Matrix Refresher

Brandon Stewart[1]

Princeton

October 18, 2018

[1] These slides borrow liberally from Matt Blackwell and Adam Glynn.

Where We've Been and Where We're Going...

- Lecture 7 presents the full matrix version of OLS. These slides provide a compact summary of the key matrix concepts that are necessary to follow along.
- Khan Academy provides a number of other great linear algebra resources.
- You might also find the following books helpful:
  - *Essential Mathematics for Political and Social Research* by Jeff Gill
  - *A Mathematics Course for Political & Social Research* by Moore and Siegel
- Both books are designed to cover the mathematical content for this particular kind of course sequence.
- For those going on to future training, matrix representations of linear regression are the most common form you will see so it is good to have the basics.

Why Matrices and Vectors?

Here's one way to write the full multiple regression model:

$$y_i = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{iK}\beta_K + u_i$$

- Notation is going to get needlessly messy as we add variables
- Matrices are clean, but they are like a foreign language
- They make take some practice but it will make everything substantially more general

Matrices and Vectors

- A matrix is just a rectangular array of numbers.
- We say that a matrix is $n \times p$ ("n by p") if it has n rows and p columns.
- Uppercase bold denotes a matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

- Generic entry: a_{ij} where this is the entry in row i and column j

Design Matrix

One example of a matrix that we'll use a lot is the design matrix, which has a column of ones, and then each of the subsequent columns is each independent variable in the regression.

$$\mathbf{X} = \begin{bmatrix} 1 & \text{exports}_1 & \text{age}_1 & \text{male}_1 \\ 1 & \text{exports}_2 & \text{age}_2 & \text{male}_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \text{exports}_n & \text{age}_n & \text{male}_n \end{bmatrix}$$

Vectors

- A vector is just a matrix with only one row or one column.
- A row vector is a vector with only one row, sometimes called a $1 \times p$ vector:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_p \end{bmatrix}$$

- A column vector is a vector with one column and more than one row. Here is a $n \times 1$ vector:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- Convention: we'll assume that a vector is column vector and (when we are being formal) vectors will be written with lowercase bold lettering (**b**).

NB: I'm not completely consistent about bolding vectors and matrices. If you are unclear about the dimensionality of an object (is it a scalar? a vector? a matrix?) just ask.

Vector Examples

One common vector that we will work with are individual variables, such as the dependent variable, which we will represent as $\mathbf{y}$:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Transpose

- There are many operations we'll do on vectors and matrices, but one is very fundamental: the transpose.
- The transpose of a matrix $\mathbf{A}$ is the matrix created by switching the rows and columns of the data and is denoted $\mathbf{A}'$. That is, the pth column becomes the pth row.

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \\ q_{31} & q_{32} \end{bmatrix} \quad \mathbf{Q}' = \begin{bmatrix} q_{11} & q_{21} & q_{31} \\ q_{12} & q_{22} & q_{32} \end{bmatrix}$$

If $\mathbf{A}$ is of dimension $n \times p$, then $\mathbf{A}'$ will be dimension $p \times n$.

Transposing Vectors

Transposing will turn a $p \times 1$ column vector into a $1 \times p$ row vector and vice versa:

$$\boldsymbol{\omega} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ -5 \end{bmatrix} \qquad \boldsymbol{\omega}' = \begin{bmatrix} 1 & 3 & 2 & -5 \end{bmatrix}$$

Write matrices as vectors

- A matrix is just a collection of vectors (row or column)
- As a row vector:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \end{bmatrix}$$

with row vectors $\mathbf{a}_1' = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix}$ $\mathbf{a}_2' = \begin{bmatrix} a_{21} & a_{22} & a_{23} \end{bmatrix}$

- Or we can define it in terms of column vectors:

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} \mathbf{b_1} & \mathbf{b_2} \end{bmatrix}$$

where $\mathbf{b_1}$ and $\mathbf{b_2}$ represent the columns of $\mathbf{B}$.

Addition and Subtraction

- To perform certain operations (e.g. addition, subtraction, multiplication, inversion and exponentiation) the matrices/vectors need to be conformable. Conformable means that the dimensions are suitable for defining that operation.
- Two matrices are conformable for addition or subtraction if they have the same dimensions.
- Let $\mathbf{A}$ and $\mathbf{B}$ both be 2×2 matrices. Then, let $\mathbf{C} = \mathbf{A} + \mathbf{B}$, where we add each cell together:

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

Scalar Multiplication

- A scalar is just a single number: you can think of it sort of like a 1 by 1 matrix.
- When we multiply a scalar by a matrix, we just multiply each element/cell by that scalar.
- Whenever we do things cell-by-cell like this we sometimes call it an elementwise/entrywise operation

$$\alpha \mathbf{A} = \alpha \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \alpha \times a_{11} & \alpha \times a_{12} \\ \alpha \times a_{21} & \alpha \times a_{22} \end{bmatrix}$$

The Linear Model with New Notation

- Remember that we wrote the linear model as the following for all $i \in [1, \dots, n]$:

$$y_i = \beta_0 + x_i\beta_1 + z_i\beta_2 + u_i$$

- Imagine we had an n of 4. We could write out each formula:

$$\begin{aligned} y_1 &= \beta_0 + x_1\beta_1 + z_1\beta_2 + u_1 \quad \text{(unit 1)} \\ y_2 &= \beta_0 + x_2\beta_1 + z_2\beta_2 + u_2 \quad \text{(unit 2)} \\ y_3 &= \beta_0 + x_3\beta_1 + z_3\beta_2 + u_3 \quad \text{(unit 3)} \\ y_4 &= \beta_0 + x_4\beta_1 + z_4\beta_2 + u_4 \quad \text{(unit 4)} \end{aligned}$$

The Linear Model with New Notation

$$
\begin{aligned}
y_1 &= \beta_0 + x_1\beta_1 + z_1\beta_2 + u_1 \quad \text{(unit 1)} \\
y_2 &= \beta_0 + x_2\beta_1 + z_2\beta_2 + u_2 \quad \text{(unit 2)} \\
y_3 &= \beta_0 + x_3\beta_1 + z_3\beta_2 + u_3 \quad \text{(unit 3)} \\
y_4 &= \beta_0 + x_4\beta_1 + z_4\beta_2 + u_4 \quad \text{(unit 4)}
\end{aligned}
$$

- We can write this as:

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \beta_0 + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \beta_1 + \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} \beta_2 + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}
$$

- Outcome is a linear combination of the the **x**, **z**, and **u** vectors

Grouping Things into Matrices

- Can we write this in a more compact form?
  Yes! Let $\mathbf{X}$ and $\boldsymbol{\beta}$ be the following:

$$\underset{(4\times3)}{\mathbf{X}} = \begin{bmatrix} 1 & x_1 & z_1 \\ 1 & x_2 & z_2 \\ 1 & x_3 & z_3 \\ 1 & x_4 & z_4 \end{bmatrix} \quad \underset{(3\times1)}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Matrix multiplication by a vector

- We can write this more compactly as a matrix (post-)multiplied by a vector:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \beta_0 + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \beta_1 + \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} \beta_2 = \mathbf{X}\boldsymbol{\beta}$$

- Multiplication of a matrix by a vector is just the linear combination of the columns of the matrix with the vector elements as weights/coefficients.
- And the left-hand side here only uses scalars times vectors, which is easy!

General Matrix by Vector Multiplication

- $\mathbf{A}$ is a $n \times p$ matrix
- $\mathbf{b}$ is a $p \times 1$ column vector
- Columns of $\mathbf{A}$ have to match rows of $\mathbf{b}$ to be conformable for matrix-vector multiplication
- Let $\mathbf{a}_j$ be the jth column of A. Then we can write:

$$\underset{(j \times 1)}{\mathbf{c}} = \mathbf{Ab} = b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_p\mathbf{a}_p$$

- $\mathbf{c}$ is linear combination of the columns of $\mathbf{A}$

Back to Regression

- Imagine that we have K covariates, such that we have $K+1$ different β (due to the intercept being β_0).
- $\mathbf{X}$ is the $n \times (p+1)$ design matrix of independent variables
- $\boldsymbol{\beta}$ be the $(K+1) \times 1$ column vector of coefficients.
- $\mathbf{X}\boldsymbol{\beta}$ will be $n \times 1$:

$$\mathbf{X}\boldsymbol{\beta} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_K \mathbf{x}_K$$

- We can compactly write the linear model as the following:

$$\underset{(n\times 1)}{\mathbf{y}} = \underset{(n\times 1)}{\mathbf{X}\boldsymbol{\beta}} + \underset{(n\times 1)}{\mathbf{u}}$$

- We can also write this at the individual level, where $\mathbf{x}'_i$ is the ith row of $\mathbf{X}$:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + u_i$$

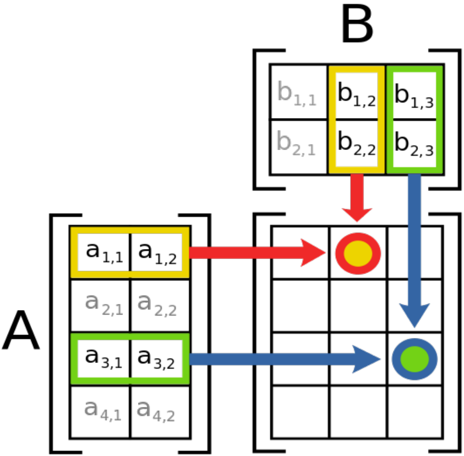
Matrix Multiplication

- What if, instead of a column vector b, we have a matrix $\mathbf{B}$ with dimensions $p \times m$.
- How do we do multiplication like so $\mathbf{C} = \mathbf{AB}$?
- Each column of the new matrix is just matrix by vector multiplication:

$$\mathbf{C} = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \cdots \quad \mathbf{c}_m] \qquad \mathbf{c}_i = \mathbf{A}\mathbf{b}_i$$

- Thus, each column of $\mathbf{C}$ is a linear combination of the columns of $\mathbf{A}$.

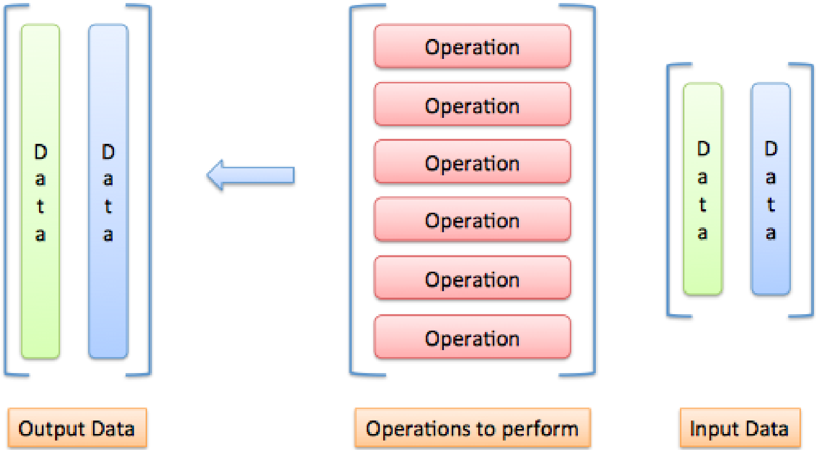
A visual approach to matrix multiplication



https://en.wikipedia.org/wiki/Matrix_multiplication#/media/File:Matrix_multiplication_diagram_2.svg

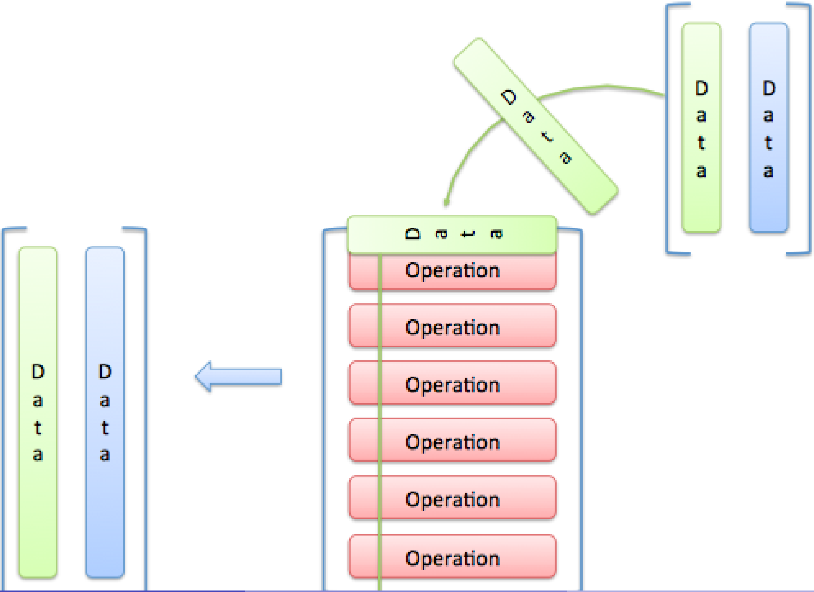
A second visual approach

Credit: `https://betterexplained.com/articles/linear-algebra-guide/`



A second visual approach

Credit: `https://betterexplained.com/articles/linear-algebra-guide/`



A dynamic approach to matrix multiplication

`https://www.khanacademy.org/math/precalculus/precalc-matrices/multiplying-matrices-by-matrices/v/matrix-multiplication-intro`

Properties of Matrix Multiplication

- Matrix multiplication is not commutative $\mathbf{AB} \neq \mathbf{BA}$
- It is associative and distributive

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$
$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$
$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Special Multiplications

- The inner product of a two column vectors **a** and **b** (of equal dimension, $p \times 1$):

$$\mathbf{a}'\mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_p b_p$$

- Special case of above: $\mathbf{a}'$ is a matrix with p columns and just 1 row, so the "columns" of $\mathbf{a}'$ are just scalars.

Sum of the Squared Residuals

- Example: let's say that we have a vector of residuals, $\widehat{\mathbf{u}}$, then the inner product of the residuals is:

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \begin{bmatrix} \widehat{u}_1 & \widehat{u}_2 & \cdots & \widehat{u}_n \end{bmatrix} \begin{bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \\ \vdots \\ \widehat{u}_n \end{bmatrix}$$

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \widehat{u}_1\widehat{u}_1 + \widehat{u}_2\widehat{u}_2 + \cdots + \widehat{u}_n\widehat{u}_n = \sum_{i=1}^{n} \widehat{u}_i^2$$

- It's just the sum of the squared residuals!

Square Matrices and the Diagonal

- A square matrix has equal numbers of rows and columns.
- The diagonal of a square matrix are the values a_{jj}:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

- The identity matrix, $\mathbf{I}$ is a square matrix, with 1s along the diagonal and 0s everywhere else.

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

- The identity matrix multiplied by any matrix returns the matrix: $\mathbf{AI} = \mathbf{A}$.

Scalar Inverses

- What is division in its simplest form? $\frac{1}{a}$ is the value such that $a\frac{1}{a} = 1$:
- For some algebraic expression: $au = b$, let's solve for u:

$$\frac{1}{a}au = \frac{1}{a}b$$
$$u = \frac{b}{a}$$

- Need a matrix version of this: $\frac{1}{a}$.

Matrix Inverses

Definition (Matrix Inverse)

If it exists, the **inverse** of square matrix $\mathbf{A}$, denoted $\mathbf{A}^{-1}$, is the matrix such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

- We can use the inverse to solve (systems of) equations:

$$\begin{aligned} \mathbf{A}\mathbf{u} &= \mathbf{b} \\ \mathbf{A}^{-1}\mathbf{A}\mathbf{u} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{I}\mathbf{u} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{u} &= \mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

- If the inverse exists, we say that $\mathbf{A}$ is invertible or nonsingular.