Matrix algebra basics

Dave Clark

January 3, 2024

Matrix basics

Matrix notation

For our purposes, matrix algebra is useful for (at least) three reasons:

- Simplifies mathematical expressions.
- Has a clear, visual relationship to the data.
- Provides intuitive ways to understand elements of the model.

Matrices

A single number is known as a *scalar*.

A *matrix* is a rectangular or square array of numbers arranged in rows and columns.

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

where each *element* of the matrix is written as $a_{row,column}$. Matrices are denoted by capital, bold faced letters, $\mathbf{A}$.

Dimensions

The *dimensions* of a matrix are the number of rows (n) and columns (m) it contains. The *dimensions* of matrices are always read n by m, row by column. We would say that matrix $\mathbf{A}$ is of order (n, m).

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 8 \\ 4 & 5 & 3 \\ 8 & 3 & 2 \end{bmatrix}$$

$\mathbf{A}$ is of order (3,3) and is a *square matrix.* If matrix $\mathbf{A}$ is of order (1,1), then it is a scalar.

Symmetric matrices

- $\mathbf{A}$ is a *square matrix* if $n = m$. This matrix is also *symmetric.*
- A *symmetric matrix* is one where each element $a_{n,m}$ is equal to its opposite element, $a_{m,n}$. In other words, a matrix is symmetric if $a_{n,m} = a_{n,m}$, $\forall$ n and m.
- All symmetric matrices are square, but not all square matrices are symmetric. A correlation matrix is and example of a symmetric matrix.

Diagonal matrices have zeros for all *off-diagonal* elements.

Diagonal

Diagonal matrices only have non-zero elements on the *main diagonal* (from upper left to lower right). That is, $a_{n,m} = 0$ if $n \neq m$.

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Scalar

$\mathbf{B}$ is a special kind of diagonal matrix known as a *scalar matrix* because all of the elements on the main diagonal are equal to each other; $a_{n,m} = k$ if $n = m$, else, zero.

$$\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Identity

$\mathbf{C}$ is a special kind of scalar matrix called the *identity matrix*; the diagonal elements of this matrix are all equal to 1 ($a_{n,m} = 1$ if $n = m$, else, zero). This is useful in some matrix manipulations we'll talk about later; it is denoted $\mathbf{I}$.

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

Rectangular matrices

A *rectangular matrix* is one where $n \neq m$.

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 3 & 5 \\ 7 & 2 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 7 \\ 3 & 5 & 2 \end{bmatrix}$$

In the first, case, $\mathbf{A}$ is of order (3,2); in the second, $\mathbf{A}$ is of order (2,3).

Vectors

A *vector* is a special kind of matrix wherein one of its dimensions is 1; the matrix has either one row or one column. Vectors are denoted by lower case, bold faced letters, **a** and may be *row* or *column* vectors.

$$= \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} \qquad \mathbf{a} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}$$

Parameter matrices (e.g. β) follow the same notation except that they are indicated by Greek rather than Roman letters.

Operations

Transposition

The *transpose* of a matrix $\mathbf{A}$ is the matrix flipped on its side such that its rows become columns and columns become rows; the transpose of $\mathbf{A}$ is denoted $\mathbf{A}'$ and is referred to as "$\mathbf{A}$ transpose."

$$\mathbf{X} = \begin{bmatrix} 1 & 3 \\ 3 & 5 \\ 7 & 2 \end{bmatrix} \qquad \mathbf{X}' = \begin{bmatrix} 1 & 3 & 7 \\ 3 & 5 & 2 \end{bmatrix}$$

so $\mathbf{X}$, a (3,2) matrix, becomes $\mathbf{X}'$, a (2,3) matrix.

Transposition

If $\mathbf{A}$ is symmetric, then $\mathbf{A} = \mathbf{A}'$ – take a look at the following symmetric matrix and you'll see why:

$$\mathbf{X} = \begin{bmatrix} 2 & 4 & 8 \\ 4 & 5 & 3 \\ 8 & 3 & 2 \end{bmatrix} \qquad \mathbf{X}' = \begin{bmatrix} 2 & 4 & 8 \\ 4 & 5 & 3 \\ 8 & 3 & 2 \end{bmatrix}$$

Transposition

Also note that the transpose of a transposed matrix results in the original matrix: $(\mathbf{X}')' = \mathbf{X}$.

Transposing a row vector results in a column vector and vice versa:

$$\mathbf{x} = \begin{bmatrix} -1 \\ -9 \\ 4 \\ 16 \end{bmatrix} \qquad \mathbf{x}' = \begin{bmatrix} -1 & -9 & 4 & 16 \end{bmatrix}$$

Trace of a Matrix

The *trace* of a matrix is the sum of the diagonal elements of a square matrix, and is denoted $tr(\mathbf{A}) = a_{11} + a_{22} + a_{33} \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}$

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 8 \\ 4 & 5 & 3 \\ 8 & 3 & 2 \end{bmatrix}$$

such that

$tr(\mathbf{A}) = 2 + 5 + 2 = 9$

Addition & Subtraction

- Addition and subtraction of matrices is analogous to the same scalar operations, but depend on the *conformatibility* of the matrices to be added or subtracted.
- Two matrices are *conformable for addition or subtraction* iff they are of the same order.
- Note that conformability means something different for addition/subtraction than for multiplication.

Addition & Subtraction

Consider the following (2,2) matrices and the problem $\mathbf{A} + \mathbf{B} = \mathbf{C}$:

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

Addition and subtraction are only possible among matrices that share the same order; otherwise, they are nonconformable.

Addition & Subtraction

Consider a second example of addition, and note that subtraction would follow directly:

$$\begin{bmatrix} 2 & 4 & 8 \\ 4 & 5 & 3 \\ 8 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 5 & 7 \\ 3 & 7 & 1 \\ 2 & 4 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 15 \\ 7 & 12 & 4 \\ 10 & 7 & 11 \end{bmatrix}$$

Matrix addition adheres to the commutative and associative properties such that:

$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (Commutative), and $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ (Associative).

Multiplication

Let's begin by multiplying a scalar and a matrix - the product is a matrix whose elements are the scalar multiplied by each element of the original matrix, so:

$$\beta \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} \beta x_{11} & \beta x_{12} \\ \beta x_{21} & \beta x_{22} \end{bmatrix}$$

Multiplication

- In order to multiply two matrices (or vectors), they must be *conformable for multiplication.*
- Two matrices are *conformable for multiplication* iff the number of columns in the first matrix is equal to the number of rows in the second matrix. That is, $\mathbf{A_{i,j}}$ and $\mathbf{B_{j,k}}$.
- An easy way to think of this is to write the dimensions of the two matrices, (i,j),(j,k) - the inner dimensions are the same, so the matrix is conformable for multiplication - moreover, the outer dimensions (i,k) give the dimension of the product matrix.

Multiplication - inner product

- To illustrate multiplication, let's start with a column vector, $\mathbf{e}$ whose order is (N,1) and take its *inner product*.
- The *inner product* of a column vector is the transpose of the column vector (thus, a row vector) post-multiplied by the column vector, so $\mathbf{e}'\mathbf{e}$. The inner product is a scalar.

Multiplication - inner product

When we transpose the column vector $\mathbf{e}$ of (N,1) order, we get a row vector $\mathbf{e}'$ of (1,N) order. Inner dimensions match, so they are conformable.

$$\mathbf{e}' = \begin{bmatrix} -1 & -9 & 4 & 16 & -10 \end{bmatrix} \qquad \mathbf{e} = \begin{bmatrix} -1 \\ -9 \\ 4 \\ 16 \\ -10 \end{bmatrix}$$

$$= (-1 \cdot -1) + (-9 \cdot -9) + (4 \cdot 4) + (16 \cdot 16) + (-10 \cdot -10) = 454$$

Least Squares foreshadowing

Note the inner product of a column vector is a scalar; the inner product of $\mathbf{e}$, is the *sum of the squares of* $\mathbf{e}$.

$$\mathbf{e}'\mathbf{e} = e_1e_1 + e_2e_2 + \dots e_Ne_N = \sum_{i=1}^{N} e_i^2$$

Multiplication - outer product

The *outer product* of a column vector is the transpose of the column vector (thus, a row vector) pre-multiplied by the column vector, so $\mathbf{e}\mathbf{e}'$. The outer product is an (N,N) matrix.

$$\mathbf{e} = \begin{bmatrix} -1 \\ -9 \\ 4 \\ 16 \\ -10 \end{bmatrix} \mathbf{e}' = \begin{bmatrix} -1 & -9 & 4 & 16 & -10 \end{bmatrix}$$

Let's multiply $\mathbf{ee}'$, a column vector of (5,1) by a row vector of (1,5); we'll obtain a square matrix of (5,5).

Least Squares foreshadowing

Let $\mathbf{e}$ represent the residuals or errors in our regression. The outer product

$$\mathbf{ee}' = \begin{bmatrix} 1 & 9 & -4 & -16 & 10 \\ 9 & 81 & -36 & -144 & 90 \\ -4 & -36 & 16 & 64 & -40 \\ -16 & -144 & 64 & 256 & -160 \\ 10 & 90 & -40 & -160 & 100 \end{bmatrix}$$

is the *variance-covariance matrix of* $\mathbf{e}$. The squares on the main diagonal (which sums to the sum of squares) and the symmetry of the off-diagonal elements.

Inverting Matrices

Inverting Matrices

- You'll have noticed that division has been conspicuously absent so far. In scalar algebra, we sometimes represent division in fractions, $\frac{1}{2}$.
- Another way to represent the same quantity is 2^{-1}, or the inverse of 2. Of course, in scalar algebra, a number multiplied by its inverse equals 1, e.g., $2 \cdot 2^{-1} = 1$ or $2 \cdot \frac{1}{2} = 1$. So, if we are given $4x = 1$ and want to solve for x what we really want to know is "what is the inverse of 4 such that $4 \cdot 4^{-1} = 1$?" We consider matrices in a similar manner.

Inverting Square Matrices

Imagine a square matrix, $\mathbf{X}$ and its inverse, denoted $\mathbf{X^{-1}}$ (read as "X inverse"). Analogous to scalar algebra, a matrix multiplied by its inverse is equal to the identity matrix, $\mathbf{I}$.

So in order to find $\mathbf{X^{-1}}$, we must find the matrix that, when multiplied by $\mathbf{X}$, produces $\mathbf{I}$. Thus, for a square matrix, its inverse (if it exists) is such that

$$\mathbf{XX^{-1}} = \mathbf{X^{-1}X} = \mathbf{I}$$

Notice the commutative property at work here - this is because $\mathbf{X}$ is square, such that $n = m$, so $\mathbf{X}$ and $\mathbf{X^{-1}}$ have the same dimensions.

Determinant

An important characteristic of square matrices is the *determinant*. The determinant, denoted $|X|$, is a scalar; every square matrix has one. We evaluate the determinant by examining the cross-products of the matrice's elements. This is simple in the 2x2 matrix, less so in larger matrices.

$$|\mathbf{X}| = \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}$$

Determinant

In the 2x2 matrix, the determinant is the cross-product of the main diagonals minus the cross-product of the off-diagonals.

$$|\mathbf{X_1}| = \begin{vmatrix} 2 & 4 \\ 6 & 3 \end{vmatrix} = 2 \cdot 3 - 4 \cdot 6 = -18$$

The determinant is -18; $|\mathbf{X_1}| = -18$.

Determinant

This matrix has a determinant of zero - this is an important case.

$$|\mathbf{X_2}| = \begin{vmatrix} 3 & 6 \\ 2 & 4 \end{vmatrix} = 3 \cdot 4 - 6 \cdot 2 = 0$$

A matrix with determinant zero is *singular*. A *singluar matrix* has no inverse.

Singular matrices

There are some important conditions that will produce a determinant of zero and thus a singular matrix:

- If all the elements of any row or column of a matrix are equal to zero, then the determinant is zero.
- If two rows or columns of a matrix are identical, the determinant is zero.
- If a row or column of a matrix is a multiple of another row or column, the determinant is zero; if one row or column is a linear combination of other rows or columns, the determinant is zero.

Least Squares Foreshadowing

For the linear model in least squares, this is important because computing estimates of β requires inverting a matrix. Let's derive β in matrix notation to see how:

The estimated model is:

$$\mathbf{y} = \mathbf{X}\hat{\beta} + \hat{}$$

Minimizing $\hat{}'\hat{}$ (skipping the math for now), we get

$$(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'\mathbf{y}$$

Foreshadowing Least Squares

The matrix $(\mathbf{X}'\mathbf{X})$ gives the sums of squares (on the main diagonal) and the sums of cross products (in the off-diagonals) of all the $\mathbf{X}$ variables; the matrix is symmetric. Since β is the unknown vector in this equation, solve for β by dividing both sides by $(\mathbf{X}'\mathbf{X})$ - in matrix terms, we premultiply each side by $(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}$:

$$(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}(\mathbf{X}'\mathbf{X})\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y}$$

Foreshadowing Least Squares

As we know, $(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}(\mathbf{X}'\mathbf{X}) = \mathbf{I}$. So,

$$\mathbf{I}\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y}$$

or

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y}$$

Estimating $\hat{\beta}$ requires *inverting* $(\mathbf{X}'\mathbf{X})$ If the matrix $(\mathbf{X}'\mathbf{X})$ is singular, then $(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}$ does not exist, and we cannot estimate $\hat{\beta}$. Least Squares fails.

Foreshadowing Least Squares

Thinking specifically of the $\mathbf{X}$ variables in the model, any of these conditions produce perfect collinearity - estimation fails because the matrix can't invert.

- If all the elements of any row or column of $\mathbf{X}$ are equal to zero, then the determinant is zero.
- If two rows or columns of $\mathbf{X}$ are identical, the determinant is zero.
- If a row or column of $\mathbf{X}$ is a multiple of another row or column, the determinant is zero; if one row or column is a linear combination of other rows or columns, the determinant is zero.

Examples

Example 1

$\mathbf{X}$ below has a row of zeros; compute the determinant using the difference of cross-products.

$$|\mathbf{X}| = \begin{vmatrix} 0 & 0 \\ 19 & 12 \end{vmatrix} = 0 \cdot 12 - 0 \cdot 19 = 0$$

You can see in $\mathbf{X}$ that because all the cross-products involve multiplying by zero, the determinant will always be zero; this applies to larger matrices as well.

Example 2

$$|\mathbf{X}| = \begin{vmatrix} 5 & 7 \\ 5 & 7 \end{vmatrix} = 5 \cdot 7 - 5 \cdot 7 = 0$$

In $\mathbf{X}$ it's easy to see that the cross-products will always be equal to one another if the rows or columns are identical, and the difference between identical values is zero.

Example 3

$$|\mathbf{X}| = \begin{vmatrix} 64 & 48 \\ 16 & 12 \end{vmatrix} = 64 \cdot 12 - 48 \cdot 16 = 768 - 768 = 0$$

Finally $\mathbf{X}$ shows that if one row or column is a linear combination of other rows or columns, the determinant will be zero; in that matrix, the top row is 4 times the bottom row.

Inversion

The matrix we need to invert, $\mathbf{X}'\mathbf{X}$ is rectangular, so inversion is more complex. I'm going to illustrate the method of *cofactor expansion* - the purpose here is to give you an idea what software does every time you estimate an OLS model.

Minor of a matrix

The *minor* of a matrix is the determinant of a submatrix created by deleting the ith row and the jth column of the full matrix, where the minor is denoted by the element where the deleted rows intersect.

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

$$\text{so the minor of } a_{11} \text{ is } |\mathbf{M_{11}}| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22} \cdot a_{33} - a_{23} \cdot a_{32}$$

Cofactor Matrix

The *cofactor* of an element (c_{ij}) is the minor with a positive or negative sign depending on whether $i + j$ is odd (negative) or even (positive). This is given by:

$$\theta_{ij} = (-1)^{i+j} |\mathbf{M_{ij}}|$$

so, from the example finding the minor of a_{11}, i=1 and $j = 1$; their sum is 2 which is even, so the cofactor is $+a_{22} \cdot a_{33} - a_{23} \cdot a_{32}$. If we find the cofactor for every element, a_{ij} in a matrix and replace each element with its cofactor, the new matrix is called the *cofactor matrix of* **A**.

Cofactor Matrix

What we have so far is the signed matrix of minors:

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

Cofactor Expansion

Find the signed determinant of each 2x2 submatrix.

$${}_{\mathbf{A}} \begin{bmatrix} \begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 3 & 3 \\ 2 & 2 \end{vmatrix} \\ -\begin{vmatrix} 4 & 2 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 2 & 2 \end{vmatrix} \\ \begin{vmatrix} 4 & 2 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ 3 & 3 \end{vmatrix} \end{bmatrix} \quad {}_{\mathbf{A}} = \begin{bmatrix} 10 & -10 & 0 \\ -12 & 0 & 6 \\ -2 & 5 & -9 \end{bmatrix}$$

Find the Determinant

Using the cofactor matrix, ${}_{\mathbf{A}}$ we perform *cofactor expansion* in order to find the determinant, $|\mathbf{A}|$. Cofactor expansion is given by:

$$|\mathbf{A}| = \sum_{i,j=1}^{n} a_{ij}\Theta_{ij}$$

which means that we take two corresponding rows or columns from $\mathbf{A}$ and ${}_{\mathbf{A}}$ and sum the products of their elements - this gives us the determinant of $\mathbf{A}$.

Find the Determinant

So taking the first row from $\mathbf{A}$ (1,4,2) and from ${}_{\mathbf{A}}$ (10,-10,0) above, we compute

$$a_{11} \cdot \Theta_{A11} + a_{12} \cdot \Theta_{A12} + a_{13} \cdot \Theta_{A13}$$

$$1 \cdot 10 + 4 \cdot -10 + 2 \cdot 0 = -30$$

This is the determinant of $\mathbf{A}$, $|\mathbf{A}|$. Note that any corresponding rows or columns from $\mathbf{A}$ and $_{\mathbf{A}}$ will produce the same result.

Adjoint Matrix

If we transpose the cofactor matrix, (cof $\mathbf{A}'$), the new matrix is called the *adjoint matrix*, denoted adj$\mathbf{A}$=(cof $\mathbf{A}'$).

$$adj\mathbf{A} = {}'_{\mathbf{A}} = \begin{bmatrix} 10 & -12 & -2 \\ -10 & 0 & 5 \\ 0 & 6 & -9 \end{bmatrix}$$

Inverting the matrix

Believe it or not, this all leads us to inverting the matrix, provided it is nonsingular.

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

The inverse of A is equal to one over the determinant of A times the adjoint matrix of A. So we're pre-multiplying $adj\mathbf{A}$ by the (scalar) determinant, -30.

Inverse of Matrix A

$$\mathbf{A}^{-1} = -\frac{1}{30}\begin{bmatrix} 10 & -12 & -2 \\ -10 & 0 & 5 \\ 0 & 6 & -9 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{5} & \frac{1}{15} \\ \frac{1}{3} & 0 & -\frac{1}{6} \\ 0 & -\frac{1}{5} & \frac{3}{10} \end{bmatrix}$$

Checking our work

We can check our work to be sure we've inverted correctly by making sure that $\mathbf{A^{-1}A} = \mathbf{I_3}$; so,

$$-\frac{1}{30}\begin{bmatrix} 10 & -12 & -2 \\ -10 & 0 & 5 \\ 0 & 6 & -9 \end{bmatrix}\begin{bmatrix} 1 & 4 & 2 \\ 3 & 3 & 1 \\ 2 & 2 & 4 \end{bmatrix} = -\frac{1}{30}\begin{bmatrix} -30 & 0 & 0 \\ 0 & -30 & 0 \\ 0 & 0 & -30 \end{bmatrix}$$

The data matrix and the model

Connecting data matrices to OLS

Let's examine these matrices a little to get a feel for what the computation of $\hat{\beta}$ involves: $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

$$\mathbf{X} = \begin{bmatrix} 1 & X_{1,2} & X_{1,3} & \cdots & X_{1,k} \\ 1 & X_{2,2} & X_{2,3} & \cdots & X_{2,k} \\ 1 & X_{3,2} & X_{3,3} & \cdots & X_{3,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n,2} & X_{n,3} & \cdots & X_{n,k} \end{bmatrix}$$

The dimensions of $\mathbf{X}$ are nxk; the sample size by the number of regressors (including the constant).

Connecting data matrices to OLS

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} N & \sum X_{2,i} & \sum X_{3,i} & \cdots & \sum X_{k,i} \\ \sum X_{2,i} & \sum X_{2,2}^2 & \sum X_{2,i}X_{3,i} & \cdots & \sum X_{2,i}X_{k,i} \\ \sum X_{3,i} & \sum X_{3,i}X_{2,i} & \sum X_{3,i}^2 & \cdots & \sum X_{3,i}X_{k,i} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum X_{k,i} & \sum X_{k,i}X_{2,i} & \sum X_{k,i}X_{3,i} & \cdots & \sum X_{k,i}^2 \end{bmatrix}$$

In $\mathbf{X}'\mathbf{X}$, the main diagonal is the sums of squares and the offdiagonals are the cross-products.

Connecting data matrices to OLS

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} \sum Y_i \\ \sum X_{2,i}Y_i \\ \sum X_{3,i}Y_i \\ \vdots \\ \sum X_{k,i}Y_i \end{bmatrix}$$

When we compute $\hat{\beta}$, $\mathbf{X}'\mathbf{y}$ is the covariation of X and y, and we pre-multiply by the inverse of $(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}$ to control for the relationships among X_k.