Matrix Algebra Review

Daniel J. Bauer & Patrick J. Curran

CenterStat.org

Matrix Algebra Review

Daniel J. Bauer & Patrick J. Curran
CenterStat.org

Objectives

- Introduce the building blocks of matrix algebra in terms of scalars, vectors and matrices
- Define core mathematical operations for vectors and matrices
- Provide understanding of basic concepts to help with navigating matrix expressions in statistical models
  - Used widely in regression, multilevel modeling, factor analysis, SEM, mixture modeling, machine learning techniques, and beyond

CenterStat 2

Why Do We Need Matrices?

- Multivariate models can be expressed compactly using matrices
  - allows organization of large numbers of observations, variables, and parameters
- Understanding matrices allows us to better understand core statistical model that underlies nearly all quantitative methods
  - also to navigate technical resources; e.g., textbooks, computer manuals, online resources, etc.
- Some computer programs are set up solely in matrix form and many programs list warnings and errors in terms of matrices
  - e.g., "psi matrix is not positive definite", or "Hessian matrix cannot be inverted" or "defined model is inadmissible due to a singularity in sigma"
- Finally, matrices make the world a happier place

CenterStat 3

But do they really?

Matrices We Already Know

- A matrix is simply a doubly-ordered arrangement of numbers
- One example is raw data matrix
  - data matrix has *N*-rows (one for each observation) and *p*-columns (one for each variable)
- Another example is correlation matrix
  - we are all familiar with the correlation matrix in which ones are on the diagonal and all bivariate correlations are on the off-diagonal
- Even a calendar is a matrix
  - the rows are weeks and the columns are days
- We are thus already quite familiar with many forms of matrices

CenterStat 4

Scalars

- To differentiate from vectors and matrices, a single "ordinary" number is referred to as a scalar
- Scalars are typically denoted by lower-case italicized letters
- The algebra of scalars is arithmetic, and arithmetic provides the rules for operating on scalars

$$a = -6$$

$$b = 5$$

$$a + b = -6 + 5 = -1$$

CenterStat 5

As we will soon see, these are called "scalars" because they are often used to "re-scale" vectors and matrices (e.g., make things larger or smaller depending on our needs and purposes).

Matrix

- A matrix is a doubly-ordered arrangement of scalars
- The rows represent one set of categories
- The columns represent the other set of categories
- Matrices are usually denoted by bold capital letters

$$\mathbf{A} = \begin{bmatrix} 6 & 1 & 3 & 5 \\ -4 & 2 & 9 & 11 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} -5 & 3 \\ 2.6 & 43 \end{bmatrix}$$

CenterStat 6

The use of a bold capital letter to denote a matrix is arbitrary, but if we all agree on this notation then anywhere in the world when we see a bold capital $\mathbf{X}$ and know that represents a doubly-ordered matrix.

Order of a Matrix

- The *order* of a matrix refers to the number of rows and columns. For example, this matrix is of order "3 by 4:"

$$\mathbf{A}_{(3\times4)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

- In general, matrices are of order $r \times c$. The first number always denotes row & the second number denotes the second column
  - thus the mnemonic $r \times c$.

CenterStat 7

The Elements of a Matrix

- The scalars within a matrix are called *elements*.
- The element in row i and column j of matrix $\mathbf{A}$ is designated a_{ij} where the first subscript designates the row and the second the column.
- Two matrices are equal if and only if they are of the same order and all corresponding elements are equal

$$\mathbf{A} = \begin{bmatrix} 3 & 4 & 8 \\ 9 & 4 & 0 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & 4 & 8 \\ 9 & 4 & 0 \end{bmatrix} \qquad \mathbf{A} = \mathbf{B}$$

$$a_{13} = 8 \qquad b_{21} = 9$$

CenterStat 8

Think of a matrix as nothing more than a storage unit you might buy at Ikea: each row represents one set of categories (say, subjects) and each column represents another (say, a response to test item), and they jointly define a cubby hole in which we can store a given item response for a given individual.

The Transpose of a Matrix

- The transpose of a matrix is formed by interchanging the elements in each row and column.
- The transpose of matrix $\mathbf{A}$ $(r \times c)$ is designated $\mathbf{A}'$ or $\mathbf{A}^T$ $(c \times r)$.
- The first row simply becomes the first column, the second row the second column, and so on.

$$\mathbf{A}_{(r \times c)} = \begin{bmatrix} 6 & 2 & 4 \\ 8 & 1 & 0 \end{bmatrix} \qquad \mathbf{A}'_{(c \times r)} = \begin{bmatrix} 6 & 8 \\ 2 & 1 \\ 4 & 0 \end{bmatrix}$$

CenterStat 9

Symmetric Matrices

- A matrix is considered *symmetric* if the matrix is equal to the transpose of the same matrix: $\mathbf{A} = \mathbf{A}'$
- Note that a matrix must be *square* to be symmetric
  - a square matrix defined as having equal number of rows and columns; e.g., $c=r$
- Important cases: covariance, correlation, and distance matrices

$$\mathbf{R}_{(3\times 3)} = \begin{bmatrix} 1 & .2 & .4 \\ .2 & 1 & .3 \\ .4 & .3 & 1 \end{bmatrix} \qquad \mathbf{R}'_{(3\times 3)} = \begin{bmatrix} 1 & .2 & .4 \\ .2 & 1 & .3 \\ .4 & .3 & 1 \end{bmatrix}$$

CenterStat 10

When a matrix is symmetric, it is common for it to be displayed in what is called "lower-triangular form" to omit the redundant elements, i.e.,

$$\mathbf{R}_{(3\times 3)} = \begin{bmatrix} 1 & & \\ .2 & 1 & \\ .4 & .3 & 1 \end{bmatrix}$$

When presented in this way, it is assumed the upper-diagonal elements exist (they really are part of $\mathbf{R}$) but are simply not shown because of redundancy with the lower-diagonal elements.

Diagonal Matrices

- A particular kind of symmetric matrix is a diagonal matrix
  - Contains non-zero values on the diagonal, zeros everywhere else
- For example, define the sample covariance matrix to be

$$\mathbf{S} = \begin{bmatrix} s_{11} & & & \\ s_{21} & s_{22} & & \\ s_{31} & s_{32} & s_{33} & \\ s_{41} & s_{42} & s_{43} & s_{44} \end{bmatrix}$$

- Then the diagonal matrix is simply

$$\mathbf{D} = diag(\mathbf{S}) = \begin{bmatrix} s_{11} & & & \\ 0 & s_{22} & & \\ 0 & 0 & s_{33} & \\ 0 & 0 & 0 & s_{44} \end{bmatrix}$$

CenterStat 11

The Identity Matrix

- An important diagonal matrix is the identity matrix
  - Ones on the diagonal and zeros everywhere else
- The identity matrix plays the same role in matrix algebra as the number 1 in arithmetic

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

CenterStat 12

Vectors: A Particular Type of Matrix

- Vectors are matrices that have one row or one column
- A vector of order $r \times 1$ is a *column vector*
- A vector of order $1 \times c$ is a *row vector*
- Vectors are usually denoted by a lower case bold letter that defines a column vector unless transposed to a row:

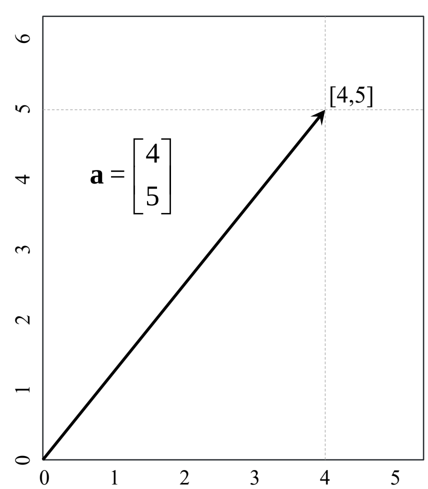
$$\mathbf{x}_{(r\times 1)} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} \qquad \mathbf{x}'_{(1\times c)} = \begin{bmatrix} x_1 & x_2 & \cdots & x_c \end{bmatrix}$$

- The prime ($'$) transposes $\mathbf{x}$ from a column to a row vector

CenterStat 13

Vector: Data Point in Space

- The term 'vector' literally describes a line projected in geometric space
- Indeed, much of multivariate statistics can be re-cast in geometric terms
- Vectors are represented by "arrows" that begin at the origin and terminate at the data point
- A vector of length two can easily be shown as a 2D plot; longer vectors generalize to higher-dimensions
  - gets hard to visualize



CenterStat 14

Vector Addition/Subtraction

- Two or more vectors of the same length can be added by adding corresponding elements

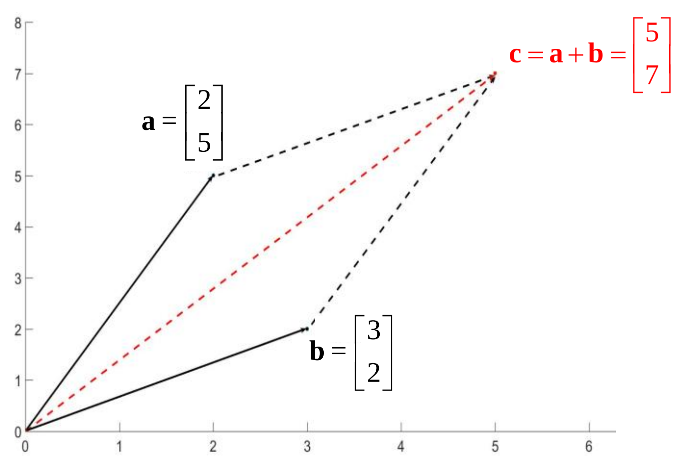
$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \qquad \mathbf{c} = \mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_N + b_N \end{bmatrix}$$

- Subtraction works the same way – just "adding" a negative number

CenterStat 15

Example: Vector Addition

- Geometrically, adding two vectors follows the "parallelogram law"
  - The resultant vector bisects the space between the two original vectors
  - The resultant vector is longer than either of the two original vectors



CenterStat 16

Matrix Operations: Addition

- Like vectors, when adding two matrices they must be of the same order
- Addition of two matrices is accomplished by adding corresponding elements, e.g., $c_{ij} = a_{ij} + b_{ij}$

$$\begin{bmatrix} 3 & 1 & 5 \\ 2 & 4 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 7 \\ 8 & 11 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 12 \\ 10 & 15 & 10 \end{bmatrix}$$

- Matrix addition is:
  - *commutative*: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
  - *associative*: $\mathbf{A} + \mathbf{B} + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- The resulting matrix always the same order as the matrices being added

CenterStat 17

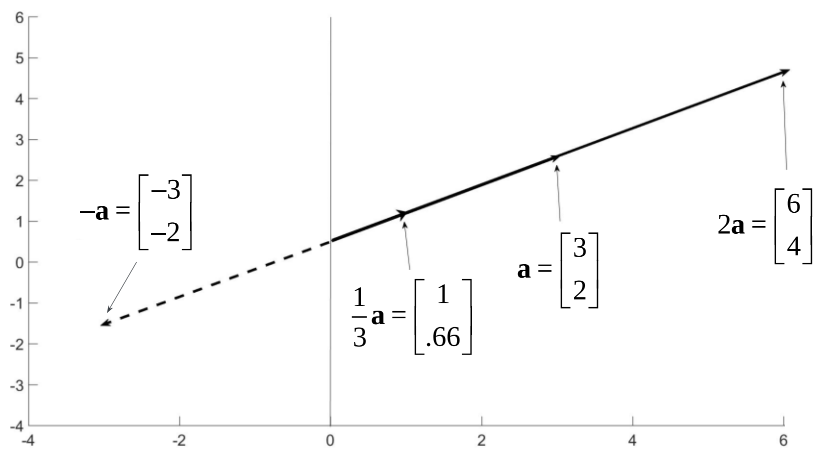
Multiplying a Vector by a Scalar

- Sometimes will see a vector multiplied by a scalar, e.g., $k\mathbf{a}$
- Each element within vector $\mathbf{a}$ is then multiplied by k
- Multiplying a vector by a scalar serves to either to stretch, shrink, or reflect the original vector
  - Most often encountered when standardizing or transforming variables
- For a scalar k and a vector $\mathbf{a}$ the following holds
  - $\mathbf{a}$ is stretched when $k > 1$
  - $\mathbf{a}$ is shrunken when $0 < k < 1$
  - $\mathbf{a}$ is reflected when $k < 0$

CenterStat 18

Earlier we noted the reason a number is called a "scalar" is that it rescales information in ways that are important to us; the final set of bullet points above formally defines this for scalar *k*.

Example: Vector × Scalar



CenterStat 19

Working with a vector of length two is helpful for visualizing in two dimensions but the ideas generalize to longer vectors in higher dimensions.

Combining Vector Addition and Scalar Multiplication

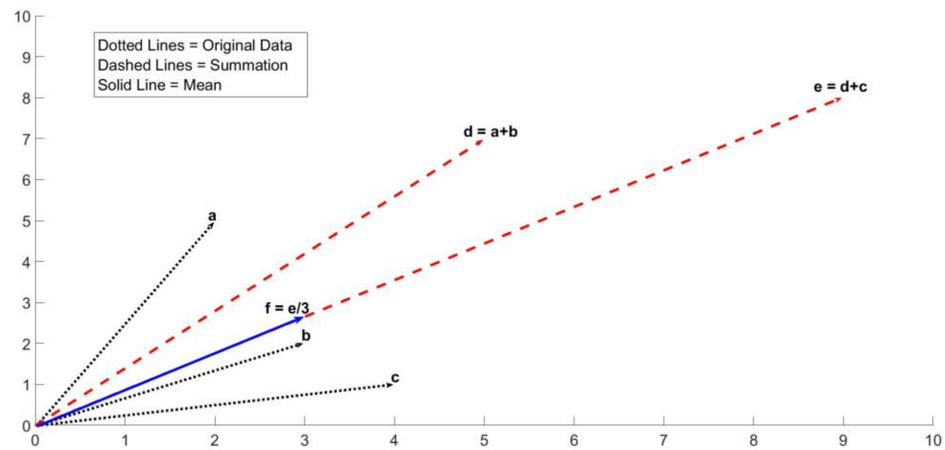
- Consider the vector of means:

$$\bar{\mathbf{x}} = \frac{\sum_{i=1}^{N} \mathbf{x}_i}{N}$$

- The numerator represents adding data vectors together
  - The summation centers the final vector among all the other vectors
- The denominator scales the vector
  - dividing by N shrinks the summed vector back into space of original vectors
  - because N is a scalar, multiplying by $1/N$ results in "elementwise" division
  - we will see in a moment that we must do additional work to divide vectors and matrices by one another

CenterStat 20

Example: Vector Mean

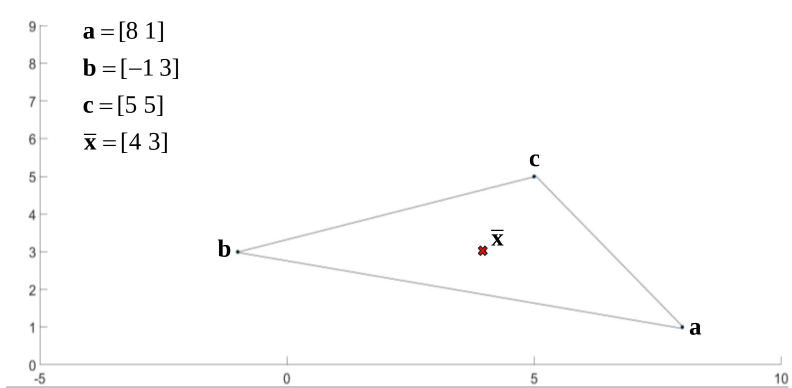


21

Here we are considering the mean vector for two variables, so that **a**, **b**, and **c** represent the values of two variables for three individual cases, **d** and **e** are intermediate steps that perform the summation of the individual data vectors, and **f** rescales the summed vector into the mean by multiplying by 1/3. The same thing occurs when there are three variables, but now with vectors in three-dimensional space. With more than three variables it gets hard to visualize but is conceptually the same.

Mean: Center of Mass

- The mean of a set of points can also be thought of as the center of mass of the points



22

Imagine you are balancing a "plate" defined by connecting the vectors on the end of a pencil – that's the center of mass, or mean vector. For this reason, the mean vector is sometimes called the "centroid"

Multiplying a Matrix by a Scalar

- Sometimes you will also see a matrix multiplied by a scalar, e.g., $k\mathbf{A}$
- Works just like vector × scalar case
- Simply multiply each element in $\mathbf{A}$ by scalar k
- For example, could specify a model in which the covariance matrices of two groups are restricted to be proportional via the constraint

$$\boldsymbol{\Sigma}_2 = k\boldsymbol{\Sigma}_1$$

- The relative pattern of variance and covariance would be the same across groups, but the magnitude of the elements (variances and covariances) in group 2 would be inflated by k (or deflated if $k < 1$)

CenterStat 23

This particular constraint, of proportional covariance matrices, comes up sometimes in finite mixture modeling. Other examples of matrices multiplied by scalars arise in other contexts. For instance, the residual covariance matrix for a typically linear regression is often represented $\sigma^2\mathbf{I}$ where $\mathbf{I}$ is an N x N identity matrix. This places σ^2 on the diagonal (indicating equal variance for all residuals or homoscedasticity), whereas all off-diagonal elements remain zero (indicating independence of the residuals across observations).

Matrix Multiplication

- Multiplying two matrices together is a common operation and a bit more complex compared to multiplying a matrix by a scalar
- ***Rule 1***: Only matrices of the form $(p \times q) * (q \times k)$ are conformable for multiplication.
  - columns in premultiplier must equal rows in postmultiplier.
- ***Rule 2***: The product matrix will have the following order:
  - $\mathbf{A}_{(p \times q)}\mathbf{B}_{(q \times k)} = \mathbf{C}_{(p \times k)}$
- ***Rule 3***: c_{ij} represents an element in row i, column j of the product matrix and results from the product of row i of the premultiplier with column j of the post multiplier
  - $c_{ij}=a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$

CenterStat 24

In scalar algebra, order of multiplication is not important: $ab = ba$. Likewise, order doesn't matter when multiplying a vector or matrix by a scalar. But when multiplying matrices together, the order of multiplication is critical: in general, $\mathbf{AB} \neq \mathbf{BA}$. Because the order matters, it is quite important to differentiate *pre-multiplication* by $\mathbf{A}$ (say, $\mathbf{AB}$) from *post-multiplication* by $\mathbf{A}$ (say, $\mathbf{BA}$).

Matrix Operations: Multiplication

$$\mathbf{A}_{(2\times3)}\mathbf{B}_{(3\times2)} = \mathbf{C}_{(2\times2)} = \begin{bmatrix} 2 & 4 & 1 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 8 \\ 19 & 17 \end{bmatrix}$$

$$c_{1,1} = (2)(1)+(4)(2)+(1)(4) = 14$$

$$c_{1,2} = (2)(3)+(4)(0)+(1)(2) = 8$$

$$c_{2,1} = (3)(1)+(0)(2)+(4)(4) = 19$$

$$c_{2,2} = (3)(3)+(0)(0)+(4)(2) = 17$$

CenterStat 25

Visualize running the index finger of your left across a given row of the first matrix while running the index finger of your right hand down the column of the second matrix, and summing the products of the pairs of matching elements along the way.

Matrix Operations: Division via Inverses

- You cannot divide one matrix by another (this operation is undefined), but you can do the equivalent by multiplying by the *inverse* of a matrix
- Remember that multiplying the inverse of a scalar with the original scalar equals one:

$$(5)(5^{-1}) = (5^{-1})(5) = 1$$

- Similarly, multiplying a matrix by its inverse results in the identity matrix:

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

- The inverse of the covariance/correlation matrix found in:
  - Mahalonobis' Distance
  - Multivariate normal density function
  - estimation of OLS regression weights

CenterStat 26

Matrix Operations: Inverses

- Say that matrix $\mathbf{A}$ and vector $\mathbf{b}$ were known and we must solve for vector $\mathbf{x}$ in the equation

$$\mathbf{Ax} = \mathbf{b}$$

- We cannot divide both sides by $\mathbf{A}$ because this is not defined in matrix algebra
- But we can compute the inverse of matrix $\mathbf{A}$ and multiply both sides by the inverse to isolate $\mathbf{x}$

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

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

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

CenterStat 27

The Determinant of a Matrix

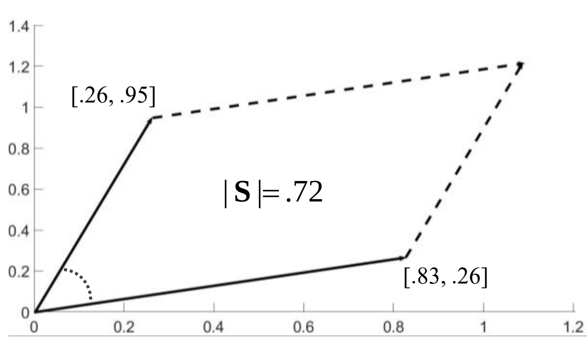
- To invert a matrix, say S, must first obtain its determinant, denoted |S|
- Determinant of covariance matrix is a scalar that reflects generalized variance
  - The larger the determinant, the greater the generalized variance
- Properties
  - Can only be computed for square matrices (e.g., correlation and covariance matrices)
  - The maximum value of the determinant of a correlation matrix is one
  - If $|\mathbf{S}|=0$, the matrix is said to be singular or "not of full rank"
  - Determinant is also the "volume" of the matrix

CenterStat 28

Warning: The notation for a determinant in matrix algebra looks just like the notation for an absolute value in scalar algebra, but they are not in any way related to one another.

Determinant: Geometric Interpretation

$$\mathbf{S} = \begin{bmatrix} .83 & .26 \\ .26 & .95 \end{bmatrix}; \ |\mathbf{S}| = .72$$



- For 2 x 2 matrix, determinant is area of parallelogram defined by the two vectors
  - With 3 x 3 would become volume, and so on
- Cosine of angle between the vectors reflects degree of correlation
- As angle becomes more acute, the determinant becomes smaller
  - Less generalized variance
  - At extreme vectors lie on top of each other and correlation is 1.0

CenterStat 29

Conceptualizing the determinant as the area (or volume, or hyper-volume) of a parallelogram is one of the coolest parts of geometric thinking. Picture the vectors above getting closer and closer to one another: the correlation between them gets higher and higher, and the determinant goes to zero.

Generalized Variance Properties

- The generalized variance is zero if and only if at least one column of the data matrix can be written as a linear combination of the others
- Consider a data matrix and its covariance matrix

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 5 \\ 4 & 1 & 6 \\ 4 & 0 & 4 \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} 3 & -1.5 & 0 \\ -1.5 & 1 & .5 \\ 0 & .5 & 1 \end{bmatrix}$$

- $|\mathbf{S}| = 0$. Why? Because $\mathbf{x}_3 = \mathbf{x}_1 + 2\mathbf{x}_2$
  - angle between vectors collapses because vectors lie on top of each other
- Problems can arise in many statistical models as determinant gets *near* zero

CenterStat 30

The Trace of a Matrix

- The trace of a matrix is the sum of the diagonal elements

$$Tr\left(\mathbf{A}_{(p\times p)}\right) = \sum_{i=1}^{p} a_{ii} = a_{11} + a_{22} + \cdots + a_{pp}$$

- Can be applied only to square matrices
- Example with a covariance matrix, results in the total variance in the data:

$$\mathbf{S} = \begin{bmatrix} 4.5 & & & \\ 1.2 & 3.1 & & \\ -.2 & -.1 & .4 & \\ -.4 & -.8 & .2 & 1.2 \end{bmatrix} \qquad Tr(\mathbf{S}) = 4.5 + 3.1 + .4 + 1.2 = 9.2$$

CenterStat 31

Computing Covariance Matrix

- Understanding the form of the covariance matrix in terms of matrices can provide insight to many procedures
- Assume that the data matrix $\mathbf{X}$ has p variables and that all the variables have been mean centered
  - the mean has been subtracted from each variable
- The sample covariance matrix can be computed as

$$\mathbf{S} = \frac{\mathbf{X}'\mathbf{X}}{N-1}$$

- The numerator is the "meat" of this equation and shows up repeatedly throughout statistics

CenterStat 32

Mean centering refers to subtracting the mean of each variable from the individual scores for that variables, i.e., for variable 1, $x'_{1i} = x_{1i} - \bar{x}_1$. Mean centering can be useful for a variety of purposes, but here it simplifies computing the covariance matrix. Recall the usual formula for a variance is of a single variable can be expressed in scalar algebra as

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{N-1} = \frac{\sum x_i'^2}{N-1}$$

We can see that the matrix analogue to this, producing the entire covariance matrix, is

$$\mathbf{S} = \frac{\mathbf{X}'\mathbf{X}}{N-1}$$

$\mathbf{X}'\mathbf{X}$ produces the sums of squares and cross products (SSCP) matrix for the data, which is then rescaled element-wise to produce the covariance matrix through scalar division by $N-1$.

Sometimes you may hear about the "corrected" versus "uncorrected" SSCP matrices. The corrected SSCP uses the mean-centered variables whereas the uncorrected SSCP is computed without mean centering.

Eigenvalues and Eigenvectors

- A particularly challenging topic involves eigenvalues and eigenvectors
- Eigenvalues and eigenvectors "repackage" a correlation or covariance matrix into orthogonal linear composites of the original variables
  - each eigenvalue reflects the variance of a composite formed using the associated eigenvector as weights
- There are as many eigenvalues as measured variables, and extracting them all simply repackages the original matrix and does us little good
  - but we can extract a *subset* of eigenvalues that repackages *most* of the covariance matrix and can thus be used for data reduction
- Eigen decomposition is complex and we won't detail this here, but eigenvalues and eigenvectors have nearly magical properties

CenterStat 33

Eigen-decompositions are critical for many things and are used in principal components analysis and factor analysis, visualization, cluster analysis, network analysis, mixture modeling, machine learning, etc. It is thus good to have some idea of what this is and how it is used.

Cool Facts About Eigenvalues

- Consider a $p \times p$ correlation (or covariance) matrix for p-variables

1. There are as many eigenvalues as there are observed variables
2. The sum of the eigenvalues equals the *trace* of the matrix
   - for a correlation matrix, this would be p, since diagonal is all 1's
3. The product of the eigenvalues equals the *determinant* of the matrix
4. The number of non-zero eigenvalues is the *rank* of the matrix

CenterStat 34

The "coolness" of these facts is a matter of some debate in the scientific community. (Patrick thinks they are pretty darn cool, but then Patrick doesn't have many friends).

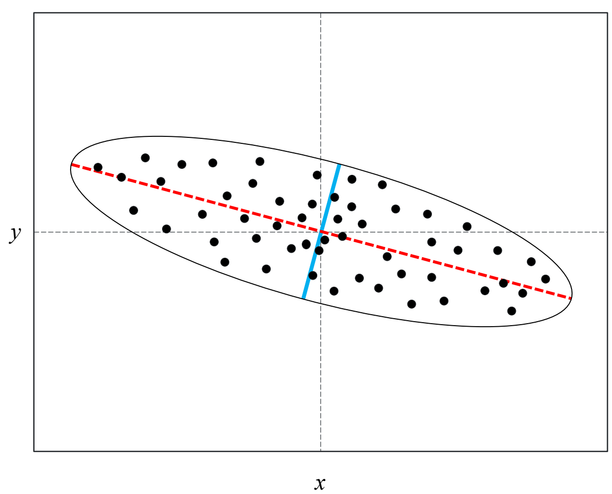
The rank of a matrix reflects the amount of non-redundant information present in that matrix. For instance, a 3 x 3 covariance matrix would be said to be of full rank if it has three non-zero eigenvalues. But if one variable is actually a linear combination of the other two, one of the eigenvalues would drop to zero and the rank would be two – there are really only two non-redundant columns in the matrix.

Interpretation for a Covariance Matrix

- The 1st eigenvector represents the direction of maximum variation and the 1st eigenvalue represents the amount of variation
- The 2nd eigenvector is the direction of maximum variation that is orthogonal to the 1st, and the 2nd eigenvalue represents its variation
- All eigenvectors are directions of variation, orthogonal to all other eigenvectors. The eigenvalues are their variances.
- Generally eigenvalues and eigenvectors found via principal components analysis (a data reduction method)

CenterStat 35

Visualizing Directions of Maximal Variation



Can visualize eigen-decomposition of covariance matrix for two variables, x and y

Direction of first principal component, red dashed, represents maximum direction of variation

Direction of second principal component, blue solid, represents maximum direction of variation orthogonal to first (at right angle to first)

Original x and y are correlated, but repackaged into orthogonal principal components

CenterStat 36

In a rather remarkable video, Greg Hancock pantomimes this in three dimensions using a croissant and bamboo skewers: https://www.youtube.com/watch?v=86ODrk1nB-g

Interpretation for Other Matrices

- Similar interpretations hold for matrices in other statistical models
- For instance, sometimes applied to a matrix of distances between observations (e.g., in multidimensional scaling, cluster analysis)
  - first eigenvector captures dimension of maximal variation in distances
  - second captures dimension of maximal variation remaining that is orthogonal to the first, etc.
- Also sometimes used with adjacency matrices in network analysis
  - e.g., eigenvector centrality
- Very common with correlation matrices in exploratory factor analysis
  - each eigenvalue represents a "factor" or "component" and the eigenvectors provide the item weights or "factor loadings"

CenterStat 37

Summary

- We need not be experts in matrix algebra, but it is useful to have a working knowledge
- A matrix is a doubly-ordered organizational structure for numbers
- Matrix algebra is a set of rules for applying mathematical functions to scalars, vectors and matrices
  - Can add, subtract, multiply and divide (via multiplication by an inverse) and solve for unknown values
- Matrices are used throughout all of multivariate statistics, sometimes "under the hood" but other times as a fundamental expression of the statistical model and the estimated parameters

CenterStat 38

A non-exhaustive list of further readings

Fieller, N. (2018). *Basics of matrix algebra for statistics with R*. Chapman and Hall/CRC.

Hohn, F. E. (2013). *Elementary matrix algebra*. Courier Corporation.

Kaw, A. (2008). *Introduction to matrix algebra*. Lulu.com.

Searle, S. R., & Khuri, A. I. (2017). *Matrix algebra useful for statistics*. John Wiley & Sons.

Related episodes of the podcast Quantitude (co-hosted by Patrick Curran and Greg Hancock)

S3E22: The Mättrix Part I: Defining & Manipulating Matrices

S3E23: The Mättrix Part II: Using Matrices to Our Advantage

S3E03: Principal Components Analysis is your PAL