Matrix Algebra Review

(after the notes of L. R Tucker)

Steven M. Boker

Department of Psychology
University of Virginia

Introduction to SEM
Psyc–8501–001

Elementary Matrix Concepts and Operations

- A **matrix** is a rectangular table of values.
- A matrix may be designated by a single symbol, such as the letter **A**. Frequently, capital letters will be used to designate matrices.
- The cell entries of a matrix may be designated by a symbol with two subscripts, the first subscript indicating the **row** and the second subscript indicating the **column** location of the cell. Frequently, lower case letters are used to designate cell entries of matrices. For any matrix designated by a capital letter, the corresponding lower case letter will be used for the cell entries. Thus, the cell entries in matrix **A** are designated by a_{ij} where i indicates the row location and j indicates the column location.

Elementary Matrix Concepts and Operations

- The size of a matrix is the number of rows and the number of columns of the matrix and is termed the **order** of the matrix. Thus, a matrix with m rows and n columns is of order m by n, or $m \times n$.
- Two matrices are equal only when they are of the **same order** and each cell entry in one matrix equals the cell entry in the same row and column of the other matrix.
- A matrix equals zero only when every cell entry equals zero. Such a matrix is frequently termed a **null matrix** .
- The **transpose** of a matrix is a new matrix produced by writing the entries in each row of the original matrix as entries in the corresponding column of the transpose matrix. An equivalent operation is to write the entries in each column of the original matrix as entries in the corresponding row of the transpose matrix. The transpose matrix is designated by a symbol for the original matrix with a prime; for example, $\mathbf{A}'$ designates the transpose of matrix $\mathbf{A}$.

Elementary Matrix Concepts and Operations

- The transpose of the transpose of a matrix equals the original matrix:

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

- **Addition** and **subtraction** of matrices may be performed on matrices of the same order by performing the addition and subtraction of cell entries in identical locations in the matrices to produce the cell entries in the resulting matrix. Thus, the equation

$$\mathbf{A} + \mathbf{B} - \mathbf{C} = \mathbf{R}$$

implies

$$a_{ij} + b_{ij} - c_{ij} = r_{ij}$$

for each and every cell.

Elementary Matrix Concepts and Operations

- Matrix addition and subtraction obey the commutative and associative laws, thus:

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

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

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

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

- Matrix addition and subtraction obey the cancellation law; thus, if

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

then

$$\mathbf{B} = \mathbf{C}.$$

Elementary Matrix Concepts and Operations

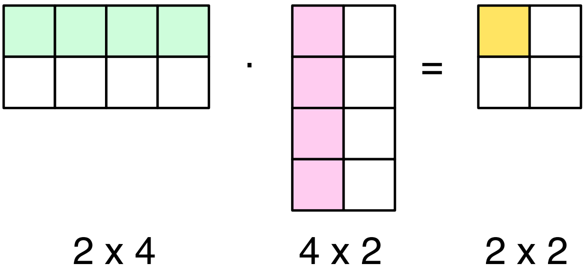
- Two matrices, such as **A** and **B**, may be **multiplied** when the number of columns in the first matrix (**A**) equals the number of rows in the second matrix (**B**) by obtaining for each combination of a row of the first matrix and a column of the second matrix the sum of products between corresponding cell entries in the row and column, thus producing the cell entry in that row and column of the product matrix. Thus, $\mathbf{AB} = \mathbf{C}$ implies that

$$a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + \cdots + a_{1m}b_{m1} = c_{11}$$

$$a_{31}b_{15} + a_{32}b_{25} + a_{33}b_{35} + \cdots + a_{3m}b_{m5} = c_{35}$$

$$\sum_{j=1}^{n} a_{ij}b_{jk} = c_{ik}.$$

Elementary Matrix Concepts and Operations



Elementary Matrix Concepts and Operations

- The product matrix resulting from multiplication of two matrices has number of rows corresponding to number of rows of the first matrix and number of columns corresponding to number of columns of the second matrix forming the product. Thus, the product of an $n \times m$ matrix $\mathbf{A}$ and an $m \times p$ matrix $\mathbf{B}$ will be an $n \times p$ matrix.
- Matrix multiplication obeys the associative law but not the commutative law; that is:

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

$$\mathbf{AB} \neq \mathbf{BA}.$$

Elementary Matrix Concepts and Operations

- In a matrix product, the first matrix is termed a **pre-multiplier** and the second matrix is termed a **post-multiplier**. Thus, in the matrix product $\mathbf{AB}$, $\mathbf{A}$ is a pre-multiplier and $\mathbf{B}$ is a post-multiplier. $\mathbf{A}$ is post-multiplied by $\mathbf{B}$ and $\mathbf{B}$ is pre-multiplied by $\mathbf{A}$.
- Matrix multiplication does not necessarily obey the cancellation law. Thus, if

$$\mathbf{AB} = \mathbf{AC}$$

  then $\mathbf{B}$ may or may not equal $\mathbf{C}$.
- Matrix addition (and subtraction) and multiplication obey the distribution law. Thus,

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

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

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

Elementary Matrix Concepts and Operations

- The transpose of a product matrix equals the product of the transposes of the original matrices in reverse order. Thus, if

$$\mathbf{AB} = \mathbf{C}$$

then

$$\mathbf{B}'\mathbf{A}' = \mathbf{C}'$$

- The transpose of a sum of matrices is the sum of the transposes of the original matrices. Thus, if

$$\mathbf{A} + \mathbf{B} - \mathbf{C} = \mathbf{R}$$

then

$$\mathbf{A}' + \mathbf{B}' - \mathbf{C}' = \mathbf{R}'$$

Elementary Matrix Concepts and Operations

- The square root of the sum of squares of the entries in a matrix, say **A**, is called the **norm** of the matrix and is indicated by $Norm(\mathbf{A})$.

$$Norm(\mathbf{A}) = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{n} a_{ij}^2}$$

  where $i = 1, 2, 3, \ldots, N$ and $j = 1, 2, 3, \ldots, n$.

- A matrix that has as many rows as it has columns is called a **square matrix**; otherwise the matrix is a **rectangular matrix**.
- In a square matrix the diagonal of the cells running from upper left to lower right is termed the **principal diagonal** of the matrix. This terminology is frequently shortened to the **diagonal** of the matrix.

Elementary Matrix Concepts and Operations

- The sum of the entries in the principal diagonal of a square matrix is termed the **trace** of the matrix and is designated by $Trace(\mathbf{A})$ where $\mathbf{A}$ is the matrix.

  $$Trace(\mathbf{A}) = \sum_{i=1}^{N} a_{ii}$$

  where $i = 1, 2, 3, \ldots, N$
- A square matrix with zero entries in all cells not on the principal diagonal is termed a **diagonal matrix**. Thus, if $\mathbf{D}$ is a diagonal matrix then

  $$d_{ij} = 0$$

  for $i \neq j$

Elementary Matrix Concepts and Operations

- Premultiplication of a matrix, say **A**, by a diagonal matrix, say **D**, multiplies the entries in each row of **A** by the diagonal entry in **D** in the diagonal cell corresponding to that row of **A**.

$$\mathbf{DA} = \mathbf{B}$$

then

$$d_{ii} a_{ij} = b_{ij}$$

- Post-multiplication of a matrix, say **A** by a diagonal matrix, say **D**, multiplies the entries in each column of **A** by the diagonal entry in **D** in the diagonal cell corresponding to that column of **A**.

$$\mathbf{AD} = \mathbf{B}$$

then

$$a_{ij} d_{jj} = b_{ij}$$

Elementary Matrix Concepts and Operations

- The product of two diagonal matrices, say **D** and **E**, is a third diagonal matrix, say **F**, with diagonal entries equal to the products of corresponding entries in the first two diagonal matrices.

$$\mathbf{DE} = \mathbf{F}$$

$$d_{ii}e_{ii} = f_{ii}$$

for $i = 1, 2, 3, \ldots, N$

- The product of diagonal matrices is invariant over reversal in the order of multiplication. Let **D**, **E**, and **F** be diagonal matrices where

$$\mathbf{DE} = \mathbf{F},$$

then

$$\mathbf{ED} = \mathbf{F}.$$

Elementary Matrix Concepts and Operations

- A diagonal matrix with all diagonal entries equal to a single value is termed a **scalar matrix**.

$$d_{ii} = k$$

  for $i = 1, 2, 3, \ldots, N$

- Multiplication, either pre-multiplication or post-multiplication, of a matrix, say $\mathbf{A}$, by a scalar matrix, say $\mathbf{K}$, results in each entry of $\mathbf{A}$ being multiplied by the constant in the diagonal cells of $\mathbf{K}$.

$$\mathbf{AK} = \mathbf{B}$$

$$a_{ij}k = b_{ij}$$

- Multiplication by a scalar matrix is invariant with reversal of order of multiplication. Let $\mathbf{K}$ be a scalar matrix, then

$$\mathbf{AK} = \mathbf{KA}$$

  Note: $\mathbf{A}$ need not be square but the matrix $\mathbf{K}$ is to be made of appropriate order separately on each side of the equation to perform the indicated multiplication.

Elementary Matrix Concepts and Operations

- A scalar matrix with its constant diagonal entry equal to unity is termed an **identity matrix** and is designated by the capital letter **I**.

$$\mathbf{K} = \mathbf{I}$$

  if $k_{ii} = 1$ for $i = j$ and $k_{ii} = 0$ for $i \neq j$

- The product of a matrix, say **A**, with an identity matrix equals the matrix **A**.

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

- A square matrix having each entry on one side of the principal diagonal equal to the symmetrically located entry on the other side of the principal diagonal is termed a **symmetric matrix**.

$$a_{ij} = a_{hk}$$

  for $k = i$ and $h = j$

Elementary Matrix Concepts and Operations

- All symmetric matrices, and only symmetric matrices equal their transposes.

$$\mathbf{A} = \mathbf{A}'$$

- A **triangular matrix** is a square matrix with zero entries on one side of the diagonal. If the non-zero entries are below the diagonal, it will be termed a **lower triangular matrix**. If the non-zero entries are above the diagonal, it will be termed an **upper triangular matrix**.

Elementary Matrix Concepts and Operations

- The product of a matrix, say $\mathbf{A}$, post-multiplied by its transpose, $\mathbf{A}'$, is a symmetric matrix which will be termed the **rows product matrix** of $\mathbf{A}$ and designated by $\mathbf{P}_R$.

$$\mathbf{P}_R = \mathbf{P}'_R = \mathbf{A}\mathbf{A}'$$

- The product of a matrix, say $\mathbf{A}$, pre-multiplied by its transpose, $\mathbf{A}'$, is a symmetric matrix which will be termed the **columns product matrix** of $\mathbf{A}$ and designated by $\mathbf{P}_C$.

$$\mathbf{P}_C = \mathbf{P}'_C = \mathbf{A}'\mathbf{A}$$

- The trace of the rows product matrix of $\mathbf{A}$ equals the trace of the columns product matrix of $\mathbf{A}$ and each trace equals the square of the norm of matrix $\mathbf{A}$.

$$Trace(\mathbf{P}_R) = Trace(\mathbf{P}_C) = Norm(\mathbf{A})^2.$$

Elementary Vector Concepts and Operations

- An array of n numbers $(x_1, x_2, \ldots, x_n)$ is termed a vector $\mathbf{V}_x$ in an n dimensional space. Each number is taken as the coordinate of the terminus of the vector on one reference axis. The vector is taken to start at the origin.
  - A row of numbers is termed a **row vector**.
  - A column of numbers is termed a **column vector**.
- A matrix of order $n \times m$ may be interpreted as either:
  - n row vectors in a space of m dimensions, or
  - m column vectors in a space of n dimensions.
- **Addition** of two vectors is accomplished by addition of corresponding coordinates.

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

$$\mathbf{V}_x + \mathbf{V}_y = \mathbf{V}_{(x+y)}.$$

Elementary Vector Concepts and Operations

- **Scalar multiplication** of a vector by a scalar is accomplished by multiplying each coordinate of the vector by the scalar.

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n).$$

$$c\mathbf{V}_x = \mathbf{V}_{cx}.$$

- **Linear combination** of vectors is accomplished as a joint use of scalar multiplication of vectors and addition of vectors.

$$\begin{aligned} & c_x(x_1, x_2, \ldots, x_n) + c_y(y_1, y_2, \ldots, y_n) + \\ & \cdots + c_z(z_1, z_2, \ldots, z_n) \\ = \quad & c_x x_1 + c_y y_1 + \ldots + c_z z_1, \\ & c_x x_2 + c_y y_2 + \ldots + c_z z_2, \cdots, \\ & c_x x_n + c_y y_n + \ldots + c_z z_n). \end{aligned}$$

$$c_x\mathbf{V}_x + c_y\mathbf{V}_y + \cdots + c_z\mathbf{V}_z = \mathbf{V}_{(c_x x + c_y y + \cdots + c_z z)}.$$

Elementary Vector Concepts and Operations

- A vector $\mathbf{V}_u$ is termed as **linearly dependent** on a group of vectors, $\mathbf{V}_x, \mathbf{V}_y, \ldots, \mathbf{V}_z$ if scalars $c_x, c_y, \ldots, c_z$ can be found so that $\mathbf{V}_u$ is a linear combination of $\mathbf{V}_x, \mathbf{V}_y, \ldots, \mathbf{V}_z$:

$$\mathbf{V}_u = c_x\mathbf{V}_x + c_y\mathbf{V}_y + \cdots + c_z\mathbf{V}_z,$$

  otherwise, $\mathbf{V}_u$ is linearly independent of the given groups of vectors.
- The subspace composed exclusively of all vectors produced by linear combinations of a given group of vectors is said to be **spanned** by this group of vectors.
- A group of vectors is termed a **basis** for a subspace if (1) the subspace is spanned by the group of vectors, and (2) each vector in the group is linearly independent of the remaining vectors in the group.

Elementary Vector Concepts and Operations

- The number of vectors in a basis for a subspace is termed the **dimensionality** of the subspace.
- The **rank** of a matrix is the dimensionality of the subspace spanned by the vectors composing the matrix.
- The **length**, L_x, of a vector, $\mathbf{V}_x$, is the positive square root of the sum of squares of the coordinates of the vector.

$$L_x = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$$

- The **inner product**, P_{xy}, of two vectors, $\mathbf{V}_x$ and $\mathbf{V}_y$, is the sum of products between corresponding coordinates of the two vectors and equals the product of the lengths of the two vectors and the cosine of the angle between the two vectors.

$$P_{xy} = (x_1 y_1 + x_2 y_2 + \cdots + x_n y_n) = L_x L_y \cos \Theta_{xy}$$

Elementary Vector Concepts and Operations

- Two vectors are **orthogonal** if their inner product is zero provided that neither vector has a zero length.
- Two subspaces are orthogonal if every vector in the basis of one subspace is orthogonal to every vector in the basis of the other subspace.
- An **orthogonal basis** of a subspace is composed of mutually orthogonal vectors.
- A vector with unit length is termed a **unit vector**, or a **normal vector**.
- An **orthonormal basis** of a subspace contains mutually orthogonal unit vectors.

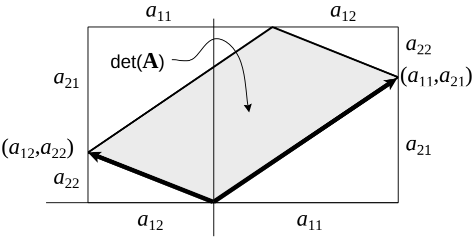
Elementary Vector Concepts and Operations

- When the basis for two subspaces combined form the basis of the whole space the two subspaces are termed **complementary subspaces**.
- Two orthogonal subspaces which are also complementary are termed **orthogonal complementary subspaces**.
- A subspace of dimensionality $n - 1$ in an n dimensional space is termed a **hyperplane** and its orthogonal complementary subspace is termed the **normal to the hyperplane**.

Inverse Matrices

- Let **A** be a square 2×2 matrix. The determinant of **A** is written as $|\mathbf{A}|$ or $\det(\mathbf{A})$, and is defined as

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21} \ .$$



The determinant of **A** is represented by the gray area of the parallelogram defined by the vectors (a_{11}, a_{21}) and (a_{12}, a_{22}).

Inverse Matrices

- The area in gray can be calculated as

$$
\begin{aligned}
|\mathbf{A}| &= (a_{11} - a_{12})(a_{21} + a_{22}) - (-a_{12}a_{22} + a_{11}a_{21}) \\
&= (a_{11}a_{21} - a_{12}a_{21} + a_{11}a_{22} - a_{12}a_{22}) - (-a_{12}a_{22} + a_{11}a_{21}) \\
&= -a_{12}a_{21} + a_{11}a_{22} \\
&= a_{11}a_{22} - a_{12}a_{21}
\end{aligned}
$$

Inverse Matrices

- Let $\mathbf{A}$ be a square matrix, if matrix $\mathbf{B}$ exists such that

$$\mathbf{AB} = \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix, then

$$\mathbf{BA} = \mathbf{I};$$

$\mathbf{B}$ is called the **inverse** of $\mathbf{A}$ and designated $\mathbf{A}^{-1}$; $\mathbf{A}$ is said to be **non-singular** which is symbolized by $|\mathbf{A}| \neq 0$.

- A square matrix which is non-singular has its rank equal to its order.

Inverse Matrices

- A group of n non-homogeneous linear equations in m unknowns:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1m}x_m & = & c_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2m}x_m & = & c_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nm}x_m & = & c_n \end{array}$$

  May be written in matrix form as:

$$\mathbf{AX} = \mathbf{C}$$

  where $\mathbf{A}$ is the $n \times m$ matrix of coefficients, a; $\mathbf{X}$ is an $m \times 1$ column vector of unknowns, x, and $\mathbf{C}$ is an $n \times 1$ column vector of constants, c.

- When the coefficient matrix $\mathbf{A}$ for a group of non-homogeneous equations is square and non-singular the solution for these equations is given by :

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{C}.$$

Inverse Matrices

- When there are more equations than unknowns in a group of non-homogeneous equations, $n > m$, then a **least squares approximation**

$$\mathbf{A}\tilde{\mathbf{X}} = \tilde{\mathbf{C}}$$

  is given by

$$\mathbf{A}'\mathbf{A}\mathbf{X} = \mathbf{A}'\mathbf{C}$$

  where $\tilde{\mathbf{C}}$ is a column vector such that $\mathbf{E} = \mathbf{C} - \tilde{\mathbf{C}}$ is also a column vector and has a minimum length,

$$L_{\mathbf{E}} = \text{a minimum.}$$

- In case the rows product moment matrix, $(\mathbf{A}'\mathbf{A})$, is non-singular, the solution for $\tilde{\mathbf{X}}$ is given by:

$$\tilde{\mathbf{X}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{C}.$$

Inverse Matrices

- When there are more unknowns than equations in a group of non-homogeneous equations, $m > n$, and the columns product matrix $(\mathbf{AA}')$ is non-singular, the solution for $\mathbf{X}$ is given by

$$\mathbf{X} = \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{C}$$

for which

$$L_{\mathbf{X}} = \text{a minimum.}$$

- The matrix $(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is termed a **left quasi-inverse**, having the property that

$$(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{A} = \mathbf{I}$$

and the matrix $\mathbf{A}'(\mathbf{AA}')^{-1}$ is termed a **right quasi-inverse**, having the property that

$$\mathbf{AA}'(\mathbf{AA}')^{-1} = \mathbf{I}.$$

Linear Transformations

- A transformation matrix, $\mathbf{T}$, is square and non-singular.
- Any matrix $\mathbf{A}$ can be transformed into a matrix $\mathbf{B}$ by multiplication by a transformation matrix $\mathbf{T}$. If $\mathbf{A}$ is premultiplied by $\mathbf{T}$

$$\mathbf{TA} = \mathbf{B}$$

  a row transformation has been performed and if $\mathbf{A}$ is postmultiplied by $\mathbf{T}$

$$\mathbf{AT} = \mathbf{B}$$

  a column transformation has been performed.
- A group of vectors produced by a transformation of another group of vectors will span the identical subspace spanned by the original group of vectors.

Linear Transformations

- A group of non-homogeneous equations

$$\mathbf{AX} = \mathbf{C}$$

may be row transformed into an equivalent group of equations by

$$\mathbf{A_T} = \mathbf{TA}$$

$$\mathbf{C_T} = \mathbf{TC}$$

then

$$\mathbf{A_T X} = \mathbf{C_T}$$

- If a square matrix $\mathbf{A}$ is transformed into an identity matrix by a series of transformations, the corresponding transformations on an identity matrix will produce the inverse of $\mathbf{A}$. If

$$(\mathbf{T_{R1}} \cdot \mathbf{T_{R2}} \ldots) \, \mathbf{A} \, (\mathbf{T_{C1}} \cdot \mathbf{T_{C2}} \ldots) = \mathbf{I}$$

then

$$(\mathbf{T_{R1}} \cdot \mathbf{T_{R2}} \ldots) \, \mathbf{I} \, (\mathbf{T_{C1}} \cdot \mathbf{T_{C2}} \ldots) = \mathbf{A}^{-1}$$

Linear Transformations

- An orthogonal transformation matrix $\mathbf{U}$ is a transformation matrix such that
$$\mathbf{U}' = \mathbf{U}^{-1}.$$
- The row vectors (or column vectors) contained in an orthogonal transformation matrix, $\mathbf{U}$, are of unit length and are mutually orthogonal so that:
$$\mathbf{U}\mathbf{U}' = \mathbf{U}'\mathbf{U} = \mathbf{I}.$$
- A columns (or rows) transformation by an orthogonal transformation matrix of a matrix of row (or column) vectors will produce an orthogonal rotation of axes.
- An orthogonal rotation of axes will not alter:
  - the lengths of the vectors in the space;
  - the inner product between vectors in the space;
  - the dimensionality of the subspace spanned by a group of vectors in the space.

Characteristic Roots and Vectors

- Almost any square matrix, say **A**, of order n, can be expressed as the product of three matrices of the following form:

$$\begin{array}{ccccccc} [a_{ij}] & = & [v_{im}] & \cdot & [\lambda_m] & \cdot & [v^{(mj)}] \\ \mathbf{A} & = & \mathbf{V} & \cdot & \mathbf{\Lambda} & \cdot & \mathbf{V}^{-1} \end{array}$$

  where $\mathbf{\Lambda}$ is a diagonal matrix with diagonal entries λ_m, **V** is a square matrix with entries v_{im}, and $\mathbf{V}^{-1}$ has entries indicated by superscripts v^{mj}.
- The λ_m are termed the **characteristic roots** of **A** and the column vectors, $\mathbf{V}_m$, in **V** are termed the **characteristic vectors** of **A**. Alternative terminology includes **eigenvalues** and **eigenvectors** or **latent roots** and **latent vectors** .
- The number of non-zero characteristic roots of **A** equals the rank of **A**.

Characteristic Roots and Vectors

- Each characteristic root, λ_m , and corresponding vector, $\mathbf{V}m$, is related to the matrix $\mathbf{A}$ in the following form:

$$\mathbf{A}\mathbf{V}_m = \lambda_m \mathbf{V}_m,$$

or

$$(\mathbf{A} - \lambda_m \mathbf{I})\mathbf{V}_m = 0$$

- The forms of (4) imply n homogeneous linear equations in the elements v_{im} of $\mathbf{V}_m$ which when manipulated algebraically to eliminate these elements v_{im} yield a polynomial equation in λ_m of rank n

$$\lambda_m^n + \alpha_{(n-1)}\lambda_m^{(n-1)} + \cdots + \alpha_t \lambda_m^t + \cdots + \alpha_0 = 0$$

where the αs are the coefficients of this equation and depend on the matrix $\mathbf{A}$. This equation is termed the characteristic equation of $\mathbf{A}$.

Characteristic Roots and Vectors

- Any characteristic vector may be multiplied by a constant and still retain its properties of (1) and (4).

  $$\mathbf{A}\mathbf{V}_m = \lambda_m \mathbf{V}_m,$$

  let

  $$\mathbf{V}^\star_m = c\mathbf{V}_m,$$

  then,

  $$\mathbf{A}\mathbf{V}^\star_m = \lambda_m \mathbf{V}^\star_m.$$

  By convention, a characteristic vector is usually defined to be of unit length.

  $$\mathbf{V}'_m \mathbf{V}_m = \mathbf{I}$$

- Solution of the equation of (1) for $\mathbf{\Lambda}$, yields the following form

  $$\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}$$

  which is termed the diagonalization of $\mathbf{A}$.

Characteristic Roots and Vectors

- Let $\mathbf{A}_t$ and $\mathbf{B}_t$ be matrices related to $\mathbf{A}$ by the equation

$$\mathbf{B}_t^{-1}\mathbf{A}\mathbf{B}_t = \mathbf{A}_t$$

and let $\mathbf{G}_{(t-1)t}$ be a matrix such that

$$\mathbf{G}_{(t-1)t}^{-1}\mathbf{A}_{(t-1)}\mathbf{G}_{(t-1)t} = \mathbf{A}_t$$

then

$$\mathbf{B}_{(t-1)}\mathbf{G}_{(t-1)t} = \mathbf{B}_t;$$

or, if there exists a series of matrices for $t = 1, 2, \ldots, s$ and

$$\mathbf{B}_0 = \mathbf{I},$$

then,

$$\mathbf{B}_s = \mathbf{G}_{0,1}\mathbf{G}_{1,2}\cdots\mathbf{G}_{(s-1),s}.$$

Characteristic Roots and Vectors

- If $\mathbf{G}_{(t-1)}$ is chosen for each t so that $\mathbf{A}_t$ is more nearly a diagonal matrix in some sense than is $\mathbf{A}_{(t-1)}$, then as $s \rightarrow \infty$

$$\mathbf{A}_s \rightarrow \mathbf{\Lambda}$$

and

$$\mathbf{B}_s \rightarrow \mathbf{V}$$

- For every characteristic root $\lambda_{m'}$, which does not equal any other characteristic root λ_m , $(m \neq m'; m = 1, 2, 3 \ldots, n)$ the corresponding characteristic vector is unique.

Characteristic Roots and Vectors

- If two or more characteristic roots, say $\lambda_p, \lambda_{(p+1)}, \ldots, \lambda_{(p+q-1)}$ for q roots, of matrix $\mathbf{A}$ are equal,

$$\lambda_p = \lambda_{(p+1)} = \cdots = \lambda_{(p+q-1)},$$

then a q dimensional subspace will be uniquely defined but the q vectors in any basis of this space will have the properties of characteristic vectors. Let one such group of vectors be column vectors in a matrix $\mathbf{V}_q$ and let $\mathbf{T}$ be a non-singular square matrix of order q, then

$$\mathbf{V}_q\mathbf{T} = \mathbf{V}_q^\star$$

yields a new group of vectors in $\mathbf{V}_q^\star$ which have the properties of characteristic vectors of $\mathbf{A}$.

Characteristic Roots and Vectors

- Addition of a scalar matrix with constant element k to a matrix $\mathbf{A}$ does not alter the characteristic vectors but adds the constant k to each root.

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$$

$$(\mathbf{A} + \mathbf{K}) = \mathbf{V}(\mathbf{\Lambda} + \mathbf{K})\mathbf{V}^{-1}.$$

- If two matrices, say $\mathbf{A}$ and $\mathbf{B}$, have identical characteristic vectors, then their sum will have the same characteristic vectors and will have characteristic roots equal to the sum of the corresponding roots for $\mathbf{A}$ and $\mathbf{B}$.

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}_A\mathbf{V}^{-1}$$

$$\mathbf{B} = \mathbf{V}\mathbf{\Lambda}_B\mathbf{V}^{-1}$$

$$(\mathbf{A} + \mathbf{B}) = \mathbf{V}(\mathbf{\Lambda}_A + \mathbf{\Lambda}_B)\mathbf{V}^{-1}.$$

Characteristic Roots and Vectors

- If a matrix **A** is multiplied by a scalar matrix **K** with constant element k, then the characteristic vectors will remain unchanged and the characteristic roots will be multiplied by the same constant.

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$$

$$(\mathbf{KA}) = \mathbf{V}(\mathbf{K}\boldsymbol{\Lambda})\mathbf{V}^{-1}.$$

- If a matrix **A** is raised to a power p by multiplying it by itself p times, the characteristic vectors will remain unchanged but the characteristic roots will be raised to the power p.

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$$

$$\mathbf{A}^2 = \mathbf{AA} = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^{-1}$$

$$\mathbf{A}^3 = \mathbf{AAA} = \mathbf{V}\boldsymbol{\Lambda}^3\mathbf{V}^{-1}, \text{ etc.}$$

Characteristic Roots and Vectors

- If a matrix $\mathbf{A}$ is inverted, then the characteristic vectors will remain unchanged but the characteristic roots will be inverted.
- If the matrix $\mathbf{A}$ is symmetric:

$$\mathbf{A} = \mathbf{A}',$$

  then the matrix $\mathbf{V}$ of characteristic vectors will be an orthogonal transformation so that

$$\mathbf{V}' = \mathbf{V}^{-1},$$

  and

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'.$$

- If the matrix $\mathbf{A}$ is symmetric $(\mathbf{A} = \mathbf{A}')$, then the trace of $\mathbf{A}$ equals the sum of the characteristic roots.

$$Trace(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{m=1}^{n} \lambda_m.$$

Characteristic Roots and Vectors

- If the matrix $\mathbf{A}$ is symmetric $(\mathbf{A} = \mathbf{A}')$, then the square of the norm of $\mathbf{A}$ equals the sum of the squares of the characteristic roots.

$$Norm^2(\mathbf{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 = \sum_{m=1}^{n} \lambda_m^2.$$

Characteristic Roots and Vectors

- If matrix $\mathbf{A}$ is symmetric ($\mathbf{A}' = \mathbf{A}$), let $\mathbf{B}_t$ be an orthogonal matrix

$$\mathbf{B}_t' = \mathbf{B}_t^{-1}$$

and let $\mathbf{A}_t$ be defined as

$$\mathbf{B}_t'\mathbf{A}\mathbf{B}_t = \mathbf{A}_t$$

then $\mathbf{A}_t$ is symmetric

$$\mathbf{A}_t' = \mathbf{A}_t,$$

and

$$Trace(\mathbf{A}_t) = Trace(\mathbf{A}),$$

$$Norm^2(\mathbf{A}_t) = Norm^2(\mathbf{A}).$$

Characteristic Roots and Vectors

- Further, let $\mathbf{G}_{(t-1)t}$ be an orthogonal matrix such that

$$\mathbf{G}'_{(t-1)t} = \mathbf{G}^{-1}_{(t-1)t}$$

$$\mathbf{G}'_{(t-1)t}\mathbf{A}_{(t-1)}\mathbf{G}_{(t-1)} = \mathbf{A}_t$$

then

$$\mathbf{B}_{(t-1)}\mathbf{G}_{(t-1)t} = \mathbf{B}_t$$

or, if there exists a series of matrices for $t = 1, 2, \ldots, s$ and

$$\mathbf{B}_0 = \mathbf{I}$$

then

$$\mathbf{B}_s = \mathbf{G}_{0,1}\mathbf{G}_{1,2}\mathbf{G}_{2,3}\cdots\mathbf{G}_{(s-1)s};$$

Characteristic Roots and Vectors

- and, if $\mathbf{G}_{(t-1)t}$ is chosen for each t so that

$$\sum_i \sum_{j \neq 1} a^2_{(t-1)ij} > \sum_i \sum_{j \neq 1} a^2_{tij}$$

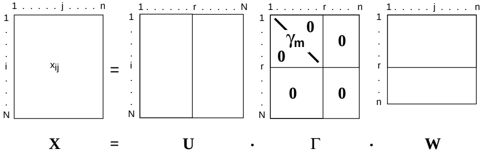
(these sums being of the squares of off-diagonal entries in $\mathbf{A}_{(t-1)}$ and $\mathbf{A}_t$), then as $s \to \infty$

$$\mathbf{A}_s \to \mathbf{\Lambda}$$

$$\mathbf{B}_s \to \mathbf{V}.$$

Principal Roots and Vectors

- Any rectangular matrix, say **X**, of order $N \times n$, can be expressed as the product of three matrices of the following form:



where **U** and **W** are orthogonal matrices so that

$$\mathbf{U}^{-1} = \mathbf{U}'$$
$$\mathbf{W}^{-1} = \mathbf{W}'$$

and $\mathbf{\Gamma}$ is zero except for a diagonal matrix of order r in the upper left corner with diagonal entries γ_m such that

$$\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \ldots \geq \gamma_r \geq 0$$

where $r \leq N$ and $r \leq n$.

Principal Roots and Vectors

- The columns of $\mathbf{U}$ will be called the **left principal vectors** of $\mathbf{X}$ and be designated $\mathbf{U}_m$; similarly, the rows of $\mathbf{W}$ will be called the **right principal vectors** of $\mathbf{X}$ and be designated $\mathbf{W}_m$ and, correspondingly, the γ_m will be called the **principal roots** of $\mathbf{X}$.
- For each $m = \{1 \dots r\}$ the product of the corresponding left principal vector, principal root, and right principal vector will produce a matrix $\mathbf{Y}_{(m)}$, of order $N \times n$, which will be designated the mth **principal component matrix** of $\mathbf{X}$.

$$\mathbf{Y}_{(m)} = \mathbf{U}_m \gamma_m \mathbf{W}_m.$$

Elements of $\mathbf{Y}_{(m)}$ will be designated $y_{(m)ij}$.

Principal Roots and Vectors

- Any two different principal component matrices of $\mathbf{X}$, say $\mathbf{Y}_{(m)}$ and $\mathbf{Y}_{(m^\star)}$, where m and $m^\star$ are unequal, are orthogonal in that every row vector in one is orthogonal to every row vector in the other and every column vector in one is orthogonal to every column vector in the other; thus:

$$\mathbf{Y}_{(m)}\mathbf{Y}'_{(m^\star)} = 0$$

$$\mathbf{Y}'_{(m)}\mathbf{Y}_{(m^\star)} = 0$$

- The matrix $\mathbf{X}$ is the sum of its principal component matrices,

$$\mathbf{X} = \sum_{m=1}^{r} \mathbf{Y}_{(m)}$$

Principal Roots and Vectors

- The square of the norm of a principal component matrix equals the square of the corresponding principal root.

$$[Norm(\mathbf{Y}_{(m)})]^2 = \gamma_m^2.$$

- The square of the norm of the matrix **X** is equal to the sum of the squares of the norm of the principal component matrices and to the sum of the squares of the principal roots.

$$[Norm(\mathbf{X})]^2 = \sum_{m=1}^{r} [Norm(\mathbf{Y}_{(m)})]^2 = \sum_{m=1}^{r} \gamma_m^2.$$

- The rank of matrix **X** equals r, the number of non-zero principal roots of **X**.

Principal Roots and Vectors

- Let ρ be an integer in the range

$$0 \leq \rho \leq r,$$

the ρth approximation to $\mathbf{X}$ is designated by $\hat{\mathbf{X}}_{(\rho)}$ and is defined as the sum of the first ρ principal component matrices:

$$\hat{\mathbf{X}}_{(\rho)} = \sum_{m=1}^{\rho} \mathbf{Y}_{(m)}.$$

- The error of approximating $\mathbf{X}$ by $\hat{\mathbf{X}}_{(\rho)}$ is designated by $\mathbf{E}_{(\rho)}$ so that

$$\mathbf{X} = \hat{\mathbf{X}}_{(\rho)} + \mathbf{E}_{(\rho)}.$$

which implies

$$\mathbf{E}_{(\rho)} = \sum_{m=\rho+1}^{r} \mathbf{Y}_{(m)}$$

Principal Roots and Vectors

- The square of the norm of $\mathbf{X}$ is resolved into the squares of the norms of $\mathbf{X}_{(\rho)}$ and $\mathbf{E}_{(\rho)}$ in a manner parallel to the resolution of $\mathbf{X}$ into $\mathbf{X}_{(\rho)}$ and $\mathbf{E}_{(\rho)}$ above; thus:

$$
\begin{aligned}
[Norm(\hat{\mathbf{X}}_{(\rho)})]^2 &= \sum_{m=1}^{\rho} \gamma_m^2 \\
[Norm(\mathbf{E}_{(\rho)})]^2 &= \sum_{m=\rho+1}^{r} \gamma_m^2 \\
[Norm(\mathbf{X})]^2 &= [Norm(\hat{\mathbf{X}}_{(\rho)})]^2 + [Norm(\mathbf{E}_{(\rho)})]^2
\end{aligned}
$$

- The matrix $\mathbf{X}_{(\rho)}$ provides a least squares approximation matrix of rank ρ to the given matrix $\mathbf{X}$.

Principal Roots and Vectors

- The characteristic roots and vectors of the rows product matrix of $\mathbf{X}$,

$$\mathbf{P}_{(i)} = \mathbf{X}\mathbf{X}'$$

are related to the principal roots and the left principal vectors of $\mathbf{X}$ in the following manner.

$$\mathbf{P}_{(i)} = \mathbf{U}\beta_{(N)}\mathbf{U}'$$

where $\beta_{(N)}$ is a diagonal matrix of order N containing the characteristic roots of $\mathbf{P}_{(i)}$ and $\mathbf{U}$ is the orthogonal matrix containing the characteristic vectors of $\mathbf{P}_{(i)}$, then

$$\begin{aligned} \beta_m &= \gamma_m^2 \quad \text{for} \quad m = 1, 2, 3, \ldots, r; \\ \beta_{m^*} &= 0 \quad \text{for} \quad m^* = (r+1), \ldots, N; \end{aligned}$$

and $\mathbf{U}$ contains the left principal vectors of $\mathbf{X}$.

Principal Roots and Vectors

- The characteristic roots and vectors of the columns product matrix of $\mathbf{X}$,

$$\mathbf{P}_{(j)} = \mathbf{X}'\mathbf{X}$$

are related to the principal roots and right principal vectors of $\mathbf{X}$ in the following manner.

$$\mathbf{P}_{(j)} = \mathbf{W}'\beta_{(n)}\mathbf{W}$$

where $\beta_{\mathbf{n}}$ is a diagonal matrix of order n containing the characteristic roots of $\mathbf{P}_{(j)}$ and $\mathbf{W}'$ is the orthogonal matrix containing the characteristic vectors of $\mathbf{P}_{(j)}$, then

$$\beta_m = \gamma_m^2 \quad \text{for} \quad m = 1, 2, 3, \ldots, r;$$
$$\beta_{m^{**}} = 0 \quad \text{for} \quad m^{**} = (r+1), \ldots, n;$$

and $\mathbf{W}$ contains the right principal vectors of $\mathbf{X}$.

Principal Roots and Vectors

- An orthogonal rotation of column axes of $\mathbf{X}$ by the matrix $\mathbf{W}'$ produces a matrix $\mathbf{Z}_{(c)}$,

$$\mathbf{X}\mathbf{W}' = \mathbf{Z}_c,$$

such that

$$\mathbf{Z}_c = \mathbf{U}\mathbf{\Gamma}$$

and the columns product matrix of $\mathbf{Z}_{(c)}$ equals the diagonal matrix $\beta_{(n)}$,

$$\mathbf{Z}'_{(c)}\mathbf{Z}_{(c)} = \beta_{(n)}.$$

Principal Roots and Vectors

- A similar orthogonal rotation of row axes of $\mathbf{X}$ may be performed with matrix $\mathbf{U}'$ to $\mathbf{Z_{(r)}}$ so that

$$\begin{aligned} \mathbf{U}'\mathbf{X} &= \mathbf{Z}_{(r)}, \\ \mathbf{Z}_{(r)} &= \mathbf{\Gamma W}, \\ \mathbf{Z}_{(r)}\mathbf{Z}'_{(r)} &= \beta_{(N)} \; . \end{aligned}$$

- The m^{**} columns of $\mathbf{Z}_{(c)}$, $m^{**} = (r+1), ..., n$, will contain all zero coordinates of the row vectors of $\mathbf{X}$. (Similarly, the m^* rows of $\mathbf{Z}_{(r)}$, $m^* = (r+1), ..., N$, will contain all zero coordinates of the column vectors of $\mathbf{X}$.)

Covariance of a Linear Combination

- Suppose we have an $N \times 2$ data matrix $\mathbf{X}$ where each column has a mean of zero.
- Further suppose we construct a multivariate linear combination $\mathbf{Y}$ such that

$$\begin{aligned} y_{i1} &= b_{11}x_{i1} + b_{21}x_{i2} + e_{i1} \\ y_{i2} &= b_{12}x_{i1} + b_{22}x_{i2} + e_{i2} \end{aligned}$$

which we write in matrix format as

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

where

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

and $\mathbf{E}$ is an $N \times 2$ matrix of residuals with zero mean.

Covariance of a Linear Combination

- We now wish to calculate the covariance matrix of the linear combination.
- Variance of a variable x_1 is the mean square deviation from the variable's mean

$$1/N \sum_{i=1}^{N} (x_1 - \bar{x_1})^2$$

- Covariance of two variables x_1 and x_2 is the mean square crossproduct from each variable's mean

$$1/N \sum_{i=1}^{N} (x_1 - \bar{x_1})(x_2 - \bar{x_2})$$

Covariance of a Linear Combination

- Since x_1 and x_2 had mean of zero, this reduces to.
  - Variance of x_1

$$1/N \sum_{i=1}^{N} x_1^2$$

  - Covariance of x_1 and x_2

$$1/N \sum_{i=1}^{N} x_1 x_2$$

Covariance of a Linear Combination

- Now, note that if we premultiply $\mathbf{Y}$ by its transpose

$$\mathbf{Z} = \mathbf{Y}'\mathbf{Y}$$

  we create a 2×2 matrix $\mathbf{Z}$, the columns product matrix of $\mathbf{Y}$.
- What are in the elements of $\mathbf{Z}$?
  - z_{11} contains the sum of squares of y_1
  - z_{22} contains the sum of squares of y_2
  - z_{12} and z_{21} contain the sum of crossproducts between y_1 and y_2.
- Now, if we premultiply $\mathbf{Z}$ by $1/N$ what is in the elements of $\mathbf{Z}$?
  - z_{11} contains the mean square of y_1
  - z_{22} contains the mean square of y_2
  - z_{12} and z_{21} contain the mean crossproducts between y_1 and y_2.
- If the variables of $\mathbf{Y}$ have mean zero, then $1/N(\mathbf{Y}'\mathbf{Y})$ is the covariance matrix of $\mathbf{Y}$.

Covariance of a Linear Combination

- So, since

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

let's try to substitute and work out the covariance of $\mathbf{Y}$.

$$\begin{aligned}
\mathbf{C_{YY}} &= 1/N(\mathbf{Y}'\mathbf{Y}) \\
&= 1/N((\mathbf{XB} + \mathbf{E})'(\mathbf{XB} + \mathbf{E})) \\
&= 1/N(((\mathbf{XB})' + \mathbf{E}')(\mathbf{XB} + \mathbf{E})) \\
&= 1/N((\mathbf{B}'\mathbf{X}' + \mathbf{E}')(\mathbf{XB} + \mathbf{E})) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB} + \mathbf{B}'\mathbf{X}'\mathbf{E} + \mathbf{E}'\mathbf{XB} + \mathbf{E}'\mathbf{E}) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB} + \mathbf{B}'0 + 0\mathbf{B} + \mathbf{E}'\mathbf{E}) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB} + \mathbf{E}'\mathbf{E}) \\
&= \mathbf{B}'(1/N\mathbf{X}'\mathbf{X})\mathbf{B} + 1/N\mathbf{E}'\mathbf{E} \\
&= \mathbf{B}'\mathbf{C_{XX}}\mathbf{B} + C_{EE}
\end{aligned}$$

Next Week

- Path Analysis.
- Components of Covariance.
- Running Mx.