Review from Bootcamp: Linear Algebra

D. Alex Hughes

October 27, 2014

Properties of Estimators

1. Unbiasedness:

$$E(\hat{\theta}) - \theta = 0$$

2. Asymptotic Unbiasedness:

$$\lim_{n\to\infty} P(|E[\hat{\theta}] - \theta| > \epsilon) \to 0; \forall\epsilon > 0$$

3. Efficiency

$$\frac{1}{MSE} = \frac{1}{E[(\hat{\theta} - \theta)^2]}$$

4. Consistency

$$\lim_{n\to\infty} P(|\hat{\theta} - \theta| > \epsilon) \to 0 \forall\epsilon > 0$$

Linear algebra

- Motivation: Linear algebra or matrix algebra avoids the mess and lets us solve for things we care about quickly, cleanly and easily.
- This is no different than algebra. Consider the difference in the following formulas for the mean:

$$\bar{x} = \frac{x_1 + x_2}{n}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3}{n}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4}{n}$$

$$\bar{x} = \frac{\sum x_i}{n}$$

Linear algebra

- Similarly, matrix algebra is a form of notation that cleans up the mess when working with more complex formulas. So suspend disbelief and concern, and treat this as a new language you are learning.
- Think of this as algebra on steroids.

Motivation

Why are we studying matrix algebra?

- We will use matrix algebra to derive the least squares estimator
- Matrices are an intuitive way to think about data. We have a set of observations (perhaps individuals) on the row, and observe many different characteristics (such as race, gender, PID, etc.) corresponding to columns
- Matrices are useful for solving systems of equations, like multiple regression
- Notation is much more compact and concise

Definition of Matrices and Vectors

Definition

A matrix is simply an arrangement of numbers in rectangular form.

Generally, a $(j \times k)$ matrix $\mathbf{A}$ can be written as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jk} \end{bmatrix}$$

Note that there are j rows and k columns, defining the *dimensionality* (order) of the matrix. Also note that the elements are double sub-scripted, with the row number first, and the column number second. In general terms, the $\mathbf{A}$ above is of order (j, k).

Examples

Example

$$\mathbf{W} = \begin{bmatrix} 1 & 3 \\ 2 & -6 \end{bmatrix}$$

is of order $(2, 2)$. This is also called a square matrix.

There are also rectangular matrices ($j \neq k$), such as:

Example

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 4 \\ 1 & 3 \\ 1 & -2 \\ 0 & 3 \end{bmatrix}$$

which is of order (4,2).

Notation

- Matrices are usually written using capital, bold-faced Roman or Greek letters. Roman is typically data, and Greek is typically parameters. This is *not* universal.

Vector

Definition

Vectors are matrices that have either one row or one column. is the same as a scalar – a regular number.

Row vectors have a single row and multiple columns.

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

Column vectors are those that have a single column and multiple rows.

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

Operations on Matrices

Addition and Subtraction
Scalar addition is simply:

$$m + n = 2 + 5 = 7$$

Addition is similarly defined for matrices.
If matrices or vectors are of the same order, then they can be added. One performs the addition element by element.

Addition

$\mathbf{A} + \mathbf{B} = \mathbf{C}$:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

Subtraction

$\mathbf{A} - \mathbf{B} = \mathbf{D}$:

$$\begin{bmatrix} 1 & 4 & -2 \\ 5 & -3 & 3 \end{bmatrix} - \begin{bmatrix} -3 & 2 & 8 \\ 2 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 4 & 2 & -10 \\ 3 & -5 & 6 \end{bmatrix}$$

Properties of Matrix Addition

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Matrix addition is commutative.
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$. Matrix addition is associative.

Transpose

Definition

To **transpose a matrix** is to exchange order subscripts. An order (j, k) matrix becomes an order (k, j) matrix.

Transposition is denoted by placing a prime after a matrix or by placing a superscript T.

$$\mathbf{Q} = \begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \\ q_{3,1} & q_{3,2} \end{bmatrix} \quad \mathbf{Q}' = \begin{bmatrix} q_{1,1} & q_{2,1} & q_{3,1} \\ q_{1,2} & q_{2,2} & q_{3,2} \end{bmatrix}$$

Note that the subscripts in the transpose remain the same, they are just exchanged.

Example

Example

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

Some Definitions

There are a few results regarding transposition that are important to remember:

- An order (j, j) matrix $\mathbf{A}$ is said to be **symmetric** iff $\mathbf{A} = \mathbf{A}'$.

$$\mathbf{W} = \begin{bmatrix} 1 & .2 & -.5 \\ .2 & 1 & .4 \\ -.5 & .4 & 1 \end{bmatrix} \quad \mathbf{W}' = \begin{bmatrix} 1 & .2 & -.5 \\ .2 & 1 & .4 \\ -.5 & .4 & 1 \end{bmatrix}$$

- $(\mathbf{A}')' = \mathbf{A}$
- For a scalar k, $(k\mathbf{A})' = k\mathbf{A}'$.
- For two matrices of the same order, the transpose of the sum is equal to the sum of the transposes. $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

Matrices and Multiplication

Scalar times a matrix. In words, a scalar α times a matrix **A** equals the scalar times each element of **A**. Thus,

$$\alpha\mathbf{A} = \alpha \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} \\ \alpha a_{2,1} & \alpha a_{2,2} \end{bmatrix}$$

So, for:

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 2 \\ 6 & 8 & 10 \end{bmatrix} \quad \frac{1}{2}\mathbf{A} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & 4 & 5 \end{bmatrix}$$

Matrices and Multiplication

Definition

Given **A** of order (m, n) and **B** of order (n, r), then the product $\mathbf{AB} = \mathbf{C}$ is the order (m, r) matrix whose entries are defined by:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

where $i = 1, \ldots, m$ and $j = 1, \ldots, r$ and $k_1 = n_2$

Matrices and Multiplication

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix}$$

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} -2 \cdot 3 + 1 \cdot 2 + 3 \cdot 1 & -2 \cdot (-2) + 1 \cdot 4 + 3 \cdot (-3) \\ 4 \cdot 3 + 1 \cdot 2 + 6 \cdot 1 & 4 \cdot (-2) + 1 \cdot 4 + 6 \cdot (-3) \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 \\ 20 & -22 \end{bmatrix} \end{aligned}$$

Matrices and Multiplication

Is multiplication of matrices commutative?

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix}$$

$$\mathbf{BA} = \begin{bmatrix} -14 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{bmatrix}$$

No: Multiplication of matrices is not commutative. In other words: $\mathbf{AB} \neq \mathbf{BA}$.

Matrices and Multiplication

Important results

- Matrix multiplication is not commutative: $\mathbf{AB} \neq \mathbf{BA}$.
- Matrix multiplication is associative:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

- Matrix multiplication is distributive:

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

- The transpose of a product can be written as

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Vectors and Multiplication

Inner product of vectors

$$\mathbf{e}'\mathbf{e} = \begin{bmatrix} e_1 & e_2 & \cdots & e_N \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

$$\mathbf{e}'\mathbf{e} = e_1 e_1 + e_2 e_2 + \cdots + e_N e_N = \sum_{i=1}^{N} e_i^2$$

Alt: outer product

Other Useful Vector Products

Let $\mathbf{i}$ denote an order $(N, 1)$ vector of ones, and $\mathbf{x}$ denote an order $(N, 1)$ vector of data.

$$\mathbf{i}'\mathbf{x} = (x_1 + x_2 + \cdots x_N) = \sum x_i$$

From this, it follows that:

$$\frac{1}{N}\mathbf{i}'\mathbf{x} = \frac{1}{N}\sum x_i = \bar{x}$$

Cross Product

$$A \times B = \hat{n}|A||B|cos(\theta)$$

- $\hat{n}$: perpendicular unit vector
- $|A|$: Length of A
- θ: angle between A & B

Trace

Sum of the diagonal elements of a square matrix.

$$\mathbf{A} = \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{matrix}$$

$$tr(\mathbf{A}) = \sum a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

Special Matrices and Their Properties

When performing scalar algebra, we know that $x \cdot 1 = x$, which is known as the identity relationship.

There is a similar relationship in matrix algebra: $\mathbf{AI} = \mathbf{A}$.

What is $\mathbf{I}$?

It can be shown that $\mathbf{I}$ is a diagonal, square matrix with ones on the main diagonal, and zeros on the off diagonal.

For example, the order three identity matrix is:

$$\mathbf{I_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Special Matrices and Their Properties

Notice that $\mathbf{I}$ is oftentimes subscripted to denote its dimensionality. Here is an example of the use of an identity matrix:

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

Special Matrices and Their Properties

One of the nice properties of the identity matrix is that it is commutative with respect to multiplication. That is,

$$\mathbf{AIB} = \mathbf{IAB} = \mathbf{ABI} = \mathbf{AB}$$

An identity in scalar algebra is $x + 0 = x$.
This generalizes to matrix algebra, with the definition of the null matrix, which is simply a matrix of zeros, denoted $\mathbf{0_{j,k}}$.
Here is an example:

$$\mathbf{A} + \mathbf{0_{2,2}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \mathbf{A}$$

Matrix Inversion

Definition:

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

This is roughly akin to division in non-matrix algebra. Actually calculating the inverse of a matrix takes several steps and has several prerequisites.

Matrix Inversion

General solution for a square matrix **A**:

$$\mathbf{A}^{-1} = \frac{1}{|A|} adj A$$

So we need to figure out $\frac{1}{|A|}$ and $adj(A)$.

Matrix Inversion

The first of these, $|\mathbf{A}|$, is called "the determinant". There's lots to learn about determinants, but we'll stick to the basics.

Most importantly, if the determinant is NOT zero, a square matrix *is* invertible.

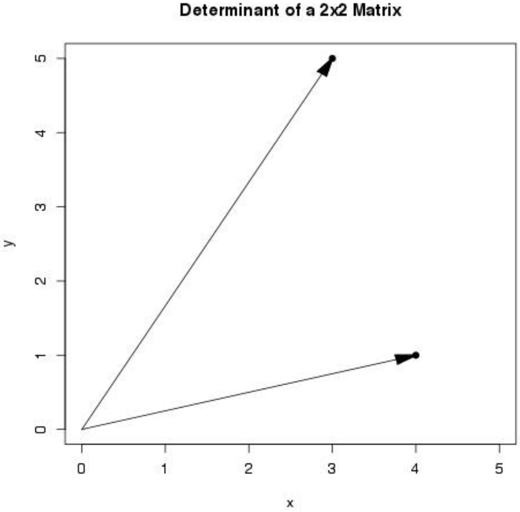
The determinant is a scalar, that is, a single number, like 5.

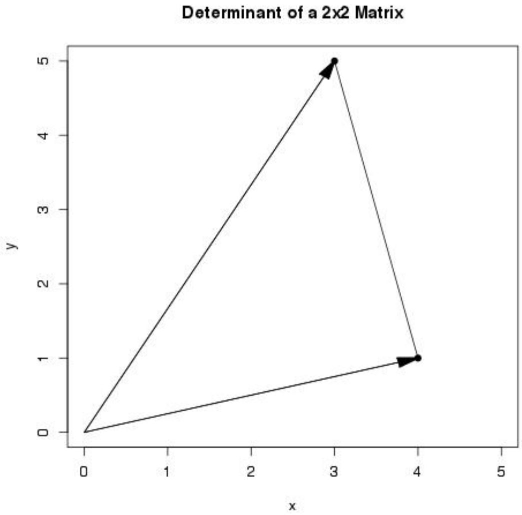
For a two by two matrix, the determinant is:

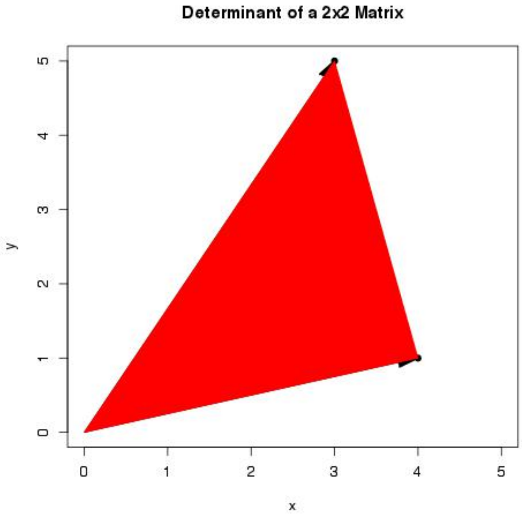
$$|\mathbf{A}| = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

So you multiply the corners and subtract one product from the other.

Graphical Intuition for Determinant of a 2x2 Matrix




Determinant of a 2x2 Matrix
y
0
1
2
3
4
5
x
0
1
2
3
4
5


Determinant of a 2x2 Matrix
y
0
1
2
3
4
5
x
0
1
2
3
4
5

Graphical Linear Algebra

- For a 2x2 matrix, the determinant is two times the area of the triangle defined by the row vectors.
- Think about this for a matrix like:

$$|\mathbf{A}| = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

Or...

$$|\mathbf{A}| = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}$$

- Most linear algebra functions can be represented graphically.
- Ask me for citations if you want a book that illustrates all these.

Calculating Determinants

With three by three matrices, the determinant is still quite manageable:

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

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

Graphically, this means adding the diagonal products from left to right, and subtracting the diagonal products from right to left. (example)

More Determinants

For bigger matrices, we have to use alternative methods, typically the *Laplace Expansion*. Basically, we break the matrix into sub-matricies, and calculate determinants of these submatricies, then combine our results. Steps:

1. Pick a row or column to work with.
2. For each element in that row, calculate the subdeterminant, also called the minor.
3. Multiply each element by its subdeterminant, determine signs, and add.

Higher Order Determinants - Example

$$\begin{vmatrix} 5 & 6 & 1 \\ 2 & 3 & 0 \\ 7 & -3 & 0 \end{vmatrix} = 5 * \begin{vmatrix} 3 & 0 \\ -3 & 0 \end{vmatrix} - 6 * \begin{vmatrix} 2 & 0 \\ 7 & 0 \end{vmatrix} + 1 * \begin{vmatrix} 2 & 3 \\ 7 & -3 \end{vmatrix}$$

Properties of Determinants

- $|A| = |A'|$
- Interchanging any two rows or columns will alter sign but not value of determinant.
- Multiplication of one row by k will change $|A|$ to $k|A|$.
- Addition/subtraction of a multiple of any row to another row will leave the value of the determ unaltered (works for col too).
- If one row or columns is a multiple of another, the value of the determinant will be zero - matrix is *singular*

Tricks for Determinants

Pick a good row or column.

$$\begin{vmatrix} 5 & 6 & 1 \\ 2 & 3 & 0 \\ 3 & -3 & 0 \end{vmatrix} = 5 * \begin{vmatrix} 3 & 0 \\ -3 & 0 \end{vmatrix} - 6 * \begin{vmatrix} 2 & 0 \\ 7 & 0 \end{vmatrix} + 1 * \begin{vmatrix} 2 & 3 \\ 7 & -3 \end{vmatrix}$$

$$\begin{vmatrix} 5 & 6 & 1 \\ 2 & 3 & 0 \\ 7 & -3 & 0 \end{vmatrix} = 1 * \begin{vmatrix} 2 & 3 \\ 7 & -3 \end{vmatrix} - 0 * \begin{vmatrix} 5 & 6 \\ 7 & -3 \end{vmatrix} + 0 * \begin{vmatrix} 5 & 6 \\ 2 & 3 \end{vmatrix}$$

Tricks for Determinants

Manipulate rows if possible

$$\begin{vmatrix} 5 & 6 & 1 \\ 2 & 4 & 7 \\ 3 & 6 & 11 \end{vmatrix} = \begin{vmatrix} 5 & 6-10 & 1 \\ 2 & 4-4 & 7 \\ 3 & 6-6 & 11 \end{vmatrix} = \begin{vmatrix} 5 & -4 & 1 \\ 2 & 0 & 7 \\ 3 & 0 & 11 \end{vmatrix}$$

Back to inverting a matrix

$$A^{-1} = \frac{1}{|A|} adj(A)$$

(adj means *adjugate*)
where

$$adj(A) = \begin{bmatrix} |C_{11}| & |C_{12}| & \cdots & |C_{1n}| \\ |C_{21}| & |C_{22}| & \cdots & |C_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |C_{n1}| & |C_{n2}| & \cdots & |C_{nn}| \end{bmatrix}'$$

Cofactors

C_{ij} is a *matrix cofactor* - the determinant of the matrix when excluding row i and column j, and $adj(A)$ is the transpose of the matrix of cofactors. The determinant of the cofactor submatrix is multiplied by 1 when $i + j$ is even, and by -1 when $i + j$ is odd.

Easy Inversion - 2X2

$$A^{-1} = \frac{1}{|A|} adj(A)$$

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$X^{-1} = \frac{1}{|3-2|} * \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$