Stat 608 – Chapter 1

Theme of The Class

- It makes sense to base inferences or conclusions only on valid models.
- A key step in any regression model, then, is to identify and address model weaknesses.
- There are two main parts of the class:
  1. Choose appropriate diagnostic procedures for building and assessing validity of regression models.
  2. Understand underlying mathematical properties of regression models in order to make appropriate decisions.

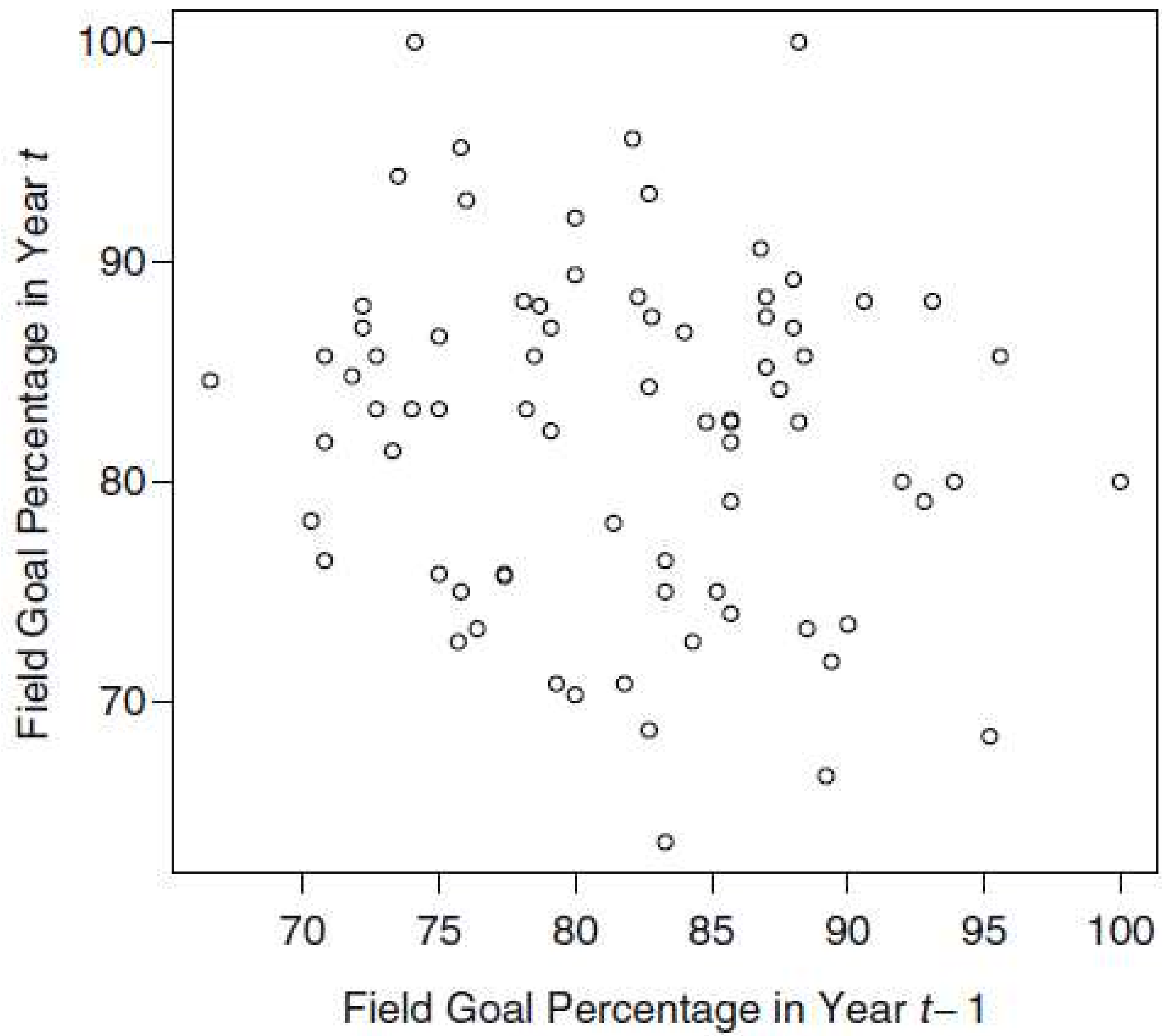
Example

In the Keeping Score column by Aaron Schatz in the Sunday November 12, 2006 edition of the *New York Times* *entitled "N.F.L. Kickers Are Judged on the* Wrong Criteria" the author makes the following claim:

There is effectively no correlation between a kicker's field goal percentage one season and his field goal percentage the next.

+

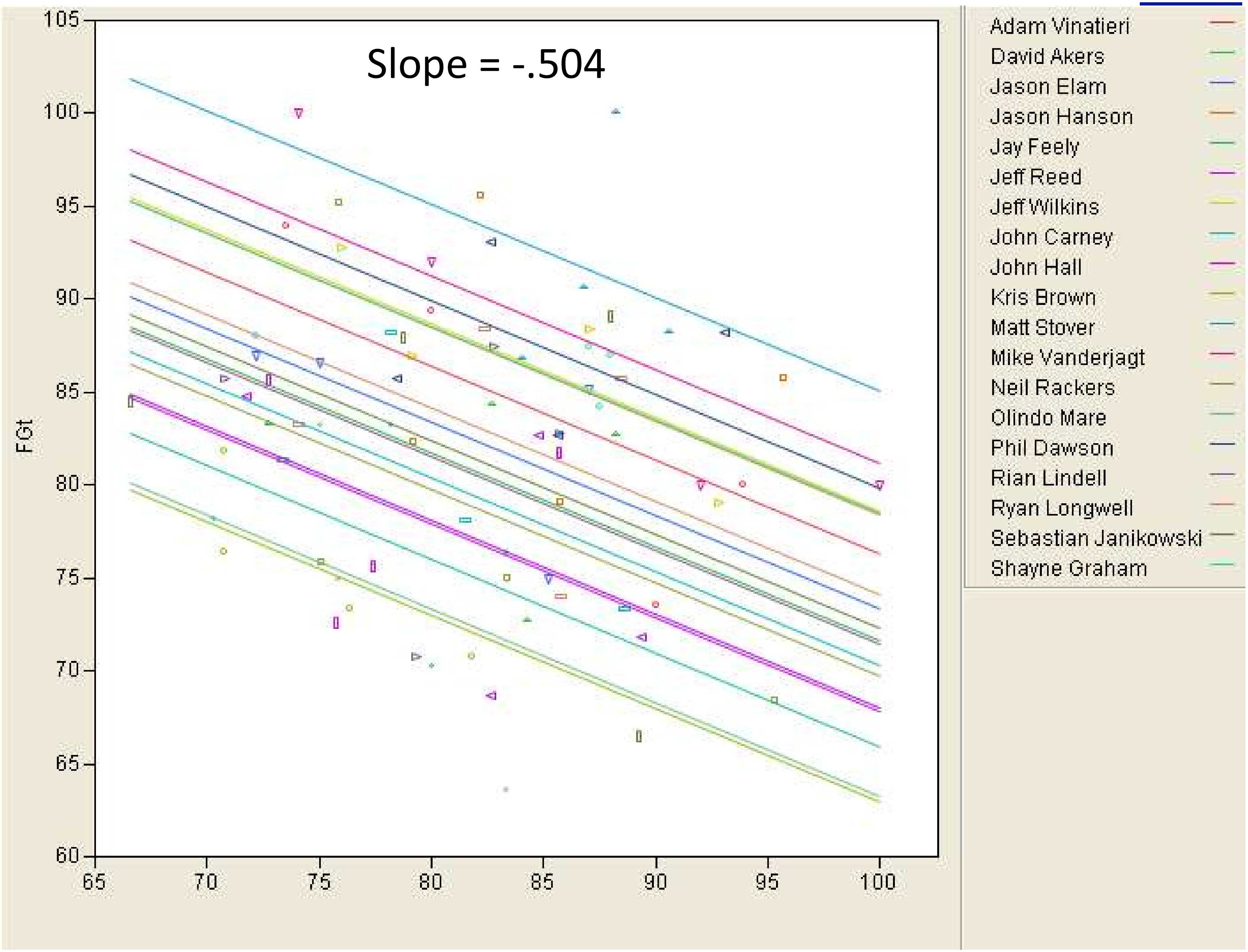
Slope = - .15



100
90
80
70
Field Goal Percentage in Year t
70
75
80
85
90
95
100
Field Goal Percentage in Year t– 1

Valid Model?

However, this approach is **fundamentally flawed as it fails to take into account the potentially** different abilities of the 19 kickers. In other words this approach is based on an **invalid model.**


Slope = -.504
FGt
105
100
95
90
85
80
75
70
65
60
65
70
75
80
85
90
95
100
Adam Vinatieri
David Akers
Jason Elam
Jason Hanson
Jay Feely
Jeff Reed
Jeff Wilkins
John Carney
John Hall
Kris Brown
Matt Stover
Mike Vanderjagt
Neil Rackers
Olindo Mare
Phil Dawson
Rian Lindell
Ryan Longwell
Sebastian Janikowski
Shayne Graham

Level of Mathematics

From former classes:

- Transpose, trace, determinant
- Addition / Subtraction
- Multiplication / Inverse
- Vector spaces, bases
- Logs
- Partial Derivatives

New material:

- Matrix Derivatives
- Expectation & Variance of Matrices

Design Matrix Setup: $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ - model

↳ A way of characterizing a linear model

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & & x_{nn} \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Intercept $\quad \underline{x_1} \quad \underline{x_2} \quad x_n$

$n \times (p+1)$ design matrix $\quad n \times 1$

$p \equiv$ # of predictor variables (covariates) (EV,)

Square Matrices

- A square matrix has the same number of rows and columns.

Square

$$\begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Not Square

$$\begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 0 & 1 \\ 2 & 4 \\ -2 & 0 \end{bmatrix}$$

Diagonal Elements

- The diagonal elements are the entries in the matrix from the top, left corner to the bottom, right corner:

$$\begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Symmetric Matrices

- A symmetric matrix is symmetric about the *diagonal*. That is, $a_{12} = a_{21}$, and $a_{23} = a_{32}$, etc. Symmetric matrices must be square.

- Which one is symmetric?

$$\begin{bmatrix} 1 & -2 & 0 \\ -2 & 3 & -5 \\ 0 & -5 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & -2 & 1 \\ 2 & -5 & 2 \\ 3 & -4 & 3 \end{bmatrix}$$

Symmetric Matrices

- The matrix X'X is symmetric.
- The variance-covariance matrix is symmetric:
  - Variances for variables 1, 2, …, p are found in the matrix at locations (1,1), (2,2), …, (p,p).
  - The covariance for variables i and j can be found at locations (i, j) and (j, i).
- The correlation matrix is also symmetric, having 1's on the diagonal and correlations on the off-diagonal.

+

Transpose

- To transpose A, use the rows of A for the columns of A'. (Or A^T, but the textbook uses A'.)

$$A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A' = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

- The equivalent of x^2 for a matrix X is X'X.

Trace

- The trace of a matrix A, tr(A), is the sum of the diagonal elements of A.

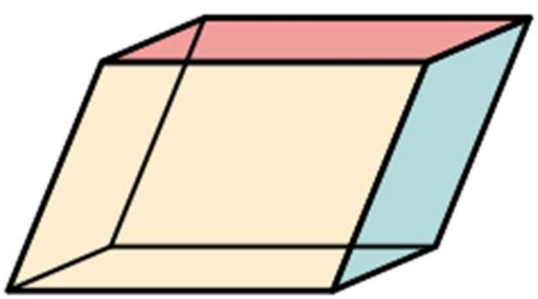
$$A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Tr(A) = 3 + 2 + 0 = 5

+

Determinant

- A determinant is a single value that characterizes a square matrix.
- The determinant is the volume of a parallelepiped formed by the vectors of the matrix.
- The determinant of a covariance matrix is the generalized variance of a set of variables.



Determinant

- The determinant of a 2X2 matrix with elements as follows is ad – bc:

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

- For example:

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

det(A) = |A| = 2(1) – 3(0) = 2

Determinant

- Determinant of a 3X3 matrix:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Determinant

- Example:

$$A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

- |A| = 3{2(0) – 1(1)} – 0{-1(0) – 1(0)} +

(-1){(-1)(1) – 2(0)} = -3 – 0 + 1= -2

+

Addition

- To add, just add each element:

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

- Matrices or vectors must be the same size to be added.

Subtraction

- To subtract, simply subtract each element:

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

- Matrices and vectors must be the same size to be subtracted.

Multiplication

- If A is a matrix, and c is a scalar, cA is found by multiplying c by every element of A:

$$c = 2$$

$$A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$cA = \begin{bmatrix} 6 & 0 & -2 \\ -2 & 4 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

Multiplication

- The dot product of two vectors is the sum of the products of the corresponding elements:

$$a = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, a \cdot b = -3 + 0 + 2 = -1$$

- Both vectors must be the same length for dot products to be defined.

- When the dot product of two vectors = 0, they are orthogonal.

- Often, we write a'b instead of the dot product.

Multiplication

- Matrix multiplication is more complicated than addition. To find the i^{th}, j^{th} element of matrix AB, take the dot product of the i^{th} row of A and the j^{th} column of B.

$$A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & -2 \end{bmatrix} \quad AB = \begin{bmatrix} 6 & -5 & 2 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- The number of columns of A must equal the number of rows of B.
- *Matrix multiplication is not commutative!*

+ START Friday 7/21/22 (Week 1, lecture 2)

~~Division~~ Inversion

- Actually, division for real numbers is multiplication by the inverse: instead of dividing by 2, multiply by ½.
- Inverses of diagonal matrices are easy:

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/5 \end{bmatrix}$$

~~Division~~ Inversion

- And inverting a 2X2 matrix isn't bad:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Rank

- The rank of a matrix X is:
  - The number of linearly independent column vectors of X.
  - The dimension of the vector space spanned (or generated) by the columns of X.
- A matrix is full rank if its rank is as large as possible (the minimum of the number of rows and the number of columns).
- A square matrix of full rank is invertible (nonsingular); if it is not of full rank, it is not invertible and is singular.
- A square matrix of full rank if and only if its determinant is non-zero.

Singular Matrices

- If the determinant of a matrix is 0, the inverse cannot be calculated, and the matrix is called *singular.*
- If two variables are perfectly correlated, one is a scalar multiple of the other, meaning there is a linear "combination" of one vector that gives the other. That is, perfectly correlated variables are not linearly independent, and X'X is singular.
- Indicator variables for *every* category of a categorical variable, plus an intercept, will make X'X singular.
- If there are more variables than observations, X'X will be singular.
- In linear models, if X'X is singular, the generalized inverse can be obtained, but the parameter estimates are not unique.

Commonly Used Matrices

- The multiplicative identity:

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

- The additive identity:

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

Projection

- To project vector **u** onto vector **v**, we use the formula:

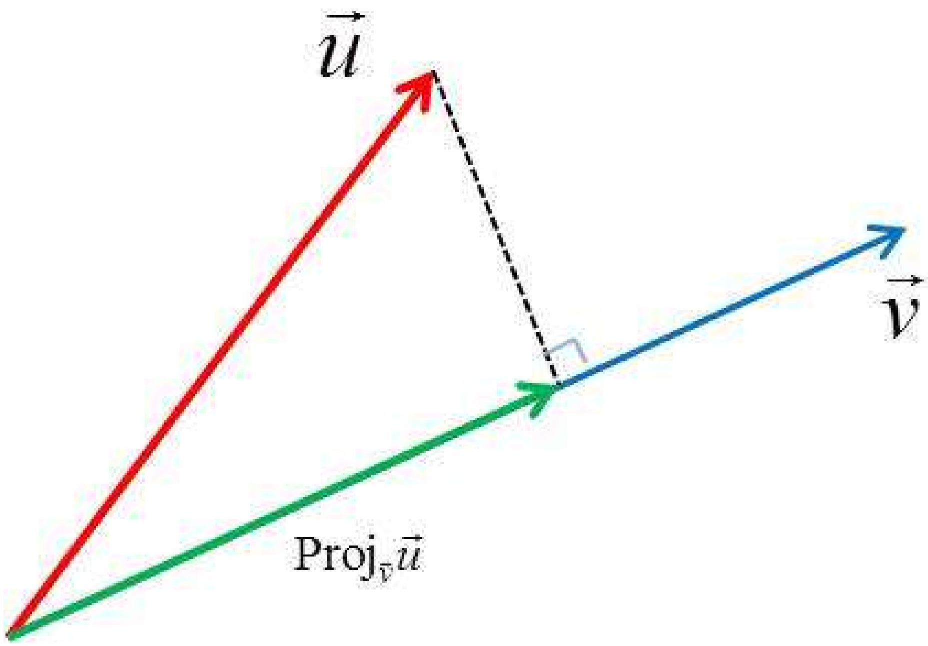
$$Proj_v u = P\mathbf{u}$$

where:

$$P = \frac{\mathbf{v}\mathbf{v}'}{\mathbf{v}'\mathbf{v}}$$

- If the matrix **V** contains vectors making up a basis for a subspace, then we can project onto the subspace using the projection matrix:

$$\mathbf{P} = \mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'$$

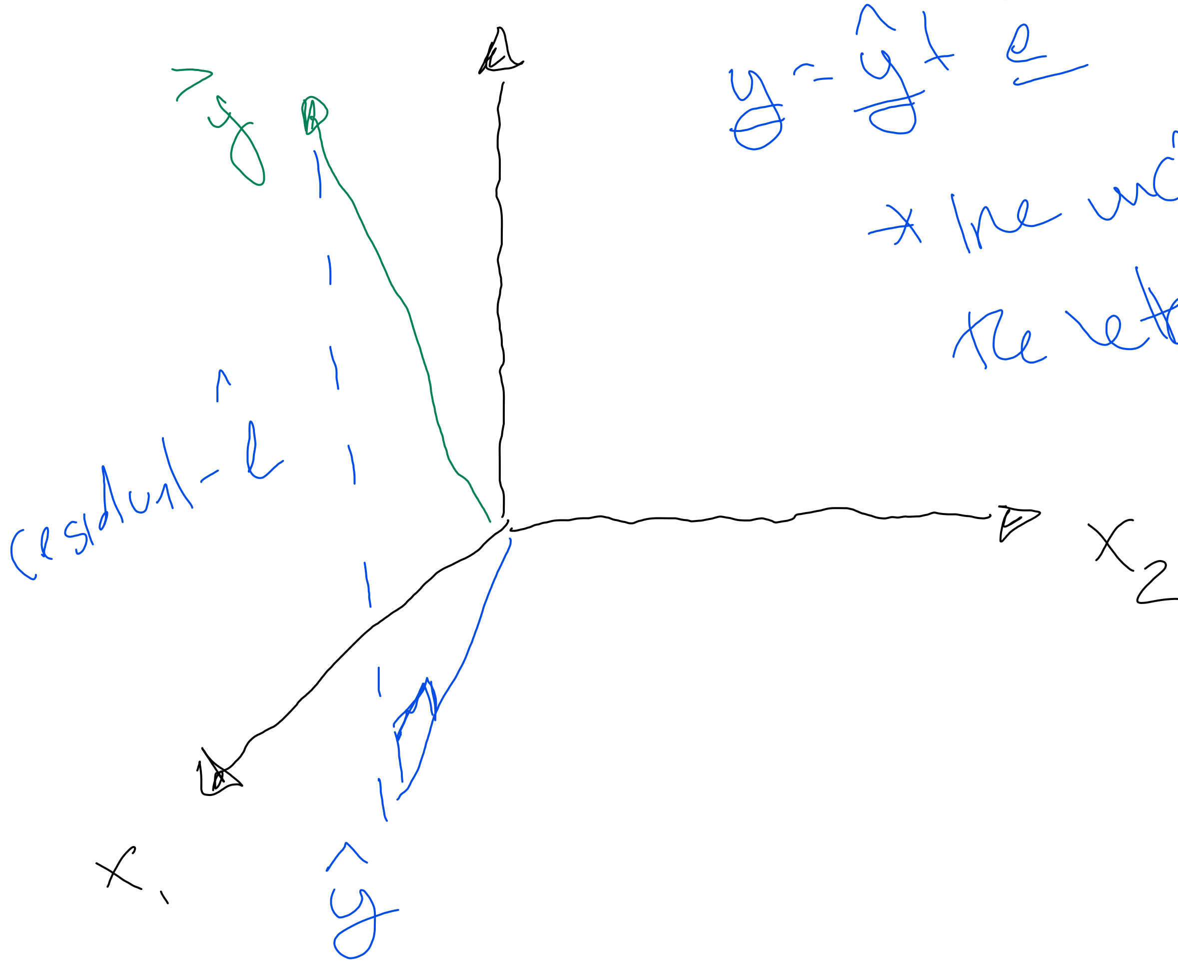


Projection

Let's work in 3-dimensional space, and suppose we want to project the vector **u** onto the subspace **V**:

$\underline{y} = \underline{\hat{y}} + \underline{\hat{e}}$

* line underneath the letters = vectors.

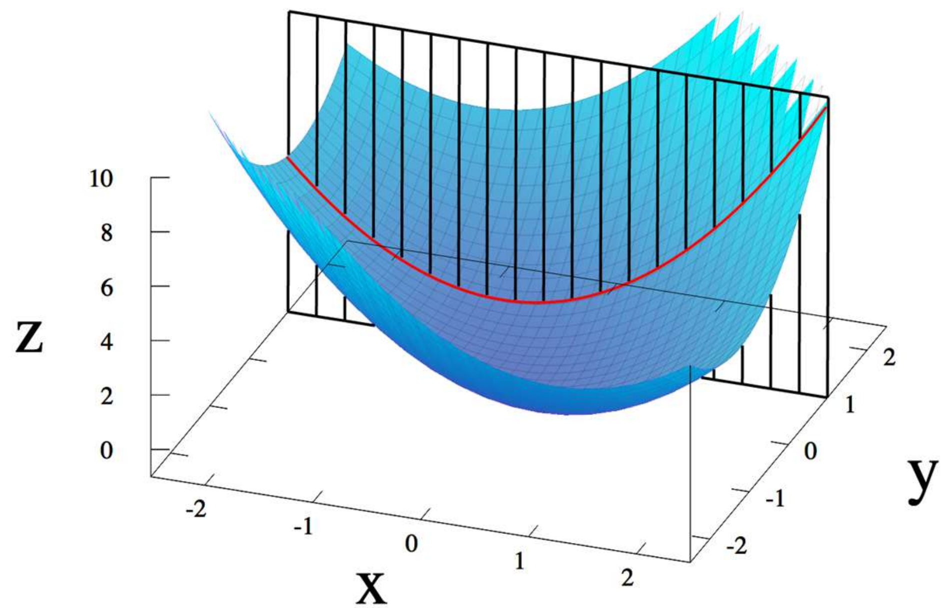


+ Partial Derivatives

- A **partial derivative** takes the derivative of a function with respect to only one variable, holding the other constant.

- The function at right is $f(x,y) = z = x^2 + xy + y^2$. To find its minimum, we can take partials with respect to both x and y, and set them equal to 0, giving us two equations with two unknowns.

- To take partial derivatives, assume that the other variable is another constant.



Partial Derivatives

$$f(x, y) = x^2 + xy + y^2$$

$$\frac{\partial f(x,y)}{\partial x} = 2x + y \stackrel{\text{set}}{=} 0$$

$$y = -2x$$

$$y = -2(0)$$

$$y = 0$$

$$\frac{\partial f(x,y)}{\partial y} = x + 2y \stackrel{\text{set}}{=} 0$$

$$x + 2(-2x) = 0$$

$$x - 4x = 0$$

$$-3x = 0$$

$$x = 0$$