Some matrix stuff

What are Matrices?

- A matrix is simply a way of organizing and manipulating information in a grid system consisting of rows and columns

- Kinds of Matrices Commonly Used with Multivariate Methods
  - *Scalar* [1 x 1]
  - *Data* [N x p]
  - *Vector* [p x 1] or [1 x p]
  - *Diagonal Matrix* is a square, symmetric matrix with values on the diagonal and zeros on the off-diagonals.
    - A common diagonal matrix is an Identity Matrix, "**I**", which is the matrix parallel of the scalar, 1

Matrix Types

- *Variance-Covariance Matrix*, "Σ" or **S**,is a [p x p] square and symmetric matrix from the population that has the variances of the p variables along the diagonal and covariances in the off-diagonals

- *Sum of Squares and Cross-Products Matrix*, "**SSCP**" is a [p x p] square, symmetrical matrix and contains the numerators of the variance-covariance, Σ matrix.

- *Correlation Matrix*, "**R**" is a [p x p] square and symmetrical matrix. Ones are placed along the diagonal, with the off-diagonals showing the magnitude and direction of relationship between pairs of variables, using a standardized metric ranging from -1 to 1.

Kinds Of Calculations With Matrices

- Adding and multiplying a matrix by a constant (scalar)
- This involves just adding or multiplying each element in the matrix by that value
- Subtracting or dividing a matrix by a constant (scalar) is done the same way
- Adding matrices requires that each matrix be the same size, that is, conformable. To add matrices, simply add corresponding elements in the two matrices
  - Row 1, Col 1 of first matrix to R1 C1 of the second matrix and so on
- Subtracting matrices involves subtracting corresponding elements from two matrices of the same size
- Multiplying matrices involves summing the products of corresponding elements in a row of the first matrix with those from a column of the second matrix.
  - This requires that the number of columns in the 1st matrix equal the number of rows in the 2nd matrix.

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 \\ 1+7 & 0+5 \\ 1+2 & 2+1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 8 & 5 \\ 3 & 3 \end{bmatrix}$$

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

$$\begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \times 3 + 0 \times 2 + 2 \times 1 & 1 \times 1 + 0 \times 1 + 2 \times 0 \\ -1 \times 3 + 3 \times 2 + 1 \times 1 & -1 \times 1 + 3 \times 1 + 1 \times 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

Dividing Matrices

- Division by a scalar is straightforward
  - Divide each element of a matrix by that scalar

- *Dividing matrices* is similar to dividing scalars in logic.
  - Matrix **A** divided by B
  - When dividing two matrices, we *multiply* the first matrix (e.g., **A**) by the inverse of the second matrix (e.g., **B**): $\mathbf{B}^{-1}\mathbf{A}$
- If **B**, the matrix for which we need an inverse, is a [2 x 2] matrix, we can calculate the inverse by dividing the adjoint of **B** by the determinant
  - The *Determinant* of a [2 x 2] matrix is a single number that provides an index of the *generalized variance* of a matrix
    - Subtract the product of the elements on the second diagonal from the product of the elements on the first diagonal
  - For an *Adjoint* of a [2 x 2] matrix, switch main diagonal elements, and multiply off-diagonal elements, in their original place, by -1
- This gets tedious when going beyond 2 x 2 but the approach is the same

$$\mathbf{B} = \begin{bmatrix} a & c \\ d & b \end{bmatrix}$$

$$\mathrm{Det}(\mathbf{B}) = ab - cd$$

$$Adj(\mathbf{B}) = \begin{bmatrix} b & -c \\ -d & a \end{bmatrix}$$

Central Themes of Variance and Covariance Applied To Matrices

- Most matrices used in multivariate methods involve some form of variances and covariances
- The **SSCP** matrix has the numerators of the variances (i.e., the sums of squares) along the diagonal and the numerators of the covariances (i.e., the cross products) along the off-diagonals
- The variance-covariance matrix, **S**, holds the variances along the diagonal and covariances in the off diagonal

Linear Combinations, Eigenvalues and Eigenvector weights

- *Linear combinations* maximize the amount of information or variance from a set of variables into a set of composite variables
- An *eigenvalue* is the variance for a linear combination
- A *trace* is the sum of the diagonal elements, which is equal to the sum of the eigenvalues of a matrix
- The specific amount of variance taken from each variable is called an *eigenvector weight* (similar to unstandardized multiple regression weight in telling how much a variable relates to the overall linear combination)

Macro-level Assessment of Matrices

- In multivariate methods, we often examine some matrix ratio of Model over residual information and assess whether it is significant
- In MANOVA and DFA we look at the ratio of between group over within group **SSCP** matrices
- In Canonical Correlation, we look at the ratio of correlations between Xs and Ys over the matrix of correlations within Xs and Ys
- Forming a ratio of matrices yields another matrix that is not as easily summarized as a single number, such as the F-ratio in ANOVA
- We usually summarize matrices by means of traces, determinants, eigenvalues

Micro-Level Assessment of Matrices

- Examine weights or means to assess the micro-aspects of an analysis
- Weights can have several forms:
  - Unstandardized (eigenvector) weights
  - Standardized weights (usually ranging from -1 to + 1)
  - Loadings that are correlations between a variable and its linear combination (e.g., see DFA, CC, FA & PCA)

Trace

– sum of diagonal elements

$$B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 6 & 1 & 0 \\ 2 & 8 & 7 \\ 2 & 4 & 5 \end{pmatrix}$$

$$Trace = 6 + 8 + 5 = 19$$

Trace

- If the matrix is an SSCP matrix then the trace is the total sum-of-squares
- If the matrix is the variance/covariance matrix than the trace is simply the sum of variances
- If it is a correlation matrix the trace is just the number of variables

Eigenvalues and Eigenvectors

- This is a way of rearranging and consolidating the variance in a matrix

$$\underset{MxM}{D}\ \underset{Mx1}{V} = \underset{1x1}{\lambda}\ \underset{Mx1}{V}$$

$$D = any\ square\ matrix$$

$$V = Eigenvector$$

$$\lambda = Eigenvalue$$

- Think of it as taking a matrix and allowing it to be represented by a scalar and a vector (actually a few scalars and vectors, because there is usually more than one solution).

Eigenvalues and Eigenvectors

- Another way to look at it is that we are trying to come up with λ and V that will allow for the following equation to be true[1]

$$(D-\lambda I)V = 0$$

- D is the matrix of interest, I the identity matrix

$$(D-\lambda I)V = 0$$

$$\left[\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right]\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\left[\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right]\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

Eigenvalues and Eigenvectors

If v1 and v2 equal zero the previous statement is true, but uninteresting.

A non-trivial solution is found when the determinate of the

leftmost matrix is set to be 0.

$$(a-\lambda)(d-\lambda)-bc=0$$

$$\lambda^2-(a+d)\lambda+ad-bc=0$$

Solving for the eigenvalues λ requires solving for the roots of this polynomial

Generalize it to $x\lambda^2-y\lambda+z=0$

To solve for λ apply:

$$\lambda=\frac{-y\pm\sqrt{y^2-4xz}}{2x}$$

Eigenvalues and Eigenvectors

$$D = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

$$\lambda^2 - (5 + 2)\lambda + 5 * 2 - 1 * 4 = 0$$

$$\lambda^2 - 7\lambda + 6 = 0$$

$$\lambda = \frac{-(-7) + \sqrt{7^2 - 4 * 1 * 6}}{2 * 1} = 6$$

$$\lambda = \frac{-(-7) - \sqrt{7^2 - 4 * 1 * 6}}{2 * 1} = 1$$

$$\lambda_1 = 6, \lambda_2 = 1$$

Eigenvalues and Eigenvectors

- Using the first eigenvalue we solve for its corresponding eigenvector

$$\begin{bmatrix} 5-6 & 1 \\ 4 & 2-6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -1 & 1 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

This gives you two equations:

$$-1v_1 + 1v_2 = 0$$

$$4v_1 - 4v_2 = 0$$

$$V_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- Using the second eigenvalue we solve for its corresponding eigenvector

$$\begin{bmatrix} 5-1 & 1 \\ 4 & 2-1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} 4 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

This gives you two equations:

$$4v_1 + 1v_2 = 0$$

$$4v_1 + 1v_2 = 0$$

$$V_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- Let’s show that the original equation holds

$$\begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix} * \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix} \text{ and } 6 * \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix} * \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \text{ and } 1 * \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

Determinant

$$D = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$|D| = a_{11}a_{22} - a_{12}a_{21}$$

- This is considered the generalized variance of a matrix, i.e. a single number representing the variability in the matrix
  - Usually signified by |X|
  - For a 2 X 2 matrix the determinate is simply the product of the main diagonal minus the product of the other diagonal
- It is equal to the product of the eigenvalues for a given matrix
- For a correlation matrix, it will range from 0 to 1

Determinant

- For larger matrices the calculations become tedious and are best left to the computer
- What it can do for us is give a sense of how much independent information is contained within a set of variables
  - Larger determinant = more unique sources of information
  - Smaller determinant = more redundancy/correlation among the variables
- If a determinate of a matrix equals 0 than that matrix cannot inverted, since the inversion process requires division by the determinate.
  - A determinant of zero is caused by redundant data
  - Multicollinearity among predictors leads to determinants near zero, and unstable (inefficient) parameter estimates or simply broken analyses
    - This is often why you get error messages in more advanced techniques, some variable(s) is a linear combination of the others

Questions in the Use of Matrices

- Under what circumstances would we subtract a constant from a matrix?
- What is the interrelationship of **SSCP**, covariance, and correlation matrices, i.e., how do we proceed from one matrix to the next?
- Why and when would we divide the between groups variance matrix by the within groups variance matrix?
- How is the concept of orthogonality related to the concept of a determinant?

Summary

- Several kinds of matrices are commonly used in multivariate methods
- (1 x 1) scalar matrix, such as N or p
  - (N x p) Data Matrix, **X**
  - Vector which is simply a row or column from a larger matrix

- Matrix calculations include:
  - Addition, subtraction, multiplication, division (i.e., inverse)

- We can form a progression from **SSCP**, to **S**, and **R** matrices, each depicting relationship between variables in increasingly standardized and interpretable form.
- It is often the **S** or **R** matrix that serves as the actual data for an analysis

Summary Table

Concepts	Description
Scalar	(1 x 1) matrix or number (e.g., N, p, M)
Data Matrix	(N x p) **X** matrix with N rows and p columns
Vector	(p x 1) column or (1 x p) row of a matrix
SSCP Matrix	(p x p) square, symmetric matrix with sums of squares on the diagonal and cross-products on the off-diagonals
S Matrix	(p x p) square, symmetric matrix with variances on the diagonal and covariances on the off-diagonals
R Matrix	(p x p) square, symmetric matrix with 1's on the diagonal and correlations ranging from -1 to +1 on the off-diagonals
I Matrix	Square, Symmetric matrix with 1/s on the diagonal and 0's on the off-diagonals. It is the matrix parallel of the scalar "1"
Transpose	Matrix formed by exchanging rows (for **A**) for columns (in **A'**)
Eigenvalue	Variance of a linear combination; the sum of the eigenvalues = the trace of the original matrix
Eigenvector weights	Coefficients that indicate how much variance to extract from a variable when forming a linear combination
Adjoint	Matrix formed by exchanging the diagonal elements and multiplying off-diagonals by -1 in a (2 x 2) matrix
Determinant	Generalized variance of a matrix (ranging from 0 - 1 for **R**) equal to the product of the eigenvalues of a matrix
Inverse	A "divisor" matrix used in dividing a matrix. In a (2 x 2) matrix, this equals the Adjoint divided by the Determinant of a matrix
Trace	Sum of the diagonal elements of a matrix