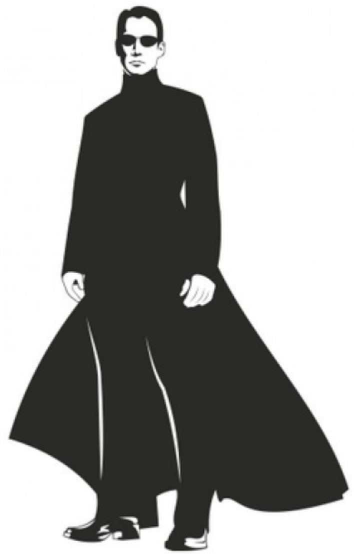
# Matrix algebra

PSY544 – Introduction to Factor Analysis

Week 2



# Prologue

- Matrix algebra is a framework for manipulating collections of numbers or algebraic symbols.

- Factor model is an algebraic system. If you understand the way it is communicated, you gain a better appreciation of what is going on.

- We have already seen the common factor model representing the structure of score $x_{ij}$ – this model applies to every $x_{ij}$ in the data matrix $\boldsymbol{X}$. Matrix algebra will allow us to express that.

# Definitions

- **Scalar**: A single value, e.g., *k = 3, z = 0.7*
- **Matrix:** A rectangular table of *elements* (numbers, symbols…):

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} 1.0 & 3.0 \\ 0.2 & 4.6 \\ 8.5 & -2.3 \end{bmatrix}$$

$a_{ij}$ = element in row *i* and column *j* of matrix ***A***

$a_{11} = 1.0$, $a_{21} = 0.2$

An uppercase letter (like ***A***) names the matrix and stands for all the elements

# Definitions

- **Data matrix: Each element is a score for an individual on a variable**

$p$ = 3 manifest variables

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix}$$

$N$ = 4 subjects

- $\boldsymbol{x_{ij}}$ **= score for the *i*-th individual on the *j*-th variable**

# Definitions

- **Order**: The size of a matrix.
- A matrix with *N* rows and *p* columns is of order *N x p*

- **Square matrix**: A matrix with the same number of rows and columns
- **Vector**: A matrix with a single column (column vector) or a single row (row vector)

$$\mathrm{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_m \end{bmatrix} \quad \text{m x 1 column vector}$$

**1 x m row vector**

$$\mathrm{w} = [w_1 \quad w_2 \quad w_3 \quad \cdots \quad w_m]$$

# Definitions

- **Transpose** of matrix ***A*** is a new matrix, ***A'***, formed by writing the rows of ***A*** as the columns of ***A'***:

If $A = \begin{bmatrix} 1.0 & 0.7 \\ 0.5 & 0.9 \\ 0.6 & 1.2 \end{bmatrix}$ then the transpose is $A' = \begin{bmatrix} 1.0 & 0.5 & 0.6 \\ 0.7 & 0.9 & 1.2 \end{bmatrix}$

- The transpose of a *P x M* matrix ***A*** is a *M x P* matrix ***A'*** (or $A^T$)
- **Equality of matrices:** Two matrices are equal if and only if they are of the same order and all their corresponding elements are equal.

# Arithmetic operations

- **Addition and subtraction:** Only possible if matrices are of the same order (*conformable* for addition)
- Corresponding elements of ***A*** and ***B*** are summed to form ***C***.
- $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ ; $c_{ij} = a_{ij} + b_{ij}$ for all *i* and *j*

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

Commutative law: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

Associative law: $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$

The transpose of a sum is the sum of the transposes: $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$

# Multiplication

- **Scalar multiplication:**

$$C = Ba = aB \; ; \; c_{ij} = a \cdot b_{ij}$$

$$2\begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix} 2 = \begin{bmatrix} 4 & 8 \\ 2 & 0 \end{bmatrix}$$

# Multiplication

- **Matrix multiplication** is **not** what you might think it is!
- **Matrix multiplication**: ***A*** and ***B*** have to be *conformable* for multiplication. Matrices are conformable if the number of columns in the first matrix equals the number of rows in the second.

$$\underset{n \times p}{\boldsymbol{A}} \quad \underset{p \times m}{\boldsymbol{B}} = \boldsymbol{AB} = \underset{n \times m}{\boldsymbol{C}}$$

- The product matrix has as many rows as the first matrix and as many columns as the second matrix.
- In this case, ***B*** was being *pre-multiplied* by ***A*** (and ***A*** was being *post-multiplied by B*)

# Multiplication

- **Matrix multiplication:** $c_{ij}$ equals the sum of products of elements in the *i*-th row of ***A*** and the *j*-th column of ***B***.

$$\begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{ip} \end{bmatrix} \text{x} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \dots \\ b_{pj} \end{bmatrix} = \begin{bmatrix} c_{ij} \end{bmatrix}$$

$$c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + \dots + a_{ip}\, b_{pj}$$

# Multiplication

- For instance, if ***A*** is a 2x2 matrix and ***B*** is a 2x3 matrix, then ***AB = C*** is a 2x3 matrix:

$$\overset{A}{\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}} \overset{B}{\begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 5 \end{bmatrix}} = \overset{C}{\begin{bmatrix} 7 & 9 & 23 \\ 4 & 5 & 14 \end{bmatrix}}$$

$$\begin{bmatrix} 2 & 3 \\ * & * \end{bmatrix} \begin{bmatrix} * & 3 & * \\ * & 1 & * \end{bmatrix} = \begin{bmatrix} * & 9 & * \\ * & * & * \end{bmatrix}$$

- Note again that order matters here! ***AB = C***, but ***BA*** is undefined.

# Multiplication

- Matrix multiplication is associative: $\boldsymbol{A}(\boldsymbol{BC}) = \boldsymbol{AB}(\boldsymbol{C})$
- Matrix multiplication is not commutative: $\boldsymbol{AB} \neq \boldsymbol{BA}$
- Matrix multiplication is distributive: $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$; $(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A} = \boldsymbol{BA} + \boldsymbol{CA}$
- The transpose of a product of matrices equals the product of the transposes in reverse order: $(\boldsymbol{AB})' = \boldsymbol{B}' \boldsymbol{A}'$

# Kinds of matrices

- **Symmetric matrix:** A square matrix $\boldsymbol{S}$ is symmetric if it's equal to its transpose

$$\boldsymbol{S} = \begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix} = \boldsymbol{S}'$$

# Kinds of matrices

- **Triangular matrix:** A square matrix ***S*** is lower triangular if all the elements above the diagonal are zero. A square matrix ***R*** is upper triangular if all the elements below the diagonal are equal to zero.

$$\boldsymbol{S} = \begin{bmatrix} 2 & 0 & 0 \\ 4 & 5 & 0 \\ 1 & 2 & 3 \end{bmatrix}$$

$$\boldsymbol{R} = \begin{bmatrix} 2 & 4 & 1 \\ 0 & 5 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

# Kinds of matrices

- **Diagonal matrix**: A square matrix ***D*** is diagonal if all the off-diagonal elements are zero.

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix}$$

# Kinds of matrices

- **Pre-multiplication** by a diagonal matrix scales rows:

$$\boldsymbol{DA} = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} d_{11}a_{11} & d_{11}a_{12} \\ d_{22}a_{21} & d_{22}a_{22} \\ d_{33}a_{31} & d_{33}a_{32} \end{bmatrix}$$

- **Post-multiplication** by a diagonal matrix scales columns:

$$\boldsymbol{BD} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} = \begin{bmatrix} d_{11}b_{11} & d_{22}b_{12} & d_{33}b_{13} \\ d_{11}b_{21} & d_{22}b_{22} & d_{33}b_{23} \end{bmatrix}$$

# Kinds of matrices

- **Identity matrix** is a diagonal matrix with all diagonal elements = 1

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

- Multiplication of a matrix by an identity matrix does not change the matrix (it's like multiplying a scalar by 1)

# Kinds of matrices

- **Orthogonal matrix**: A square matrix ***T*** is orthogonal if ***TT' = I*** or ***T'T = I***

- **Correlation matrix** is a square, symmetric matrix with unit diagonals and off-diagonal elements that satisfy $-1 \leq r_{ij} \leq 1$. Also, it has to be nonnegative definite (we will define that later)

# Functions of matrices

- The **Determinant** of a square matrix ***A*** is a scalar function of the elements of ***A***. It is denoted as $|\boldsymbol{A}|$ or det(***A***) and is a single number (scalar).
- The determinant has many functions which we will not cover here (neither will we cover the definition or computation)
- If a matrix has determinant equal to zero, the matrix is called *singular*. This is an indication that there is redundancy among the rows / columns of the matrix – if the determinant is zero, some columns (or rows) of the matrix can be expressed as linear combinations of other columns (rows). In other words, the columns (rows) are linearly dependent.

# Functions of matrices

- **A singular matrix:**

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 & 4 \\ 4 & 1 & 5 \\ 2 & 3 & 5 \end{bmatrix}$$

**(The last column is the sum of the first two columns)**

# Functions of matrices

- **Trace:** The trace of a square matrix ***A***, tr(***A***), is the sum of its diagonal elements.

- **Rank:** The column rank of ***A*** is equal to the total number of linearly independent columns of ***A***. The row rank of ***A*** is equal to the total number of linearly independent rows of ***A***.

  The rank of an N x K matrix is at most the minimum of N or K, min(N,K)
- A matrix whose rank is equal to min(N,K) is *full rank*
- A matrix whose rank is less than min(N,K) is *rank deficient*

# Functions of matrices

- **Inverse:** If $\boldsymbol{A}$ is a square matrix and is not singular (i.e., its determinant is non-zero), then it has a unique inverse $\boldsymbol{A}^{-1}$ such that:

$$\boldsymbol{AA^{-1} = A^{-1}A = I}$$

- The inverse of a matrix plays a role similar to that of a reciprocal in scalar algebra: $x * \frac{1}{x} = 1$
- Post-multiplying $\boldsymbol{A}$ by the inverse of $\boldsymbol{B}$ is analogous to „dividing“ $\boldsymbol{A}$ by $\boldsymbol{B}$ (assuming the matrices are conformable for multiplication)

# Functions of matrices

- **Solving** equations:

Consider the equation ***Ax = b***, where ***A*** is a N x N non-singular matrix***,*** ***b*** is a N x 1 vector and ***x*** is a N x 1 vector. We know the elements of ***A*** and ***b*** and wish to solve for ***x***:

$$\boldsymbol{Ax = b}$$

$$\boldsymbol{A^{-1}Ax = A^{-1}b}$$

$$\boldsymbol{Ix = A^{-1}b}$$

$$\boldsymbol{x = A^{-1}b}$$

# Functions of matrices

- **Solving** equations:

*4x + 5y = 4*

*3x + 1y = 3*

- In matrix form:

$$\begin{bmatrix} 4 & 5 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

solving:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{11} & \frac{5}{11} \\ \frac{3}{11} & -\frac{4}{11} \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

# Functions of matrices

- **Eigenvalues** and **Eigenvectors**
- Suppose that ***S*** is a square symmetric matrix of order *p*. If ***u*** is a column vector of order *p* and *v* is a scalar, such that:

$$\boldsymbol{Su} = v\boldsymbol{u}$$

...then *v* is said to be an *eigenvalue* (or characteristic root) of ***S*** and ***u*** is said to be an *eigenvector* (or characteristic vector) of ***S***.

- ***S*** will have *p* eigenvalues and *p* associated eigenvectors.

# Functions of matrices

- **Eigenvalues** and **Eigenvectors**

- If all *p* eigenvalues are positive, the matrix is *positive definite*. If one or more eigenvalues are zero and the rest is positive, the matrix is *nonnegative definite*. If one or more eigenvalues are negative, the matrix is *negative definite*.

- The determinant of $\boldsymbol{S}$, det($\boldsymbol{S}$), equals the product of the eigenvalues of $\boldsymbol{S}$ Thus, if one or more eigenvalues are zero, the matrix is singular.

# Functions of matrices

- **Eigenvalues** and **Eigenvectors**

- The eigenvalues can be arranged in descending order as the diagonal elements in a diagonal matrix ***D***, and the corresponding eigenvectors can be arranged as columns of matrix ***U***. Then:

  ***U*** is orthogonal, that is, ***U'U = I***
  The "eigenstructure" of ***S*** can be given in this form: ***SU = UD***
  It also holds that ***S = UDU'***

# Linear combinations of random variables

- Matrix equations are handy for representing linear combinations of random variables
- Let ***x*** be a column vector of order $p$ containing scores for a random individual on variables $x_1, x_2, \ldots, x_p$
- Let ***z*** be a column vector of order $m$ containing scores for a random individual on variables $z_1, z_2, \ldots, z_m$
- We will represent the variables in ***x*** as linear functions of the variables in ***z*.** Let ***A*** be a matrix of order $p$ x $m$ containing coefficients $a_{jk}$ representing the linear effects of $z_k$ on $x_j$
- Let $\boldsymbol{\mu}$ be a column vector of order $p$ containing fixed constants $\mu_1, \mu_2, \ldots, \mu_p$

# Linear combinations of random variables

- **Then, we can represent the variables in *x* as linear functions of the variables in *z* and the constants in *μ* using the following matrix equation:**

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{A}\mathbf{z}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_p \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_p \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pm} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \dots \\ z_m \end{bmatrix}$$

**…thus:**

$$x_p = \mu_p + a_{p1}z_1 + a_{p2}z_2 + \dots + a_{pm}z_m$$

- **The matrix equation actually contains the whole set of linear equations**

# An intermezzo – expected values

- Wiki: “The *expected value* of random variable is the long-run average value of repetitions of the experiment it represents”

...so, the *expected value* is the variable’s **mean**.

- $E[X] = \mu$

# An intermezzo – expected values

- Now, consider the (scalar) formula for the **variance** of a random variable:

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N}$$

…which is the “mean squared deviation from the mean”, right?

- As an expected value: $E[(X - \mu)^2]$

# An intermezzo – expected values

- **Now consider the (scalar) formula for covariance**

$$\sigma_{XY} = \frac{\sum(x - \mu_x)(y \ - \mu_y)}{N}$$

**...which is the "mean cross-product of deviations from the mean" (sorta)**

- **As an expected value: E[(X – $\mu_x$)(Y - $\mu_y$)]**

# Covariance matrix

- Now suppose that $\boldsymbol{x}$ is a vector of order $p$ containing scores on $p$ variables for a random individual selected from some population, and $\boldsymbol{\mu}$ is a vector of order $p$ containing the population means of these $p$ variables.
- Then, vector $(\boldsymbol{x} - \boldsymbol{\mu})$ stands for the vector $\boldsymbol{x}$ with the population means subtracted (it represents deviations from the mean)

- Let's multiply this vector by its transpose:

$$(\boldsymbol{x} - \boldsymbol{\mu})\ (\boldsymbol{x} - \boldsymbol{\mu})'$$

# Covariance matrix

- Let's multiply this vector by its transpose:

$$(\boldsymbol{x} - \boldsymbol{\mu})\ (\boldsymbol{x} - \boldsymbol{\mu})'$$

- …and take the expectation:

$$\mathrm{E}[(\boldsymbol{x} - \boldsymbol{\mu})\ (\boldsymbol{x} - \boldsymbol{\mu})']$$

# Covariance matrix

$$\mathrm{E}[(\boldsymbol{x}-\boldsymbol{\mu})\,(\boldsymbol{x}-\boldsymbol{\mu})']$$

- **Expanding, we get the expectation of:**

$$\begin{bmatrix}(x_1-\mu_1)\\(x_2-\mu_2)\\\vdots\\(x_p-\mu_p)\end{bmatrix}\begin{bmatrix}(x_1-\mu_1) & (x_2-\mu_2) & \cdots & (x_p-\mu_p)\end{bmatrix}$$

- **…which gives us the variance/covariance matrix of the manifest variables**

# Covariance matrix

$$E[(\boldsymbol{x} - \boldsymbol{\mu})\,(\boldsymbol{x} - \boldsymbol{\mu})']$$

- The variance-covariance matrix is a $p$ x $p$ symmetric matrix with variances on the diagonal and covariances off the diagonal