UCI Paul Merage School of Business

Leadership for a Digitally Driven World™

MFIN 290: Financial Econometrics

Lecture 1-2

Math!

- Working with Matrices, Matrix Calculus
- Text Reference: Greene – Appendices A, B (posted on site as pdf)
- If you haven't seen matrices, this will be challenging. **Please don't panic**.
- We use matrices mostly at the outset of the course and you will be expected to be comfortable with this representation and some standard derivations with a graduate training in Econometrics.

Matrices

A matrix is a rectangular array of numbers:

$$\mathbf{A} = [a_{ik}] = [\mathbf{A}]_{ik} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix}$$

Matrices

$$\mathbf{A} = [a_{ik}] = [\mathbf{A}]_{ik} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix}$$

Consists of k column vectors and n row vectors

Referred to as "an n x k matrix"

When you are first getting acclimated, it can be helpful to write the dimensions below each matrix…

Matrices

Data is formatted such that each observation is a row and each feature observed is a column:

	Column				
Row	1 Year	2 Consumption (billions of dollars)	3 GNP (billions of dollars)	4 GNP Deflator	5 Discount Rate (N.Y Fed., avg.)
1	1972	737.1	1185.9	1.0000	4.50
2	1973	812.0	1326.4	1.0575	6.44
3	1974	808.1	1434.2	1.1508	7.83
4	1975	976.4	1549.2	1.2579	6.25
5	1976	1084.3	1718.0	1.3234	5.50
6	1977	1204.4	1918.3	1.4005	5.46
7	1978	1346.5	2163.9	1.5042	7.46
8	1979	1507.2	2417.8	1.6342	10.28
9	1980	1667.2	2633.1	1.7864	11.77

Source: Data from the *Economic Report of the President* (Washington, D.C.: U.S. Government Printing Office, 1983).

Manipulation of Matrices

Equality:

$$\boldsymbol{A} = \boldsymbol{B} \;\; if\ and\ only\ if\ [\boldsymbol{a_{ik}}] = [\boldsymbol{b_{ik}}]\ for\ all\ i\ and\ k$$

Transposition:

$$\boldsymbol{A} = [\boldsymbol{a_{ik}}] \Rightarrow \boldsymbol{A}' = [\boldsymbol{a_{ki}}]$$

Symmetry: matrix A is symmetric if

$$\boldsymbol{A} = \boldsymbol{A}'$$

Q: if A is n x k, what is the dimension of A’?

MATLAB

Command Window

```
>> A = [1,2;3,4]

A =

     1     2
     3     4

>> A'

ans =

     1     3
     2     4

>> A+A'
```

MATLAB is designed to program with matrices.

You enter matrices directly, and use spaces or commas to separate columns, semicolons to separate rows.

It is very good for simulation and manipulation. You can do econometrics here, but I have always found it awkward (i.e., not good for learning).

Manipulation of Matrices

```
Command Window
>> A = [1,2;3,4]

A =

     1     2
     3     4

>> A'

ans =

     1     3
     2     4

>> A+A'
```

```
>> A+A'

ans =

     2     5
     5     8

>> A-A

ans =

     0     0
     0     0

>> (A-A)'

ans =

     0     0
     0     0
```

Adding a matrix to its transpose yields a symmetric matrix (can you prove this?)

If matrices are equal, then the difference will be a conformable matrix of zeros

Manipulation of Matrices

Addition/Subtraction:

$$C = A + B = [a_{ik} + b_{ik}]$$
$$D = A - B = [a_{ik} - b_{ik}]$$

Addition here has the usual properties: associative, commutative, transposes filter through:

$$A + B = B + A$$
$$(A + B) + C = A + (B + C)$$
$$(A + B)' = A' + B'$$
$$A + 0 = A$$

Where the matrix $Z = 0 \Rightarrow z_{ik} = 0$ for all i and k

Manipulation of Matrices

Multiplication:

Two vectors a and b are multiplied using the inner (or dot product)

$$a = [a_k]_{k=1}^{K}$$
$$b = [b_k]_{k=1}^{K}$$

$$a'b = [a_1, a_2, \dots, a_k] \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_k \end{bmatrix}$$

$$a \cdot b = a_1 b_1 + \dots + a_k b_k = \sum a_i b_i$$

Can see we have $a'b = b'a$

These need to have the same number of elements (“dimension”)

Manipulation of Matrices - MATLAB

The Paul Merage School of Business

Command Window

```
>> x = [1,2,3,4]

x =

     1     2     3     4

>> y = [1;1;2;2]

y =

     1
     1
     2
     2

>> x*y

ans =

    17
```

Commas separate columns

Semicolons separate rows

$$1 * 1 + 2 * 1 + 3 * 2 + 4 * 2 = 17$$

Example #2: Inner products

Examples

Multiply the vector x with a column vector of ones such that it can be multiplied (the dimensions of the vector of ones “conform” to x).

$$x'\mathbf{1} = [\sum_n x_n \times 1 = \sum_n x_n$$

$$\Rightarrow \bar{x} = \frac{1}{n} x'\mathbf{1}$$

Inner product of vector x with itself? Can you write the variance in matrix notation?

$$x'x = \sum_n x_n \times x_n = \sum_n x_n^2$$

Covariance?

HINT: $\mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$

Manipulation of Matrices

Multiplication:

If we have an $n\ x\ k$ matrix A and a $k\ x\ m$ matrix B and we define $C = AB$. C will be dimension $n\ x\ m$.

The ij^{th} element of C is given by the inner product of the i^{th} row of A and the j^{th} column of B

$$C = [c_{ij}] = [\sum_k a_{ik} b_{kj}]$$

To multiply two matrices then, the number of columns in the first must equal the number or rows in the second! **This means that $\boldsymbol{AB}$ need not equal $\boldsymbol{BA}$** (indeed, BA may not be defined, as in this case when $m \neq n$!).

We will often "premultiply" or "postmultiply" by matrices. Keep track of the dimensions! When in doubt, write them out.

Manipulation of Matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \; B = \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$$

$$C = AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$$

$$C_{11} = 1(2) + 2(4) = 10$$

Manipulation of Matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \; B = \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$$

$$C = AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1(2) + 2(4) & 1(1) + 2(0) \\ 3(2) + 4(4) & 3(1) + 4(0) \end{bmatrix} = \begin{bmatrix} 10 & 1 \\ 22 & 3 \end{bmatrix}$$

What is BA? What is B’A’?

Example: Matrix Multiplication

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

$$\underline{B'A'} = \underline{\begin{bmatrix} 10 & 22 \\ 1 & 3 \end{bmatrix}}$$

$$(AB)' = B'A'$$

Manipulation of Matrices

```
Command Window
>> A = [1,2;3,4]

A =

     1     2
     3     4

>> A'

ans =

     1     3
     2     4

>> A+A'
```

```
>> B = [2 1;4 0]

B =

     2     1
     4     0

>> A*B

ans =

    10     1
    22     3

>> A(1,:)*B(:,1)

ans =

    10
```

"*" Is the multiplication operator. Be careful here. A good idea to check dimensions

(i, j) refers to elements. Using a colon here refers to all of them. Can also use "1:end" to refer to everything

Manipulation of Matrices

```
>> B*A

ans =

     5     8
     4     8

>> B'*A'

ans =

    10    22
     1     3
```

$AB \neq BA$ in general, even if conformable

$(AB)' = B'A'$ always

Manipulation of Matrices

In matrix multiplication, the identity matrix is analogous to the scalar 1. For any conformable matrix or vector **A**, $\mathbf{AI} = \mathbf{A}$. In addition, $\mathbf{IA} = \mathbf{A}$, although if **A** is not a square matrix, the two identity matrices are of different orders.

A conformable matrix of zeros produces the expected result: $\mathbf{A0} = \mathbf{0}$.

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

Some general rules for matrix multiplication are as follows:

Associative law: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.
Distributive law: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.
Transpose of a product: $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.
Transpose of an extended product: $(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$.

Manipulation of Matrices

The inverse of a matrix A, if it exists, is written as A^{-1} and is defined (if and only if)

$$AA^{-1} = I$$

This implies:

$$(A^{-1})^{-1} = A$$
$$(A^{-1})' = (A')^{-1}$$

And if A and B both have inverses, then

$$(AB)^{-1} = B^{-1}A^{-1}$$

Manipulation of Matrices

The inverse of a matrix A, if it exists, is written as A^{-1} and is defined (if and only if)

$$AA^{-1} = I$$

This implies:

$$(A^{-1})^{-1} = A$$
$$(A^{-1})' = (A')^{-1}$$

If A and B both have inverses, then

$$(AB)^{-1} = B^{-1}A^{-1}$$

$AA^{-1} = I$

$(AA^{-1})' = I' = I$

$A^{-1'}A' = I' = I$

$A^{-1'} = (A')^{-1}$

Manipulation of Matrices

The inverse of a matrix A, if it exists, is written as A^{-1} and is defined

$$AA^{-1} = I$$

This implies:

$$(A^{-1})^{-1} = A$$
$$(A^{-1})' = (A')^{-1}$$

If A and B both have inverses, then

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$AB(AB)^{-1} = I = A\underbrace{\cancel{BB^{-1}}}_{I}A^{-1} = \cancel{AIA^{-1}} = I$$

Matrix Inverse

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Intuition: How to solve for A^{-1}

We know A^{-1} has to be 2x2 with elements $\begin{matrix} w & x \\ y & z \end{matrix}$ such that $AA^{-1} = I$

$$AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Four equations, four unknowns!

Matrix Inverse

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Show that $B = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ is the inverse of A

$$AB = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & db-bd \\ -cd+cd & -cb+da \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Matrices as Equations

Given an n x k matrix A (such as independent variables in a dataset) and k x 1 vector b (such as OLS coefficients),

we can write an n x 1 vector c (such as a dependent variable prediction) as

$$c = Ab$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

Example: OLS

One of the key OLS assumptions is that the independent variables and errors are orthogonal.

This means that the inner product: $X'e = 0$ (this is related to the covariance.. Can you see it?)

$$y = Xb + e$$
$$e = y - Xb$$
$$X'e = X'(y - Xb) = X'y - X'Xb = 0$$

$\Rightarrow X'y = X'Xb$ (we will return to this next lecture!)

Q: What are the dimensions of $X'y, X'X,$ and $X'Xb$ if there are n observations and k different x's?

Example: OLS

One of the key OLS assumptions is that the independent variables and errors are orthogonal.

This means that the inner product: $X'e = 0$ (this is related to the covariance.. Can you see it?)

$$y = Xb + e$$
$$e = y - Xb$$
$$X'e = X'(y - Xb) = X'y - X'Xb = 0$$

$\Rightarrow X'y = X'Xb$ (we will return to this next lecture!)

Let $A = X'X$. If it exists, we have $A^{-1} = (X'X)^{-1}$. Premultiplying both sides by A^{-1}:

$A^{-1}X'y = A^{-1}X'Xb$

$(X'X)^{-1}X'y = (X'X)^{-1}X'Xb$

$(X'X)^{-1}X'y = Ib = b$

Example (Greene, App. A)

$$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

We can interpret this in two ways. First, it is a compact way of writing the three equations

$$\begin{aligned} 5 &= 4a + 2b + 1c, \\ 4 &= 2a + 6b + 1c, \\ 1 &= 1a + 1b + 0c. \end{aligned}$$

Second, by writing the set of equations as

$$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = a \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

Matrices as Equations

Second, by writing the set of equations as

$$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = a \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

Can write this equation as a linear combination of the columns of the original matrix

Matrices

Treat the fields in your data as parts of that equation!

	Column				
Row	***1 Year***	***2 Consumption (billions of dollars)***	***3 GNP (billions of dollars)***	***4 GNP Deflator***	***5 Discount Rate (N.Y Fed., avg.)***
1	1972	737.1	1185.9	1.0000	4.50
2	1973	812.0	1326.4	1.0575	6.44
3	1974	808.1	1434.2	1.1508	7.83
4	1975	976.4	1549.2	1.2579	6.25
5	1976	1084.3	1718.0	1.3234	5.50
6	1977	1204.4	1918.3	1.4005	5.46
7	1978	1346.5	2163.9	1.5042	7.46
8	1979	1507.2	2417.8	1.6342	10.28
9	1980	1667.2	2633.1	1.7864	11.77

Source: Data from the *Economic Report of the President* (Washington, D.C.: U.S. Government Printing Office, 1983).

Linear Dependence

The column vectors X and y are "linearly dependent" if there is some linear combination of the columns of X that produce y…

Equivalently, if there is some conformable vector a (that is not all zeros) such that:

$$y = Xa = \sum_{n} a_i \boldsymbol{x}_i$$

Where:

$$y \text{ } is \text{ } n \text{ } x \text{ } 1$$
$$X \text{ } is \text{ } n \text{ } x \text{ } k$$
$$a \text{ } is \text{ } k \text{ } x \text{ } 1$$

Linear Dependence

This usually comes up with some sort of mistake. Value measured in dollars at date t and yen at date t are linearly dependent based on the FX rate

Dummy variables equal to one for each category and an intercept term are linearly dependent (later)

Linear Dependence

Example

$$\begin{matrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 * \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 * \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Linear Dependence

Example

$$\begin{matrix} 5 & 2 & 6 \\ 4 & 3 & 9 \\ 1 & 2 & 6 \end{matrix}$$

$$\begin{bmatrix} 6 \\ 9 \\ 6 \end{bmatrix} = 3 * \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}$$

Linear Dependence

Example

$$\begin{matrix} 5 & 2 & 1 \\ 4 & 3 & 5 \\ 1 & 2 & 5 \end{matrix}$$

$$\begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = 3 * \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} - 1 * \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$$

Linear Independence

The columns of X are linearly independent if none of them can be written as a linear combination of the others.

That is, the only solution a to $Xa = \sum_n a_i \boldsymbol{x}_i$ has $a_i = 0$ for all i

Note: This ensures that the matrix $(X'X)$ has an inverse

Lots of ways that you can show this… for us, it basically means that each column of X contains at least SOME new information when compared to the rest of the X's. If it's not a lot, that can cause problems, but the inverse will still exist (at least numerically)

Positive Definite

A matrix C is called positive definite if the value

$x'Cx > 0$ for any vector x

Q: if we have n observations, what is the dimension of C? $x'Cx$?

Variance covariance matrices are positive definite

Lots of other ways to show this. We don't talk about eigenvalues much in this course, but they are very interesting and very helpful for a few things. A positive definite matrix has all positive eigenvalues.

Matrix Calculus

Take the scalar $y = f(x_1, \ldots, x_n) = f(\boldsymbol{x})$

The vector of partial derivatives is called the gradient and is given l

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \partial y/\partial x_1 \\ \partial y/\partial x_2 \\ \cdots \\ \partial y/\partial x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_n \end{bmatrix}$$

$f(x_1, x_2) = x_1^2 + x_2^2$

$\frac{\partial f}{\partial x} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$

Matrix Calculus

Example

$$y = f(x_1, x_2) = f(\boldsymbol{x}) = x_1^2 - 3x_2 + 5x_1x_2$$

$$\frac{\partial y}{\partial x_1} = 2\mathrm{x}_1 + 5\mathrm{x}_2$$

$$\frac{\partial y}{\partial x_2} = -3 + 5\mathrm{x}_1$$

Matrix Calculus

Take the scalar valued (i.e., the output is a number) function $y = f(x_1, \dots, x_n) = f(\boldsymbol{x})$ where $\boldsymbol{x}$ is a column vector

$$x = \begin{matrix} x_1 \\ \dots \\ x_n \end{matrix}$$

The vector of partial derivatives is called the "gradient" and is given by:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \partial y/\partial x_1 \\ \partial y/\partial x_2 \\ \cdots \\ \partial y/\partial x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_n \end{bmatrix}$$

In this case, it is a column vector – but the shape is determined by the denominator. If X was a matrix, it would be a matrix as well

Matrix Calculus

Example

$$y = f(x_1, x_2) = f(\boldsymbol{x}) = x_1^2 - 3x_2 + 5x_1x_2$$

$$\frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} 2\mathrm{x}_1 + 5\mathrm{x}_2 \\ -3 + 5\mathrm{x}_1 \end{bmatrix}$$

What is the second derivative matrix dy/dxdx'?

Matrix Calculus

The matrix of second derivatives is written as:

$$\mathbf{H} = \begin{bmatrix} \partial^2 y/\partial x_1 \partial x_1 & \partial^2 y/\partial x_1 \partial x_2 & \cdots & \partial^2 y/\partial x_1 \partial x_n \\ \partial^2 y/\partial x_2 \partial x_1 & \partial^2 y/\partial x_2 \partial x_2 & \cdots & \partial^2 y/\partial x_2 \partial x_n \\ \cdots & \cdots & \cdots & \cdots \\ \partial^2 y/\partial x_n \partial x_1 & \partial^2 y/\partial x_n \partial x_2 & \cdots & \partial^2 y/\partial x_n \partial x_n \end{bmatrix} = [f_{ij}]$$

Here, the ij^{th} element of H is the derivative of $f(\boldsymbol{x})$ w.r.t. the i^{th} and then the j^{th} element

Young's Theorem: H is almost always symmetric!

Matrix Calculus

The matrix of second derivatives is written as:

$$\mathbf{H} = \begin{bmatrix} \partial^2 y/\partial x_1 \partial x_1 & \partial^2 y/\partial x_1 \partial x_2 & \cdots & \partial^2 y/\partial x_1 \partial x_n \\ \partial^2 y/\partial x_2 \partial x_1 & \partial^2 y/\partial x_2 \partial x_2 & \cdots & \partial^2 y/\partial x_2 \partial x_n \\ \cdots & \cdots & \cdots & \cdots \\ \partial^2 y/\partial x_n \partial x_1 & \partial^2 y/\partial x_n \partial x_2 & \cdots & \partial^2 y/\partial x_n \partial x_n \end{bmatrix} = [f_{ij}]$$

That is, each column of H is the derivative of the gradient vector w.r.t. each element in $\boldsymbol{x}'$

$$\mathbf{H} = \left[\frac{\partial(\partial y/\partial \mathbf{x})}{\partial x_1} \frac{\partial(\partial y/\partial \mathbf{x})}{\partial x_2} \cdots \frac{\partial(\partial y/\partial \mathbf{x})}{\partial x_n} \right] = \frac{\partial(\partial y/\partial \mathbf{x})}{\partial(x_1 \quad x_2 \cdots x_n)} = \frac{\partial(\partial y/\partial \mathbf{x})}{\partial \mathbf{x}'} = \frac{\partial^2 y}{\partial \mathbf{x} \partial \mathbf{x}'}$$

What is H for this example?

Example

$$y = f(x_1, x_2) = f(\boldsymbol{x}) = x_1^2 - 3x_2 + 5x_1x_2$$

$$\frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} 2\mathrm{x}_1 + 5\mathrm{x}_2 \\ -3 + 5\mathrm{x}_1 \end{bmatrix}$$

Matrix Calculus

Example

$$x = \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix}; a = \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix}; y = a'x$$

$$y = a'\boldsymbol{x} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = a_1 x_1 + a_2 x_2 + a_3 x_3 = \sum_n a_n x_n$$

Note: $\boldsymbol{x}'a = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \sum_n x_n a_n$

Matrix Calculus

Example

$$x = \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix};\ a = \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix};\ y = a'x$$

$$\frac{\partial y}{\partial \boldsymbol{x}} = \frac{\partial(a'\boldsymbol{x})}{\partial \boldsymbol{x}} = \frac{\partial(\sum_n x_n a_n)}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial(\sum_n x_n a_n)}{\partial x_1} \\ \cdots \\ \frac{\partial(\sum_n x_n a_n)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ \cdots \\ a_n \end{bmatrix} = a$$

Which means:

$$\frac{\partial(a'\boldsymbol{x})}{\partial \boldsymbol{x}} = a\ (\neq a'!)$$

Matrix Calculus

Which means:

$$\frac{\partial(a'\boldsymbol{x})}{\partial\boldsymbol{x}} = a \; (\neq a'!)$$

$$\frac{\partial(a'\boldsymbol{x})}{\partial\boldsymbol{x}'} = a'$$

Matrix Calculus

A linear function between y and the column vector x is written:

$$y = a'\boldsymbol{x} = \boldsymbol{x}'a = \sum_n x_n a_n$$

$$\frac{\partial (a'\boldsymbol{x})}{\partial \boldsymbol{x}} = a \; (\neq a'!)$$

Matrix Calculus

With a set of linear functions (say, a dataset):

$$\boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x}$$

$$\begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} = \begin{matrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{matrix} * \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} = \begin{matrix} a_{11}x_1 + a_{21}x_2 + a_{31}x_3 \\ ? \\ ? \end{matrix}$$

Note that we have flipped the order of the subscripts on a since this is a transpose…

Matrix Calculus

$$\boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x}$$

$$\begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} = \begin{matrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{matrix} * \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} = \begin{matrix} a_{11}x_1 + a_{21}x_2 + a_{31}x_3 \\ ? \\ ? \end{matrix}$$

Each element (row) of y is given by

$$y_i = a_i'\boldsymbol{x}$$

Where $a_i{}'$ is the i^{th} row of $A' \Rightarrow a_1' = [a_{11} \quad a_{21} \quad a_{31}]$

Matrix Calculus

$$\Rightarrow \frac{\partial y_i}{\partial \boldsymbol{x}} = \frac{\partial a_i' \boldsymbol{x}}{\partial \boldsymbol{x}}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{21}x_2 + a_{31}x_3 \\ a_{12}x_1 + a_{22}x_2 + a_{32}x_3 \\ a_{13}x_1 + a_{23}x_2 + a_{33}x_3 \end{bmatrix}$$

$$\frac{\partial y_1}{\partial \boldsymbol{x}} = \begin{matrix} \frac{\partial [a_{11}x_1 + a_{21}x_2 + a_{31}x_3]}{\partial x_1} \\ \frac{\partial [a_{11}x_1 + a_{21}x_2 + a_{31}x_3]}{\partial x_2} \\ \frac{\partial [a_{11}x_1 + a_{21}x_2 + a_{31}x_3]}{\partial x_3} \end{matrix} = \begin{matrix} a_{11} \\ a_{21} \\ a_{31} \end{matrix} = a_1$$

Matrix Calculus

$$\frac{\partial y_i}{\partial \boldsymbol{x}} = \frac{\partial a_i' \boldsymbol{x}}{\partial \boldsymbol{x}} = (a_i')' = a_i$$

You should convince yourselves with a similar derivation that

$$\begin{bmatrix} \frac{\partial y_1}{\partial x'} \\ \cdots \\ \frac{\partial y_n}{\partial x'} \end{bmatrix} = \begin{bmatrix} a_1{}' \\ \cdots \\ a_n{}' \end{bmatrix}$$

Matrix Calculus

$$\frac{\partial y_i}{\partial \boldsymbol{x}} = \frac{\partial a_i' \boldsymbol{x}}{\partial \boldsymbol{x}} = (a_i')' = a_i$$

This means:

$$\frac{\partial y}{\partial x} = \frac{\partial Ax}{\partial x'} = \begin{bmatrix} a_1{}' \\ \dots \\ a_n{}' \end{bmatrix} = A$$

$$\frac{\partial x' A'}{\partial x} = A'$$

Matrix Calculus

A quadratic form is written

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij}.$$

For example,

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix},$$

so that

$$\mathbf{x}'\mathbf{A}\mathbf{x} = 1x_1^2 + 4x_2^2 + 6x_1x_2.$$

Matrix Calculus

$$x'Ax = x_1^2 + 4x_2^2 + 6x_1x_2$$

$$\frac{\partial x'Ax}{\partial x} = \begin{bmatrix} \frac{\partial x'Ax}{\partial x_1} \\ \frac{\partial x'Ax}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 + 6x_2 \\ 8x_2 + 6x_1 \end{bmatrix}$$

$$= 2 \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{\partial x'Ax}{\partial x} = 2Ax$$

Matrix Calculus

$$x'Ax = x_1^2 + 4x_2^2 + 6x_1x_2$$

$$\frac{\partial x'Ax}{\partial x} = \begin{bmatrix} \frac{\partial x'Ax}{\partial x_1} \\ \frac{\partial x'Ax}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 + 6x_2 \\ 8x_2 + 6x_1 \end{bmatrix} =$$

$$\begin{bmatrix} 2 & 6 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2Ax$$ which is the general result when **A** is a symmetric matrix.

And we have $\frac{\partial x'Ax}{\partial x_1 \partial x_2} = 6 = \frac{\partial x'Ax}{\partial x_2 \partial x_1}$ $(symmetric)$

Recap

I recommend you practice this with a few examples until you convince yourself/can remember them readily. We will use these identities in several places in this course.

Greene Appendices A and B are a good summary resource and are on the course website.

Final Project

Time to discuss

find out who you want to work with, talk about what kinds of questions you find interesting, what kind of data or questions you want to look at (and talk about in future job interviews!)