w}{\partial z}$$

On the other hand, in the ordinary chain rule one can build the product to the right or to the left because scalar multiplication is commutative.

There is SO much more you could learn about matrix differential; it is a powerful tool as you can imagine. However, we have covered more than enough. The goal is for you to know these are here and you should be able to read these matrices when you come across them in your notes or readings. And maybe someday use these in your own research?