# PubH 7405:
# REGRESSION ANALYSIS

## SL REGRESSION IN MATRIX TERMS

# BASIC REGRESSION MATRICES

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_n \end{bmatrix}$$

**X is special, called "Design Matrix"**

**The dimension of "Design Matrix" X could be changed to handle more than one predictor: more columns; the first column (filled with "1") is not included when doing "Regression through the origin" (i.e. no intercept).**

**X is called the Design Matrix.**

There are two reasons for the name:

**(1)** By the model, the values of X are under the controlled of investigators: **entries are fixed/designed,**

**(2)** The design/choice is **consequential**: the larger the variation in x's the more precise the estimate of the slope.

# REGRESSION OPERATIONS

$$\mathbf{Y}'\mathbf{Y} = [y_1 \quad y_2 \quad \cdots \quad y_n]\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = [\sum y_i^2]$$

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

**"Order" is important; cannot form YX'**

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_n \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

Here, we can form **XX'** but it is a different n-by-n matrix which we do not need.

If Y is a (column) vector, Y’Y is like a “Sum of Squares” – a  scalar. We’ll use it to form SSE

Y’Y, X’Y, and X’X form “sufficient statistics”

# INVERSE OF A 2x2 MATRICE

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

$$\mathbf{A}^{-1} = \begin{bmatrix} \dfrac{d}{D} & \dfrac{-b}{D} \\ \dfrac{-c}{D} & \dfrac{a}{D} \end{bmatrix}$$

$\mathrm{D} = (\mathrm{ad} - \mathrm{bc})$ is the "**determinant**" of A; or $|\mathrm{A}|$

A singular matrix does not have an inverse because its “determinants” is zero:

$$\mathbf{A} = \begin{bmatrix} 2 & 6 \\ 7 & 21 \end{bmatrix}$$

We have:

$$(3)\begin{bmatrix} 2 \\ 7 \end{bmatrix} + (-1)\begin{bmatrix} 6 \\ 21 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and: $D = (2)(21) - (6)(7) = 0$

# REGRESSION EXAMPLE

$$\mathbf{X'X} = \begin{bmatrix} n & \sum x \\ \sum x & \sum x^2 \end{bmatrix}$$

$$D = n\sum x^2 - (\sum x)(\sum x)$$

$$= n\sum (x - \bar{x})^2$$

$$\neq 0$$

$$
(\mathbf{X'X})^{-1} = \begin{bmatrix} \dfrac{\sum x^2}{n\sum (x-\bar{x})^2} & \dfrac{-\sum x}{n\sum (x-\bar{x})^2} \\ \dfrac{-\sum x}{n\sum (x-\bar{x})^2} & \dfrac{n}{n\sum (x-\bar{x})^2} \end{bmatrix}
$$

$$
= \begin{bmatrix} \dfrac{1}{n} + \dfrac{\bar{x}^2}{\sum (x-\bar{x})^2} & \dfrac{-\bar{x}}{\sum (x-\bar{x})^2} \\ \dfrac{-\bar{x}}{\sum (x-\bar{x})^2} & \dfrac{1}{\sum (x-\bar{x})^2} \end{bmatrix}
$$

Entries are functions of **mean and variance of X**

# SIMPLE LINEAR REGRESSION MODEL

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i ; i = 1,2,\ldots,n$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbf{Y}_{nx1} = \mathbf{X}_{nx2} \boldsymbol{\beta}_{2x1} + \boldsymbol{\varepsilon}_{nx1} ; \text{or}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

# BASIC MATRIX COMPONENTS

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

**Y** is the outcome/**dependent** (vector)

**X** is referred to as the "**Design Matrix**"

**ε** is the **error** (also a column vector)

# LEAST SQUARE METHOD

$$\text{Data}: \{(x_i, y_i)\}_{i=1}^{n}$$

$$Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

We solve to obtain normal equations:

$$\frac{\delta Q}{\delta \beta_0} = -2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0$$

$$\frac{\delta Q}{\delta \beta_1} = -2 \sum_{i=1}^{n} x_i (y_i - b_0 - b_1 x_i) = 0$$

**"Normal Equations" is a system of 2 equations with 2 unknowns: $b_0$ & $b_1$**

# “SSE” IN MATRIX NOTATION

Sum of squared errors :

$$Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

In matrix notation :

$$\begin{aligned} Q &= (\mathbf{Y} - \mathbf{X\beta})^{'}(\mathbf{Y} - \mathbf{X\beta}) \\ &= (\mathbf{Y'} - \mathbf{\beta' X'})(\mathbf{Y} - \mathbf{X\beta}) \\ &= \mathbf{Y'Y} - \mathbf{\beta' X' Y} - \mathbf{Y' X\beta} + \mathbf{\beta}^{'}\mathbf{X}^{'}\mathbf{X\beta} \\ &= \mathbf{Y}^{'}\mathbf{Y} - 2\mathbf{\beta}^{'}\mathbf{X}^{'}\mathbf{Y} + \mathbf{\beta}^{'}\mathbf{X}^{'}\mathbf{X\beta} \end{aligned}$$

**Recall:** **Y’Y** is like a Sum of Squares

**Note**: **Y’Xβ** is a **1x1 matrix** , hence equal to its transpose; that is **Y’X β = β’X’Y**

# NORMAL EQUATIONS

Normal Equations :

$$\sum \mathrm{y_i} = nb_0 + b_1 \sum x_i$$

$$\sum \mathrm{x_i y_i} = b_0 \sum x_i + b_1 \sum x_i^2$$

In matrix notations :

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i yi \end{bmatrix}$$

$$\mathbf{(X'X)b = X'Y}$$

We can prove the “normal equations” by taking derivative of SSE:

$$\mathbf{SSE = Y'Y-2}\beta\mathbf{'X'Y+}\beta\mathbf{'X'X}\beta$$

To do that, we need derivative of

(1) $\beta'[X'Y]$, and

(2) $\beta'[X'X]\beta$

Example #1:

$$\boldsymbol{\beta}'\mathbf{A} = \begin{bmatrix}\beta_0 & \beta_1\end{bmatrix}\begin{bmatrix}a \\ b\end{bmatrix}$$

$$= \begin{bmatrix}a\beta_0 + b\beta_1\end{bmatrix}$$

$$\frac{d}{d\boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}) = \begin{bmatrix}a \\ b\end{bmatrix}$$

$$= \mathbf{A}$$

(Example: A is $\mathbf{X}'\mathbf{Y}$, a vector)

Example #2 :

$$\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta} = \begin{bmatrix}\beta_0 & \beta_1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$$

$$= \begin{bmatrix}a\beta_0 + c\beta_1 & b\beta_0 + d\beta\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$$

$$= \left[\beta_0(a\beta_0 + c\beta_1) + \beta_1(b\beta_0 + d\beta)\right]$$

$$= [a\beta_0^2 + c\beta_0\beta_1 + b\beta_0\beta_1 + d\beta_1^2]$$

$$\frac{d}{d\boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta}) = \begin{bmatrix}2a\beta_0 + (c+b)\beta_1\\ (c+b)\beta_0 + 2d\beta_1^2\end{bmatrix}$$

$$= \left\{\begin{bmatrix}a & b\\ c & d\end{bmatrix} + \begin{bmatrix}a & c\\ b & d\end{bmatrix}\right\}\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$$

$$= 2\mathbf{A}'\boldsymbol{\beta} = 2\mathbf{A}\boldsymbol{\beta} \text{ if } \mathbf{A} \text{ is symmetric - like } \mathbf{X}'\mathbf{X}$$

Direct Derivation:

$$Q = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{'}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

$$= (\mathbf{Y}'\text{-}\boldsymbol{\beta}'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

$$= \mathbf{Y}^{'}\mathbf{Y} - 2\boldsymbol{\beta}^{'}\mathbf{X}^{'}\mathbf{Y} + \boldsymbol{\beta}^{'}\mathbf{X}^{'}\mathbf{X}\boldsymbol{\beta}$$

$$\frac{\boldsymbol{\delta}}{\boldsymbol{\delta\beta}}(Q) = \begin{bmatrix} \dfrac{\delta Q}{\delta \beta_0} \\ \dfrac{\delta Q}{\delta \beta_1} \end{bmatrix}$$

X’X is symmetric

$$= -2\mathbf{X}^{'}\mathbf{Y} + 2\mathbf{X}^{'}\mathbf{X}\boldsymbol{\beta}$$

**Note**:

$$= 0 \Leftrightarrow (\mathbf{X}^{'}\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y}$$

# LEAST SQUARE ESTIMATES

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i yi \end{bmatrix}$$

$$(\mathbf{X'X})\mathbf{b} = \mathbf{X'Y}$$

$$(\mathbf{X'X})^{-1}(\mathbf{X'X})\mathbf{b} = (\mathbf{X'X})^{-1}(\mathbf{X'Y})$$

$$\mathbf{b} = (\mathbf{X'X})^{-1}(\mathbf{X'Y})$$

# RANDOM VECTORS & MATRICES

A random vector or a random matrix contains **elements which are random variables.**

# RANDOM VECTORS IN SLR

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

# EXPECTED VALUES

$$\mathbf{Y} = \mathbf{E}(\mathbf{Y}) + \boldsymbol{\varepsilon}$$

$$\mathbf{E}(\mathbf{Y}) = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_n) \end{bmatrix}; \mathbf{E}(\boldsymbol{\varepsilon}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

# VARIANCE-COVARIANCE MATRIX

The variances (of elements of a random matrix) and the covariances between any two elements (of elements of a random matrix) are assembled in the **variance-covariance matrix** – denoted by either **Var(Y)** or $\boldsymbol{\sigma}^2(\mathbf{Y})$ – or $\sum$

# EXAMPLE: (BIVARIATE) VECTOR

$$\text{Variable}: \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\text{Mean}: \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\text{Variance} - \text{Covariance Matrix}: \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

We have for **random** **variables**:

$E(a_* Y) = a_* E(Y)$

$Var(a_* Y) = a^2{}_* var(Y)$

What about **random** **vectors**? Say:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\mathbf{AY} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}Y_1 + a_{12}Y_2 \\ a_{21}Y_1 + a_{22}Y_2 \end{bmatrix}$$

$$E(\mathbf{AY}) = \begin{bmatrix} E(a_{11}Y_1 + a_{12}Y_2) \\ E(a_{21}Y_1 + a_{22}Y_2) \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}E(Y_1) + a_{12}E(Y_2) \\ a_{21}E(Y)_1 + a_{22}E(Y_2) \end{bmatrix}$$

$$= \mathbf{AE(Y)}$$

$$\mathbf{AY} = \begin{bmatrix} a_{11}Y_1 + a_{12}Y_2 \\ a_{21}Y_1 + a_{22}Y_2 \end{bmatrix}$$

$$\sigma^2(\mathbf{AY}) = \begin{bmatrix} \sigma^2(a_{11}Y_1 + a_{12}Y) & \sigma(a_{11}Y_1 + a_{12}Y, a_{21}Y_1 + a_{22}Y) \\ \sigma(a_{11}Y_1 + a_{12}Y, a_{21}Y_1 + a_{22}Y) & \sigma^2(a_{21}Y_1 + a_{22}Y) \end{bmatrix}$$

$$= \cdots$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \sigma^2(Y_1) & \sigma(Y_1, Y_2) \\ \sigma(Y_1, Y_2) & \sigma^2(Y_2) \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$= \mathbf{A}\boldsymbol{\sigma}^2(\mathbf{Y})\mathbf{A}'$$

Can verify backward too starting with $\mathbf{A}\boldsymbol{\sigma}^2(\mathbf{Y})\mathbf{A}'$

# REGRESSION EXAMPLES

$$\mathbf{Y} = \mathbf{E}(\mathbf{Y}) + \boldsymbol{\varepsilon}$$

$$\mathbf{Var}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$$

$$\mathbf{Var}(\mathbf{Y}) = \begin{bmatrix} \sigma^2(Y_1) & \sigma(Y_1, Y_2) & \cdots & \sigma(Y_1, Y_n) \\ \sigma(Y_2, Y_1) & \sigma^2(Y_2) & \cdots & \sigma(Y_2, Y_n) \\ \vdots & \vdots & \vdots & \vdots \\ \sigma(Y_n, Y_1) & \sigma(Y_n, Y_2) & \cdots & \sigma^2(Y_n) \end{bmatrix}$$

$$= \sigma^2\mathbf{I}$$

$$\hat{\mathbf{Y}} = \begin{bmatrix} b_0 + b_1 x_1 \\ b_0 + b_1 x \\ \vdots \\ b_0 + b_1 x_n \end{bmatrix}$$

FITTED VALUES

$$= \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$= \mathbf{Xb}$$

# THE HAT MATRIX

$$\hat{\mathbf{Y}} = \mathbf{Xb}$$

$$= \mathbf{X}[(\mathbf{X'X})^{-1}\mathbf{X'Y}]$$

$$= [\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}]\mathbf{Y}$$

$$= \mathbf{HY}$$

$\mathbf{H} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$ is called the "**Hat Matrix**"

What is the dimension of **H**?

$$\hat{\mathbf{Y}} = [\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$$

$$= \mathbf{HY}$$

$\mathbf{X}$ is nx2

$(\mathbf{X}'\mathbf{X})^{-1}$ is 2x2

$\mathbf{X}'$ is 2xn

$\mathbf{H}$ is n-by-n

$$\begin{aligned}\hat{\mathbf{Y}} &= \mathbf{Xb} \\ &= \mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}] \\ &= [\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y} \\ &= \mathbf{HY}\end{aligned}$$

Recall: Each fitted value is a linear combination of responses

$$b_0 = \sum \left\{ \frac{1}{n} - \bar{x} k_i \right\} y_i$$

$$b_1 = \sum k_i y_i$$

$$\hat{Y} = b_0 + b_1 x$$

$$= \sum \left\{ \frac{1}{n} + (x - \bar{x}) k_i \right\} y_i$$

$$\hat{\mathbf{Y}} = \mathbf{HY}$$

$$h_{ij} = ?$$

# IDEMPOTENCY

the "**Hat Matrix**":

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

is "**idempotent**":

$$\mathbf{HH} = \mathbf{H}$$

# RESIDUALS

$$\hat{\mathbf{Y}} = \mathbf{Xb}$$
$$= \mathbf{HY}$$

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$$
$$= \mathbf{Y} - \mathbf{Xb}$$
$$= \mathbf{Y} - \mathbf{HY}$$
$$= (\mathbf{I} - \mathbf{H})\mathbf{Y}$$

Like the Hat Matrix **H**, (**I-H**) is symmetric & idempotent

# VARIANCE OF RESIDUALS

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$$

$$\boldsymbol{\sigma}^2(\mathbf{e}) = (\mathbf{I} - \mathbf{H})\boldsymbol{\sigma}^2(\mathbf{Y})(\mathbf{I} - \mathbf{H})'$$

$$= (\mathbf{I} - \mathbf{H})(\sigma^2\mathbf{I})(\mathbf{I} - \mathbf{H})'$$

$$= \sigma^2(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})'$$

$$= \sigma^2(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})$$

$$= \sigma^2(\mathbf{I} - \mathbf{H})$$

$$\overset{\wedge}{=} MSE(\mathbf{I} - \mathbf{H})$$

# REGRESSION COEFFICIENTS

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$= [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$$

$$= \mathbf{A}\mathbf{Y}$$

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

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

$$\mathbf{b} = [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y})$$

$$= \begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2} & \frac{-\bar{x}}{\sum (x-\bar{x})^2} \\ \frac{-\bar{x}}{\sum (x-\bar{x})^2} & \frac{1}{\sum (x-\bar{x})^2} \end{bmatrix} \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$= \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

Can verify: same results, eg.

$$b_1 = \frac{-\bar{x}\sum y + \sum xy}{\sum (x-\bar{x})^2}$$

$$= \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

## VARIANCE OF REGRESSION COEFFICIENTS

$$\mathbf{b} = \mathbf{A}\mathbf{Y}$$

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

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

$$\boldsymbol{\sigma}^2(\mathbf{b}) = \mathbf{A}\boldsymbol{\sigma}^2(\mathbf{Y})\mathbf{A}'$$

$$= \mathbf{A}\sigma^2\mathbf{I}\mathbf{A}'$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

$$\mathbf{s}^2(\mathbf{b}) = MSE(\mathbf{X}'\mathbf{X})^{-1}$$

$$\mathbf{s}^2(\mathbf{b}) = MSE(\mathbf{X}'\mathbf{X})^{-1}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2} & \frac{-\bar{x}}{\sum (x-\bar{x})^2} \\ \frac{-\bar{x}}{\sum (x-\bar{x})^2} & \frac{1}{\sum (x-\bar{x})^2} \end{bmatrix}$$

$$\mathbf{s}^2(\mathbf{b}) = MSE \begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2} & \frac{-\bar{x}}{\sum (x-\bar{x})^2} \\ \frac{-\bar{x}}{\sum (x-\bar{x})^2} & \frac{1}{\sum (x-\bar{x})^2} \end{bmatrix}$$

$$\mathbf{s}^2(\mathbf{b}) = MSE\begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2} & \frac{-\bar{x}}{\sum (x-\bar{x})^2} \\ \frac{-\bar{x}}{\sum (x-\bar{x})^2} & \frac{1}{\sum (x-\bar{x})^2} \end{bmatrix}$$

$$s^2(b_0) = MSE\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2}\right]$$

$$s^2(b_1) = \frac{MSE}{\sum (x-\bar{x})^2}$$

$$s(b_0, b_1) = \frac{-MSE(\bar{x})}{\sum (x-\bar{x})^2}$$

Same results for Variances;

Covariance is new - Only need Mean & Variance of X

# THE MEAN RESPONSE

**Recall the following**:

Let $X = x_h$ denote the level of X for which we wish to estimate the mean response, i.e. $E(Y|X=x_h)$; this $x_h$ may be a value which occurred in the sample, or it may be some other value of the predictor variable within the scope of the model. The point estimate of the response is:

$$E(Y \mid X = x_h) = \hat{Y}_h$$

$$= b_0 + b_1 x_h$$

Recall:

$$Var(\hat{Y}_h) = \sigma^2 \left\{ \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}$$

$$s^2(\hat{Y}_h) = MSE \left\{ \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}$$

# In matrix terms:

$$\hat{Y} = b_0 + b_1 x_h$$

$$= [1 \; x_h] \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$= \mathbf{X_h^{'} b}$$

$$\sigma^2(\hat{Y}) = \mathbf{X_h^{'} \boldsymbol{\sigma}^2(b) X_h}$$

$$= [1 \; x_h] \begin{bmatrix} \sigma^2(b_0) & \sigma(b_0, b_1) \\ \sigma(b_0, b_1) & \sigma^2(b_1) \end{bmatrix} \begin{bmatrix} 1 \\ x_h \end{bmatrix}$$

$$= \sigma^2(b_0) + 2x_h \sigma(b_0, b_1) + x_h^2 \sigma^2(b_1)$$

$$s^2(\hat{Y}) = s^2(b_0) + 2x_h s(b_0, b_1) + x_h^2 s^2(b_1)$$

$$= MSE\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x-\bar{x})^2}\right] + 2x_h\left[\frac{-MSE(\bar{x})}{\sum (x-\bar{x})^2}\right] + x_h^2\left[\frac{MSE}{\sum (x-\bar{x})^2}\right]$$

$$= MSE\left[\frac{1}{n} + \frac{(x_h-\bar{x})^2}{\sum (x-\bar{x})^2}\right]$$

Very same result!

# PRICTION OF NEW OBSERVATION

**Also recall the following:**

Let $X = x_h$ denote the level of X under investigation, at which the mean response is $E(Y|X=x_h)$. Let $Y_{h(new)}$ be the value of the new individual response of interest. The point estimate is still the same as that of the mean response $E(Y|X=x_h)$:

$$\hat{Y}_{h(new)} = b_0 + b_1 x_h$$

$$= \hat{Y}_h$$

$$Var(\hat{Y}_{h(new)}) = \sigma^2 \left\{ 1 + \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}$$

$$s^2(\hat{Y}_{h(new)}) = MSE \left\{ 1 + \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}$$

$$= MSE \{ 1 + \mathbf{X}_\mathbf{h}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_\mathbf{h} \}$$

$$= MSE + s^2(\hat{Y}_h)$$

# ANALYSIS OF VARIANCE

Recall that, in the decomposition:

$(Y_i - \overline{Y}) = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \overline{Y})$,

the first term reflects the variation around the regression line; the part than cannot be explained by the regression itself, and the second term measures of the **variation in Y associated with or explained by the regression line.**

$$SST = \sum (Y - \bar{Y})^2$$

$$= \sum (Y - \bar{Y})^2 = \sum [(Y - \hat{Y}) + (\hat{Y} - \bar{Y})]^2$$

$$= \sum (Y - \hat{Y})^2 + \sum (\hat{Y} - \bar{Y})^2 + 2\sum (Y - \hat{Y})(\hat{Y} - \bar{Y})$$

$$= SSE + SSR + 2\sum e_i \hat{Y}_i + 2\bar{Y} \sum e_i$$

$$= SSE + SSR$$

0

In matrix terms, we can write – with **J** being the square nxn matrix with all elements equal to 1:

$$SST = \mathbf{Y'Y} - (\frac{1}{n})\mathbf{Y'JY}$$

$$SSE = \mathbf{e'e}$$

$$= (\mathbf{Y} - \mathbf{Xb})(\mathbf{Y} - \mathbf{Xb})'$$

$$= \mathbf{Y'Y} - \mathbf{b'X'Y}$$

$$SSR = \mathbf{b'X'Y} - (\frac{1}{n})\mathbf{Y'JY}$$

# SSE: VERIFICATION

$$
\begin{aligned}
SSE &= \mathbf{e'e} \\
&= (\mathbf{Y} - \mathbf{Xb})(\mathbf{Y} - \mathbf{Xb})' \\
&= \mathbf{Y'Y} - \mathbf{b'X'Y}
\end{aligned}
$$

# QUADRATIC FORMS

A quadratic form is a polynomial containing terms involving the squares of the observations and the cross product. For example:

$(5Y_1^2 + 6Y_1Y_2 + 4Y_2^2)$.

In matrix term, we can express the above quadratic form as follow – with A being a symmetric matrix:

$$[Y_1 \; Y_2]\begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \mathbf{Y'AY}$$

All three ANOVA sums of squares (SST, SSR, and SSE) are quadratic forms (H is the hat matrix):

$$SST = \mathbf{Y}'[\mathbf{I} - (\frac{1}{n})\mathbf{J}]\mathbf{Y}$$

$$SSE = \mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y}$$

$$SSR = \mathbf{Y}'[\mathbf{H} - (\frac{1}{n})\mathbf{J}]\mathbf{Y}$$

**Example**:

SBP versus AGE

| Age (x) | SBP (y) | x-sq | y-sq | xy |
|---|---|---|---|---|
| 42 | 130 | 1764 | 16900 | 5460 |
| 46 | 115 | 2116 | 13225 | 5290 |
| 42 | 148 | 1764 | 21904 | 6216 |
| 71 | 100 | 5041 | 10000 | 7100 |
| 80 | 156 | 6400 | 24336 | 12480 |
| 74 | 162 | 5476 | 26244 | 11988 |
| 70 | 151 | 4900 | 22801 | 10570 |
| 80 | 156 | 6400 | 24336 | 12480 |
| 85 | 162 | 7225 | 26244 | 13770 |
| 72 | 158 | 5184 | 24964 | 11376 |
| 64 | 155 | 4096 | 24025 | 9920 |
| 81 | 160 | 6561 | 25600 | 12960 |
| 41 | 125 | 1681 | 15625 | 5125 |
| 61 | 150 | 3721 | 22500 | 9150 |
| 75 | 165 | 5625 | 27225 | 12375 |
| Totals 984 | 2193 | 67954 | 325929 | 146260 |

$$\mathbf{X'X} = \begin{bmatrix} n & \sum x \\ \sum x & \sum x^2 \end{bmatrix} = \begin{bmatrix} 15 & 984 \\ 984 & 67{,}954 \end{bmatrix}$$

$$D = (15)(67{,}954) - (984)^2 = 51{,}054$$

$$\mathbf{X'X}^{-1} = \begin{bmatrix} \dfrac{67{,}954}{51{,}064} & \dfrac{-984}{51{,}064} \\ \dfrac{-984}{51{,}064} & \dfrac{15}{51{,}064} \end{bmatrix}$$

$$= \begin{Bmatrix} 1.3310 & -.0193 \\ -.0193 & .0003 \end{Bmatrix}$$

$$\mathbf{X'X} = \begin{bmatrix} 15 & 984 \\ 984 & 67{,}954 \end{bmatrix}$$

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 1.3310 & -.0193 \\ -.0193 & .0003 \end{bmatrix}$$

$$\mathbf{X'Y} = \begin{bmatrix} 2{,}193 \\ 146{,}260 \end{bmatrix}$$

$$\mathbf{b} = (\mathbf{X'X})^{-1}(\mathbf{X'Y})$$

$$= \begin{bmatrix} 1.3310 & -.0193 \\ -.0193 & .0003 \end{bmatrix} \begin{bmatrix} 2{,}193 \\ 146{,}260 \end{bmatrix}$$

# VARIANCE OF REGRESSION COEFFICIENTS

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y})$$

$$\mathbf{s}^2(\mathbf{b}) = MSE(\mathbf{X}'\mathbf{X})^{-1}$$

$$= MSE\begin{bmatrix} 1.3310 & -.0193 \\ -.0193 & .0003 \end{bmatrix}$$

# SUMMARIES

- All results can be put in the matrix forms
- If we can inverse a matrix and can multiply two matrices, we can get all numerical results – even without a packaged computer program.
- In matrix their forms, results can be easier generalized; the only change needed is the Design Matrix ( & its dimension) so as to handle more than one predictors.

# Readings & Exercises

- Readings: A thorough reading of the text's sections 5.8-5.13 (pp.193-209) is highly recommended.
- Exercises: The following exercises are good for practice, all from chapter 5 of text: 5.5,5.6, and 5.24-5.26.

# Due As Homework

**14.1** Use the follow Weights (Y, kg) and Heights (X, cm):

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Height** | 162 | 168 | 174 | 176 | 180 | 180 | 182 | 184 | 186 | 186 |
| **Weight** | 65 | 65 | 84 | 63 | 75 | 77 | 82 | 65 | 80 | 91 |

to form these matrices: (a) Y'Y, (b) X'X, (c) X'Y, (d) Vector of estimated regression coefficients, (e) Vector of residuals, (f) SSR, (g) SSE, (h) Estimated variance-Covariance matrix of b.

**14.2** (optional) Answer the 8 questions of Exercise 14.1 using the following Weights and Heights of 10 women:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Height** | 152 | 156 | 158 | 160 | 162 | 162 | 164 | 164 | 166 | 168 |
| **Weight** | 52 | 50 | 47 | 48 | 52 | 55 | 54 | 59 | 60 | 65 |