Maths Day Two: Linear Algebra

Princeton Sociology Methods Camp

Varun Satish[1]

[1] These slides are the collective effort of everyone who has contributed to the Princeton Sociology Methods Camp.

Outline

1. Vectors
   - How can we use vectors to represent data
   - Vector operations: addition/subtraction and multiplication (the dot product)
   - Motviating the use of matrices
2. Matrices
   - Moving from vectors to matrices
   - Types of matrices
   - Matrix operations: addition/subtracting and multiplication
   - The importance of checking dimensions and other information about matrix multiplication

Data source for vectors and matrix: senators' co-sponsorship of bills during 2004 session

Source for data: James H. Fowler: Connecting the Congress: A Study of Cosponsorship Networks , *Political Analysis* 14 (4): 456-487 (Fall 2006) and Legislative Cosponsorship Networks in the U.S. House and Senate, *Social Networks* 28 (4): 454-465 (October 2006). Cleaned senate network data provided as part of Skyler Cranmer, ICPSR 2016 Network Analysis workshop.

Motivation for vectors

- Compact way of storing information, instead of:

- Use:

	Gregg	Alexander	Isaakson	Lieberman
Burr	1	1	1	0

Vector notation

- **Example**: two vectors: John McCain's versus Russ Feingold's cosponsorship of bills with Hillary Clinton, Lincoln Chaffee, Joe Biden, and Strom Thurmond, stored in that order in the vector
- Let:

$$John McCain's cosponsorship = \mathbf{u} = \begin{bmatrix} HRC & LC & JB & ST \end{bmatrix} = \begin{bmatrix} 1 & 1 & 10 & 5 \end{bmatrix}$$

$$Russ\ Feingold's\ cosponsorship = \mathbf{v} = \begin{bmatrix} HRC & LC & JB & ST \end{bmatrix} = \begin{bmatrix} 2 & 2 & 8 & 1 \end{bmatrix}$$

- **Bold** = entire vector; non-bold = element of vector. For instance:
  - **u**: all of John McCain's sponsorship information
  - u_3: John McCain's sponsorship with Joe Lieberman (ten bills)
  - v_4: Russ Feingold's sponsorship with Strom Thurmond (one bill)

Vector notation: continued

Can arrange data in multiple ways depending on purpose:

- Row vector:
$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 10 & 5 \end{bmatrix}$$
- Column vector (in this case):
$$\mathbf{u}^T = \begin{bmatrix} 1 \\ 1 \\ 10 \\ 5 \end{bmatrix}$$
- Or vice versa (so if $\mathbf{u}$ were a column vector, $\mathbf{u}^T$ would be a row vector)

Vector Operations

We can:

1. Add/subtract vectors
2. Multiply vectors

Vector addition and subtraction

We can add and subtract vectors that **have the same dimensions**.

$$\mathbf{u}^T = \begin{bmatrix} 1 \\ 1 \\ 10 \\ 5 \end{bmatrix}, \mathbf{v}^T = \begin{bmatrix} 2 \\ 2 \\ 8 \\ 1 \end{bmatrix}$$

This occurs by applying the operation to the individual elements of each vector. For example, here:

$$\mathbf{u}^T + \mathbf{v}^T = \begin{bmatrix} 1+2 \\ 1+2 \\ 10+8 \\ 5+1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 18 \\ 6 \end{bmatrix}$$

Question: What about $\mathbf{u}^T + \mathbf{v}$?

Vector multiplication

- We can multiply vectors with scalars
- We can multiply vector with other vectors
- For the purposes of methods camp we are going to focus only on **dot products** and not on cross products

Multiplying by a scalar

We don't need to worry about the dimensions of the vector, we just multiply each element by the scalar. For example:

$$2 \cdot \mathbf{u}^T = \begin{bmatrix} 2 * 1 \\ 2 * 1 \\ 2 * 10 \\ 2 * 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 20 \\ 10 \end{bmatrix}$$

Multiplying vectors: The dot product

The dot product

- With addition and subtraction we needed to check that the vectors had the same dimensions.
- With the dot product, we also have to check the dimensions of the vectors, but we are looking for something different
- We need the number of **columns** of the first vector to match the number of **rows** of the second vector.
- We can multiply a 1×4 vector by a 4×1 vector, but not the other way around.

The dot product: example

$$\mathbf{u} = \underset{\color{red}{1\times 4}}{\begin{bmatrix}1 & 1 & 10 & 5\end{bmatrix}},\ \mathbf{v}^T = \underset{\color{red}{4\times 1}}{\begin{bmatrix}2\\2\\8\\1\end{bmatrix}}$$

Since these dimensions agree, then the dot product is given as:

$$\mathbf{uv}^T = u_1v_1 + u_2v_2 ... u_nv_n = \sum_{i=1}^{n} u_i * v_i$$

Let's try and work this out on the board...

Dot product: find all combinations

$$\textit{Paul Wellstone} = \mathbf{u} = \begin{bmatrix}1 & 1 & 8 & 5\end{bmatrix}$$
$$\textit{Joe Lieberman} = \mathbf{v} = \begin{bmatrix}2 & 2 & 8 & 1\end{bmatrix}$$
$$\textit{Dianne Feinstein} = \mathbf{z} = \begin{bmatrix}3 & 1 & 1 & 2\end{bmatrix}$$

On your worksheet, find the dot product by hand for the other two combinations: Wellstone and Feinstein, Lieberman and Feinstein.

Working towards matrices

Recall, the vectors **u**, **v**, and **z** represent the number of bills each senator cosponsored with Hillary Clinton, Lincoln Chaffee, Joe Biden and Strom Thurmond.

$$\begin{array}{r} \begin{bmatrix} HC & LC & JB & ST \end{bmatrix} \\ \textit{Paul Wellstone} = \mathbf{u} = \begin{bmatrix} 1 & 1 & 8 & 5 \end{bmatrix} \\ \textit{Joe Lieberman} = \mathbf{v} = \begin{bmatrix} 2 & 2 & 8 & 1 \end{bmatrix} \\ \textit{Dianne Feinstein} = \mathbf{z} = \begin{bmatrix} 3 & 1 & 1 & 2 \end{bmatrix} \end{array}$$

Is there a more convenient way to express this information?

Working towards matrices

This is one of the motivations for our use of matrices!

- Matrices are made up of rows and columns.
- They are mathematical objects which, for our purposes, can be thought of as being constructed by "stacking" row vectors
- In our example, each row represents one of Paul Wellstone, Joe Lieberman or Dianne Feinstein
- Columns represent Hillary Clinton, Lincoln Chaffee, Joe Biden, and Strom Thurmond.
- Elements represent the number of bills cosponsored by the senators represented by the rows and columns

Working towards matrices

$$\mathbf{A} = \begin{bmatrix} HC & LC & JB & ST \\ 1 & 1 & 8 & 5 \\ 2 & 2 & 8 & 1 \\ 3 & 1 & 1 & 2 \end{bmatrix}$$

- Note that the rows of this matrix represent Paul Wellstone, Joe Lieberman, and Dianne Feinstein, respectively.
- $a_{i,j}$ refers to the element of matrix **A** corresponding with row i and column j.
- For example, $a_{1,4} = 5$
- For the rest of the lecture, we will treat this as a 3×4 matrix (the first row are labels just expositional purposes).

Summing up the vectors section

1. We presented vectors as a useful way of storing information
2. We saw how we can use addition/subtraction and multiplication on vectors
3. Importantly, we saw the “dot product” which the way we will be multipliying vectors
4. We saw that matrices can be constructed by “stacking” vectors

Matrices

Outline: moving from vectors to matrices

1. Types of matrices
   - Square matrices, symmetric matrices, the identity matrix
2. Matrix algebra
   - Addition/subtraction and scalar multiplication
   - Matrix multiplication

Square matrices

- If we denote matrix dimensions as $m \times n$:
  - Square: $m = n$
  - Rectangular: $m \neq n$
  - Let's visualize this on the board...
- Typical square matrices in social science: matrices to summarize pairwise measures (e.g., our co-sponsorship data; a correlation matrix summarizing correlations between any two variables in a dataset; etc.)
- Typical rectangular matrices in social science: rows = observations, columns = predictor variables (unless your number of observations happens to equal number of covariates)

Symmetry

A common use of a matrix in the social sciences is to represent social network data.

The set up?

- We have information about which senators are friends and which aren't
- Senators can either be friends or not, but we can't measure the strength of their friendship
- In this case we can summarize information about the social network of senators by constructing a matrix made up of 0s and 1s where a 1 indicates that two senators are friends and 0s indicates that they aren't
  - We do this by constructing a matrix where row 1 refers to the same senator as column 1, column 2 refers to the same senator as column 2, and so forth...

Symmetry

	Hillary Clinton	*Lincoln Chaffee*	*Joe Biden*	*Strom Thurmond*
Hillary Clinton	0	1	0	1
Lincoln Chaffee	1	0	1	0
Joe Biden	0	1	0	0
Strom Thurmond	1	0	0	0

- This matrix is **symmetric**
- By definition, this means that $a_{i,j} = a_{j,i}$ for all i, j.
- Informally, this means that if we were to split the matrix in two along the diagonal, the two halves are mirror images.
- Importantly, for later down the line, it also means that the $\mathbf{A} = \mathbf{A}^T$
- **Question:** Is this matrix square? Why or why not?

The identity matrix

- This is a special matrix that comes up in a lot of statistical theory. We won't delve too much into the details now; it suffices to know what this looks like.
- The identity matrix has elements $a_{ij} = 0$ for all $i \neq j$ and $a_{ij} = 1$ for all $i = j$
- More formally, the notation for the identity matrix is $\mathbf{I}_n$
- In this context, $\mathbf{I}_4$:

	Hillary Clinton	*Lincoln Chaffee*	*Joe Biden*	*Strom Thurmond*
Hillary Clinton	1	0	0	0
Lincoln Chaffee	0	1	0	0
Joe Biden	0	0	1	0
Strom Thurmond	0	0	0	1

Moving onto Matrix Algebra

- So far we have become a bit more familiar with what a matrix looks like and how we define their dimensions
- Now, like vectors, we will see how we can apply mathematical operations to these objects

Matrix Algebra: Overview

- Basic operations: addition/subtraction, scalar multiplication
- More complex operations: multiplication, transpose
- For each, we'll cover:
  - Motivation/preview of applications
  - Mechanics: what counts as conformable matrices for the purposes of the operation and how to perform

Addition/subtraction

- **Thing to remember**: For addition/subtraction, the two matrices have to have the same dimensions. Otherwise the operation is not possible.
- If you have two matrices with the same dimensions, addition/subtraction happens by just applying the operation to element-by-element.

$$\mathbf{A} = \begin{bmatrix} 0 & 3 & 9 \\ 0 & 0 & 6 \\ 9 & 3 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 2 & 12 \\ 0 & 0 & 3 \\ 7 & 5 & 0 \end{bmatrix}$$

Addition/subtraction

- **How to do**: element by element addition/subtraction

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} 0 & 3-2 & 9-12 \\ 0 & 0 & 6-3 \\ 9-7 & 3-5 & 0 \end{bmatrix}$$

Multiplying by a scalar

- Similar motivation as in vector case: can simultaneously rescale all the elements by some constant
- This works for a matrix of <u>any dimension</u>
- $$\mathbf{A} = \begin{bmatrix} 0 & 3 & 9 \\ 0 & 0 & 6 \\ 9 & 3 & 0 \end{bmatrix}, \left(\frac{1}{5}\right) \cdot \mathbf{A} = \begin{bmatrix} 0 & \frac{3}{5} & \frac{9}{5} \\ 0 & 0 & \frac{6}{5} \\ \frac{9}{5} & 1 & 0 \end{bmatrix}$$

  ,

Matrix multiplication

- The most important point to remember about matrix multiplication is that you need to be **really careful** about the dimensions of the matrices
- Even when implementing in R you need to pay attention to the dimensions of the matrices you are multiplying

Matrix multiplication

Why do we need to multiply matrices?

- **Example:** Linear regression (learn much more in Soc 500), where we can begin with typical way of writing the regression equation, and rewrite using matrices and vectors:

$$Y = \beta_0 + \beta_1 X_1 \dots \beta_n X_n$$

- How can we begin to write this using vectors, matrices, and matrix multiplication?

Matrix multiplication: checking conformability

Important: number of *columns* in first matrix must equal number of *rows* in second matrix, so:

- Matrix multiplication **does not** work:

$$\beta = \underset{4\times 1}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}, \ \mathbf{X} = \underset{4\times 4}{\begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \end{bmatrix}}$$

- Matrix multiplication **does** work:

$$\mathbf{X}\beta = \underset{4\times 4}{\begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \end{bmatrix}} \underset{4\times 1}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}$$

- Dimensions of resulting matrix: same number of *rows* as first matrix and same number of *columns* as the second matrix (in this case, 4×1)

Matrix multiplication: checking dimensions

- The easiest way to think through whether or not we can multiply two matrices is to write out the dimensions of each in the correct order (**Note:** the *order* that matrices are multiplied matters).
  - i.e. $\mathbf{AB} \neq \mathbf{BA}$
- If the inside numbers match, matrix multiplication will work
- The resulting matrix will be of size denoted by the outside numbers.
- For example, suppose $\mathbf{A}$ is 4×3 and $\mathbf{B}$ is 3×2.
- $4 \times 3 \quad 3 \times 2$
- Let's try a quick example on the board...

Now that we have conformable parts – how to multiply?

- Another example: here, result will be a 4×2 matrix:

$$\mathbf{AB} = \underset{4\times 3}{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}}$$

- Formally, we take the dot product of each row vector from the first matrix and column vector from the second matrix (red = changes from row to row of results):

$$\mathbf{AB} = \underset{4\times 2}{\begin{bmatrix} (a_{11} \quad a_{12} \quad a_{13}) \bullet (b_{11} \quad b_{21} \quad b_{31}) & (a_{11} \quad a_{12} \quad a_{13}) \bullet (b_{12} \quad b_{22} \quad b_{32}) \\ (a_{21} \quad a_{22} \quad a_{23}) \bullet (b_{21} \quad b_{22} \quad b_{32}) & (a_{21} \quad a_{22} \quad a_{23}) \bullet (b_{12} \quad b_{22} \quad b_{32}) \\ (a_{31} \quad a_{32} \quad a_{33}) \bullet (b_{21} \quad b_{22} \quad b_{32}) & (a_{31} \quad a_{32} \quad a_{33}) \bullet (b_{12} \quad b_{22} \quad b_{32}) \\ (a_{41} \quad a_{42} \quad a_{43}) \bullet (b_{21} \quad b_{22} \quad b_{32}) & (a_{41} \quad a_{42} \quad a_{43}) \bullet (b_{12} \quad b_{22} \quad b_{32}) \end{bmatrix}}$$

- Informally, we can 1) draw out a shell matrix with correct dimensions for results; 2) circle rows in first, columns in second and proceed

On your worksheet, complete the following questions:

$$\mathbf{Y} = \begin{bmatrix} 3 & 1 & -2 \\ 6 & 3 & 4 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 4 & 2 \\ 3 & 0 \\ 1 & 2 \end{bmatrix}$$

1. Write out dimensions of each
2. Arrange multiplication in a way that makes matrices conformable to multiply *and* that results in a 3×3 matrix
3. Arrange multiplication in a way that makes matrices conformable to multiply *and* that results in a 2×2 matrix. Then multiply by hand.
4. Multiply by hand

Matrix multiplication: some properties

Note: these properties assume the matrices are conformable so that should always be your first step! Full list is on Gill page 112.

- Associative: $(XY)Z = X(YZ)$
- Distributive for addition: $(X + Y)Z = XZ + YZ$
- What's missing from above?
  - **NOT COMMUTATIVE FOR MULTIPLICATION**: except in certain cases, $XY \neq YX$
  - Rare case where multiplication *is* commutative: identity matrix $\times$ any other (conformable) matrix. Example:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}$$

Summing up

1. We walked through what a vector is, how they can be used to store information, and how to apply mathematical operations (i.e addition/subtraction to them)
2. We then saw how we can think of matrices as "stacked" vectors
3. We looked at different types of matrices
4. We then saw how we can apply the same operations to matrices, allbeit with different rules
5. We burned into our brains the importance of **checking the dimensions of vectors and matrices before multiplying!**