Matrix Algebra

Today's Goals:

- Matrices, vectors, and scalars
- Basic matrix operations
- Linear combinations in matrix form

A *matrix* is a rectangular array of numbers with *n* rows and *m* columns. It is symbolized with a bold, upper case letter, and subscripted to indicate its *order*.

$$\mathbf{G}_{3,2} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{pmatrix}$$

Rows

Columns

$$G_{3,2} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{bmatrix}$$

The individual elements in a matrix are called *scalars*, subscripted to indicated their position in the matrix.

$$G_{3,2} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{bmatrix}$$

The columns in a matrix are called *vectors*, and are symbolized with lower case, bold letters. This matrix has two vectors ($\mathbf{g_1}$, $\mathbf{g_2}$), each containing three scalars.

$$\mathbf{g}_{3,1} = \begin{bmatrix} g_{11} \\ g_{21} \\ g_{31} \end{bmatrix}$$

This vector is also a 3 x 1 matrix. When displayed as a row, it is symbolized differently:

$$\mathbf{g}'_{1,3} = \begin{bmatrix} g_{11} & g_{21} & g_{31} \end{bmatrix}$$

This is called the *transpose* of the vector.

The standard data matrix . . .

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	. . .	V_K
P_1												
P_2												
P_3												
P_4												
P_5												
⋮												
P_N												

can be transposed . . .

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	. . .	P_N
V_1												
V_2												
V_3												
V_4												
V_5												
⋮												
V_K												

to change the object of measurement and shift the focus of the analysis

$$\mathbf{G}_{3,2} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \end{pmatrix}$$

This matrix is *rectangular*. When the number of rows and columns are equal, the matrix is *square*:

$$\mathbf{F}_{3,3} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$$

Data matrices are typically rectangular; correlation matrices and covariance matrices are always square.

Square matrices have some additional useful properties.

$$\mathbf{F}_{3,3} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$$

A square matrix has a *main diagonal*. The sum of the elements of the main diagonal is called the *trace* of the matrix.

$$\mathrm{trace}(F) = \sum_{i=1}^{r} f_{i,i}$$

If $f_{i,j} = f_{j,i}$ for all i and j, the matrix is *symmetric*.

$$\mathbf{F}_{3,3} = \begin{pmatrix} 1 & 3 & -4 \\ 3 & 11 & 7 \\ -4 & 7 & 2 \end{pmatrix}$$

F is a symmetric matrix with a trace of 14. Correlation matrices and covariance matrices are symmetric.

If all elements of a symmetric matrix except the main diagonal are zero, the matrix is a *diagonal* matrix:

$$\mathbf{F}_{3,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

F is a symmetric, diagonal matrix with a trace of 14.

Matrix Addition and Subtraction:

- Matrices of the same order can be added and subtracted.
- These operations take place element by element.

$$F_{3,2} = \begin{pmatrix} 1 & 3 \\ 3 & 11 \\ -4 & 7 \end{pmatrix} \qquad H_{3,2} = \begin{pmatrix} 4 & -1 \\ 6 & 2 \\ 12 & 8 \end{pmatrix}$$

$$F + H = K_{3,2} = \begin{pmatrix} 1+4 & 3+(-1) \\ 3+6 & 11+2 \\ -4+12 & 7+8 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 9 & 13 \\ 8 & 15 \end{pmatrix}$$

$$F_{3,2} = \begin{pmatrix} 1 & 3 \\ 3 & 11 \\ -4 & 7 \end{pmatrix} \qquad H_{3,2} = \begin{pmatrix} 4 & -1 \\ 6 & 2 \\ 12 & 8 \end{pmatrix}$$

$$F - H = K_{3,2} = \begin{pmatrix} 1-4 & 3-(-1) \\ 3-6 & 11-2 \\ -4-12 & 7-8 \end{pmatrix} = \begin{pmatrix} -3 & 4 \\ -3 & 9 \\ -16 & -1 \end{pmatrix}$$

Matrix addition is commutative and associative:

A + B = B + A

A + B + C = (A + B) + C = A + (B + C)

Matrix subtraction is distributive:

A – (B + C) = A – B – C

A – (B – C) = A – B + C

A matrix of zeros is called a *zero matrix* or a *null matrix*. It is used in solving equations:

$$\underset{\mathbf{X}}{\begin{pmatrix} 5 & 7 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}} + \underset{\mathbf{Y}}{\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \end{pmatrix}} = \underset{\mathbf{Z}}{\begin{pmatrix} 7 & -1 \\ 7 & 4 \\ 8 & 3 \end{pmatrix}}$$

$$\underset{\mathbf{X}}{\begin{pmatrix} 5 & 7 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}} + \underset{\mathbf{Y}}{\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \end{pmatrix}} = \underset{\mathbf{Z}}{\begin{pmatrix} 7 & -1 \\ 7 & 4 \\ 8 & 3 \end{pmatrix}}$$

$$\underset{\mathbf{-X}}{\begin{pmatrix} -5 & -7 \\ -4 & -2 \\ -6 & -1 \end{pmatrix}} + \underset{\mathbf{X}}{\begin{pmatrix} 5 & 7 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}} + \underset{\mathbf{Y}}{\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \end{pmatrix}} = \underset{\mathbf{-X}}{\begin{pmatrix} -5 & -7 \\ -4 & -2 \\ -6 & -1 \end{pmatrix}} + \underset{\mathbf{Z}}{\begin{pmatrix} 7 & -1 \\ 7 & 4 \\ 8 & 3 \end{pmatrix}}$$

$$\underbrace{\begin{pmatrix} -5 & -7 \\ -4 & -2 \\ -6 & -1 \end{pmatrix} + \begin{pmatrix} 5 & 7 \\ 4 & 2 \\ 6 & 1 \end{pmatrix}} + \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \end{pmatrix} = \underbrace{\begin{pmatrix} -5 & -7 \\ -4 & -2 \\ -6 & -1 \end{pmatrix} + \begin{pmatrix} 7 & -1 \\ 7 & 4 \\ 8 & 3 \end{pmatrix}}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ y_{31} & y_{32} \end{pmatrix} = \begin{pmatrix} 2 & -8 \\ 3 & 2 \\ 2 & 2 \end{pmatrix}$$

Scalar multiplication is accomplished by multiplying every element of a matrix by a constant scalar:

$$\mathbf{F}_{3,3} = \begin{pmatrix} 1 & 3 & -4 \\ 3 & 11 & 7 \\ -4 & 7 & 2 \end{pmatrix} \qquad k = 5$$

$$k\mathbf{F} = \begin{pmatrix} 5 \times 1 & 5 \times 3 & 5 \times -4 \\ 5 \times 3 & 5 \times 11 & 5 \times 7 \\ 5 \times -4 & 5 \times 7 & 5 \times 2 \end{pmatrix}$$

Matrix multiplication requires attending to a few important rules:

- The order of multiplication is important.
- Matrices can only be multiplied if the number of columns of the first matrix is equal to the number of rows of the second matrix.
- The resulting matrix has an order equal to the number of rows of the first matrix and the number of columns of the second matrix.

$A_{ji}B_{ik} = C_{jk}$

These two must be the same.

The elements of **C** are defined as:

$$c_{j,k} = \sum_i a_{j,i} b_{i,k}$$

$$\begin{pmatrix} 4 & -1 \\ 7 & 0 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} 1 & 7 & 0 \\ 4 & 2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 26 & -6 \\ 7 & 49 & 0 \\ 14 & 46 & 12 \end{pmatrix}$$

$\mathbf{A}_{3,2}$ $\mathbf{B}_{2,3}$ $\mathbf{C}_{3,3}$

$(4 \times 1) + (-1 \times 4) = 0$

$(6 \times 0) + (2 \times 6) = 12$

Matrix multiplication is not commutative:

- It may not be possible:
  $\mathbf{A}_{2,3}\mathbf{B}_{3,5} = \mathbf{C}_{2,5}$ but $\mathbf{B}_{3,5}\mathbf{A}_{2,3}$ is impossible
  $=$ $\neq$
- When possible, the results may not be equal: $\mathbf{A}_{1,3}\mathbf{B}_{3,1} = \mathbf{C}_{1,1}$ but $\mathbf{B}_{3,1}\mathbf{A}_{1,3} = \mathbf{C}_{3,3}$
  $\neq$

Matrix multiplication is associative:

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

Matrix multiplication is distributive:

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

But order is important:

$$\mathbf{XA} + \mathbf{BX} \neq \mathbf{X}(\mathbf{A} + \mathbf{B})$$

Every matrix has a *transpose* that is obtained by exchanging the rows and columns:

$$\begin{pmatrix} 4 & -1 \\ 7 & 0 \\ 6 & 2 \end{pmatrix} \qquad \begin{pmatrix} 4 & 7 & 6 \\ -1 & 0 & 2 \end{pmatrix}$$

$\mathbf{X}$ $\qquad$ $\mathbf{X}'$ (or $\mathbf{X}^T$)

Transposes are useful for arranging a matrix so that matrix multiplication is possible.

Example: A common statistical requirement is to generate the sums of squares and cross-products for a data matrix. If $\mathbf{X}_{n,v}$ is a matrix of deviation scores, then $\mathbf{X}_{nv}\mathbf{X}_{nv}$ is not possible. But, $\mathbf{X}'_{vn}\mathbf{X}_{nv}$ can be carried out.

Sum of deviation cross-products

$$\underbrace{\begin{bmatrix} d_{11} & d_{21} & d_{31} & d_{41} & d_{51} \\ d_{12} & d_{22} & d_{32} & d_{42} & d_{52} \end{bmatrix}}_{\mathbf{D}'} \underbrace{\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \\ d_{41} & d_{42} \\ d_{51} & d_{52} \end{bmatrix}}_{\mathbf{D}} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

D is a matrix of deviation scores for 5 individuals on 2 variables.

Sum of squared deviations

$\mathbf{X}_{nv}$ is a People x Variables matrix of deviation scores.

$\mathbf{X}'_{vn}\mathbf{X}_{nv} =$

$\sum_i x_{i,1}^2$			
$\sum_i x_{i,2}\, x_{i,1}$	$\sum_i x_{i,2}^2$		
$\sum_i x_{i,v}\, x_{i,1}$			$\sum_i x_{i,v}^2$

This matrix is one step away from what other matrix?

The *identity matrix* is a diagonal matrix with ones on the main diagonal:

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

The identity matrix is often a useful target matrix in statistics.

Multiplication by diagonal matrices is especially important in statistics and is used to accomplish rescaling (expanding, shrinking, standardizing).

Post-multiplication of a matrix **X** by a diagonal matrix **D** results in the columns of **X** being multiplied by the corresponding diagonal element in **D**.

$$\underset{\mathbf{X}}{\begin{pmatrix} 4 & 7 \\ 7 & 6 \end{pmatrix}} \underset{\mathbf{D}}{\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}} \quad \underset{\mathbf{Y}}{\begin{pmatrix} 12 & 14 \\ 21 & 12 \end{pmatrix}}$$

The first column of X is multiplied by the diagonal element in the first column of D

The second column of X is multiplied by the diagonal element in the second column of D

Pre-multiplication of a matrix **X** by a diagonal matrix **D** results in the rows of **X** being multiplied by the corresponding diagonal element in **D**.

$$\underset{\mathbf{D}}{\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}} \underset{\mathbf{X}}{\begin{bmatrix} 4 & 7 \\ 7 & 6 \end{bmatrix}} = \underset{\mathbf{Y}}{\begin{bmatrix} 12 & 21 \\ 14 & 12 \end{bmatrix}}$$

The first row of X is multiplied by the diagonal element in the first row of D

The second row of X is multiplied by the diagonal element in the second row of D

Scalar multiplication is just multiplication by a diagonal matrix with a constant in the diagonal.

$$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 4 & 7 \\ 7 & 6 \end{bmatrix} = \begin{bmatrix} 12 & 21 \\ 21 & 18 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 7 \\ 7 & 6 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 12 & 21 \\ 21 & 18 \end{bmatrix}$$

Variance-covariance matrices and correlation matrices can be characterized by a single number called the *determinant* that represents the "generalized variance."

For the correlation matrix, this number can take on values from 0 to 1. When all variables are independent (an identity matrix), the determinant is 1. As variables increase in their interdependence, the determinant approaches 0. A *singular matrix* has a determinant of 0.

The determinant thus indexes the redundancy among variables in a correlation matrix.

The determinant of the matrix, **A**, is symbolized:

$$|\mathbf{A}|$$

The determinant of an identity matrix, $|\mathbf{I}|$, equals 1.

Some *square* matrices have an inverse such that $\mathbf{AA^{-1} = I}$. The inverse is useful in solving matrix equations:

$$\mathbf{Y} = \mathbf{BR}$$

Solving for **B**:

$$\mathbf{YR^{-1}} = \mathbf{B}\boxed{\mathbf{RR^{-1}}}$$

$$\mathbf{YR^{-1}} = \mathbf{B}$$

This is equal to an identity matrix and so drops out.

$$X_{10,20} = (Z_{10,5} + B_{10,5})\, A_{5,20}$$

solve for B:

$$X_{10,20} = Z_{10,5}A_{5,20} + B_{10,5}A_{5,20}$$

$$X_{10,20} - Z_{10,5}A_{5,20} = B_{10,5}A_{5,20}$$

$$(X_{10,20} - Z_{10,5}A_{5,20})A'_{20,5} = B_{10,5}A_{5,20}A'_{20,5}$$

$$(X_{10,20} - Z_{10,5}A_{5,20})A'_{20,5}(A_{5,20}A'_{20,5})^{-1} =$$

$$B_{10,5}\underbrace{(A_{5,20}A'_{20,5})(A_{5,20}A'_{20,5})^{-1}}_{I}$$

If I create the following linear combinations of the original variables:

$$\underbrace{\begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{1,2} & X_{2,3} \\ X_{3,1} & X_{2,2} & X_{3,3} \\ \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,2} & X_{n,3} \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}}_{\mathbf{W}} = \mathbf{Y}$$

What do the new variables represent? What are the variances and covariances of the new variables?

Everything we need to know about the variances and covariances of the linear combinations is contained in the variance-covariance matrix of the original variables (Σ).

$\Sigma =$

	V_1	V_2	V_3
V_1	σ^2_1	σ_{12}	σ_{13}
V_2	σ_{21}	σ^2_2	σ_{23}
V_3	σ_{31}	σ_{32}	σ^2_3

If $\mathbf{Y} = \mathbf{XW}$, then $\Sigma_y = \mathbf{W'}\,\Sigma\,\mathbf{W}$

Another common set of transformations that will be made can be summarized as:

LXM

In which **X** is a Groups x Variables matrix of means, **L** is matrix of weights used to create linear combinations of groups (e.g., group contrasts), and **M** is a matrix of weights used to create linear combinations of variables.

By carefully constructing **L** and **M**, we can make any group comparisons for any combinations of variables.

Next time:

Univariate statistics and the use of **L** and **M.**