SIMPLE LINEAR REGRESSION WITH MATRIX ALGEBRA

MODEL: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$

MATRIX MODEL: Y = XB + E

$$\text{or} \quad \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Where,

Y is the vector of dependent variables

XB is the linear model with data values and parameter estimates in separate matrices.

the "1"s correct for the intercept

E is the vector of random deviations or random errors

THE VALUES NEEDED FOR THE MATRIX SOLUTION ARE:

$Y'Y$ which is equivalent to USSY and equal to Y_i

$$Y'Y = [\,Y_1 \quad Y_2 \quad \dots \quad Y_n\,] \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$X'X$ which produces various intermediate sums

$$X'X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$$= \begin{bmatrix} n & \Sigma X_i \\ \Sigma X_i & \Sigma X_i^2 \end{bmatrix}$$

$X'Y$ which produces sum of Y and cross products

$$X'Y \;=\; X'X \;=\; \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix}$$

Using these three intermediate matrices ($Y'Y$, $X'X$ and $X'Y$) we can proceed with the matrix solutions.

The least squares solution for linear regression is based on the solutions of normal equations.

The normal equations for a Simple linear Regression are

$$\begin{aligned} nb_0 \;+\; \Sigma X_i b_1 \;&=\; \Sigma Y_i \\ \Sigma X_i b_0 \;+\; \Sigma X_i^2 b_1 \;&=\; \Sigma X_i Y_i \end{aligned}$$

x^2

which can be expressed as a matrix equation

$$\begin{bmatrix} n & \Sigma X_i \\ \Sigma X_i & \Sigma X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix}$$

from our previous calculations we can recognize this matrix equation as

$$(X'X)\,B = X'Y$$

where B is the vector of regression coefficients. In order to obtain the fitted model we must solve for B.

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

$$I*B = (X'X)^{-1}(X'Y)$$

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

where $A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

then $(X'X)^{-1} = \frac{1}{n\Sigma X_i^2 - (\Sigma X_i)^2} \begin{bmatrix} \Sigma X_i^2 & -\Sigma X_i \\ -\Sigma X_i & n \end{bmatrix}$

and

$$B = \begin{bmatrix} \frac{\Sigma X_i^2}{n\Sigma X_i^2 - (\Sigma X_i)^2} & \frac{-\Sigma X_i}{n\Sigma X_i^2 - (\Sigma X_i)^2} \\ \frac{-\Sigma X_i}{n\Sigma X_i^2 - (\Sigma X_i)^2} & \frac{n}{n\Sigma X_i^2 - (\Sigma X_i)^2} \end{bmatrix} * \begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

once we have fitted the line using the matrix equation, we proceed to obtain the Analysis of Variance table and our tests of hypothesis.

ANOVA TABLE in matrix form

	uncorrected		corrected	
Source	d.f.	USS	d.f.	SS
Regression	2	$B'X'Y$	1	$B'X'Y$ - CF
Error	n-2	$Y'Y-B'X'Y$	n-2	$Y'Y-B'X'Y$
Total	n	$Y'Y$	n-1	$Y'Y$-CF

The correction factor (CF) is the same as has been previously discussed. It can be calculated as either $\frac{(\sum_{i=1}^{n} Y_i)^2}{n}$ or $n\bar{Y}^2$

Additional calculations are

$$R^2 = \frac{B'X'Y\text{-CF}}{Y'Y\text{ - CF}}$$

the F test and t-test are given by $F = \frac{\text{MSRegression}}{\text{MSError}} = \frac{(B'X'Y - CF)/dfReg}{(Y'Y - B'X'Y)/dfError}$

The Variance - covariance matrix can be calculated as

$$MSE * (X'X)^{-1} = MSE* \begin{bmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{bmatrix} = \begin{bmatrix} MSE*c_{00} & MSE*c_{01} \\ MSE*c_{10} & MSE*c_{11} \end{bmatrix}$$

$$= \begin{bmatrix} S^2_{b_0} & S_{b_0b_1} \\ S_{b_1b_0} & S^2_{b_1} \end{bmatrix} \text{ or } \begin{bmatrix} S^2_{b_0} & S_{b_{01}} \\ S_{b_{10}} & S^2_{b_1} \end{bmatrix}$$

REVIEW OF SIMPLE LINEAR REGRESSION WITH MATRIX ALGEBRA

first obtain $Y'Y$, $X'X$, $X'Y$, and $(X'X)^{-1}$

then

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

$$SSReg = B'X'Y - CF$$

$$SSTotal = Y'Y - CF$$

$$SSE = SSTotal - SSReg = Y'Y - B'X'Y$$

$$\hat{Y}_L = \hat{Y}_L = L'B$$

$$S^2_{\hat{Y}_x} = L(X'X)^{-1}L'(MSE)$$

$$S^2_{\hat{Y}_x} = [1 + L(X'X)^{-1}L'](MSE) = L(X'X)^{-1}L'(MSE) + MSE$$

Variance-Covariance Matrix $= (X'X)^{-1}(MSE)$

Prediction of the true mean (population mean) and of new observations.

SLR

a) Predicting the true population mean at x_o

Model: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$

The true population mean at x_o is: $\beta_0 + \beta_1 x_o$

The predicted value ($\hat{Y}_o$) is: $b_0 + b_1 x_o = \bar{Y} + b_1(x_o - \bar{X})$

Variance (error) for $\hat{Y}_o$ is: $\sigma^2 \left[\frac{1}{n} + \frac{(x_o-\bar{X})^2}{S_{XX}} \right]$

Confidence Interval: $(b_0 + b_1 x_o) \pm t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(x_o-\bar{X})^2}{S_{XX}} \right]}$

b) Predicting a new observation at x_o

The true observation x_o is: $Y_o = \beta_0 + \beta_1 x_o + \epsilon_o$

The predicted value ($\hat{Y}_o$) is: $b_0 + b_1 x_o$

Variance (error) for $\hat{Y}_o$ is: $\sigma^2 \left[1 + \frac{1}{n} + \frac{(x_o-\bar{X})^2}{S_{XX}} \right]$

Confidence Interval: $(b_0 + b_1 x_o) \pm t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_o-\bar{X})^2}{S_{XX}} \right]}$

MLR - matrix algebra generalization

a) Predicting the true population mean

Model: $Y_i = X\beta + \epsilon$

The true population mean at x_o is: $X\beta$

The predicted value is: XB

Variance (error) is: $\sigma^2 X(X'X)^{-1}X'$

Confidence Interval: $x_0 b \pm t_{\frac{\alpha}{2}, n-p} \hat{\sigma}^2 X(X'X)^{-1}X'$

b) Predicting a observation

The true mean at x_o is: $Y_o = X_0\beta + \epsilon_o$

The predicted value is: x_0 b

Variance (error) is: $\sigma^2[1 + X(X'X)^{-1}X']$

Confidence Interval: $x_0 b \pm t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2[1 + X(X'X)^{-1}X']}$

More ANOVA table stuff

We have seen the partitioning of Y_i into components

SSTotal $=$ SSCF $+$ SSRegression $+$ SSError

ANOVA TABLE for SLR

SOURCE	d.f.	SS
Mean	1	$n\bar{Y}^2$
Regression	1	$\Sigma(\hat{Y}_i - \bar{Y})^2$
Error	n-2	$\Sigma(Y_i - \hat{Y}_i)^2$
Total (uncorrected)	n	ΣY_i^2

SS(Mean) = SS(b_0) $\equiv$ $SS_R(b_0)$

SS(Regression) = SS(b_1 | b_0) $\equiv$ $SS_R(b_1 \mid b_0)$

SS(Error) = SSE = SSResiduals

$SS_R(b_0,b_1) = b'X'Y$ = combined sum of squares for SSMean and SSReg

The usual ANOVA TABLE for SLR

SOURCE	d.f.	SS	
$SS_R(b_1 \mid b_0)$	1	$\Sigma(\hat{Y}_i - \bar{Y})^2$	MSReg $= \Sigma(\hat{Y}_i - \bar{Y})^2$ / 1
Error	n-2	$\Sigma(Y_i - \hat{Y}_i)^2$	MSError $= \Sigma(Y_i - \hat{Y}_i)^2$ / n-2
Total Corrected	n-1	$\Sigma(Y_i - \bar{Y})^2$	$S^2_{Y_i} = \Sigma(Y_i - \bar{Y})^2$ / n-1

Multiple Regression $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$

where $E(\epsilon_i) = 0$

then $E(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$

This model then predicts any point on a PLANE, where the axes of the plane are X_1 and X_2. The response variable, Y_i, gives the height above the plane.

If we choose any particular value of X_1 or X_2, then we essentially take a slice of the plane.

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i}$$

hold X_2 constant

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 C_{X2}$$
$$\hat{Y}_i = (b_0 + b_2 C_{X2}) + b_1 X_{1i}$$

Lets suppose we take that slice at a point where X_1=10 for the following model. We are then holding X_1 constant at 10, and examining the line (a slice of the plane) at that point.

$$\hat{Y}_i = 5 + 2X_{1i} + 3X_{2i}$$
$$\hat{Y}_i = 5 + 2*10 + 3X_{2i}$$
$$\hat{Y}_i = 25 + 3X_{2i}$$

This is a simple linear function. At every particular value of either X_1 or X_2, the function for the other X_k will be simple linear for this model.

NOTE that the interpretation for the regression coefficients is the same as before, except that now we have one regression coefficient per independent variable.

General Linear Regressions

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + ... + \beta_{p\text{-}1} X_{p\text{-}1i} + \epsilon_i$$

this is no longer a simple plane; it describes a hyperplane. However, we could still hold all X_k constant except one, and describe a simple linear function.

Calculations for Multiple Regression

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + ... + b_kX_{ki} + e_i$$

where there are k different independent variables

1) as before, the equation can be solved for e_i , and partial derivatives taken with respect to each unknown (b_0, b_1, b_2, . . .).

2) the partial derivatives can be set equal to zero, and solved simultaneously to get k+1 equations with k+1 unknowns.

3) the normal equations derived are;

$$\begin{array}{lllll} nb_0 + & \Sigma X_1b_1 & + \Sigma X_2b_2 & ... \Sigma X_kb_k & = \Sigma Y_i \\ \Sigma X_1b_0 + & \Sigma X_1^2b_1 & + \Sigma X_1X_2b_2 & ... \Sigma X_1X_kb_k & = \Sigma X_1Y_i \\ \Sigma X_2b_0 + & \Sigma X_1X_2b_1 & + \Sigma X_2^2b_2 & ... \Sigma X_2X_kb_k & = \Sigma X_2Y_i \\ \vdots \quad \vdots & \vdots & \vdots & \vdots & \vdots \\ \Sigma X_kb_0 + & \Sigma X_1X_kb_1 & + \Sigma X_2X_kb_2 & ... \Sigma X_k^2b_k & = \Sigma X_kY_i \end{array}$$

which can be factored out to an equation of matrices

$$\begin{bmatrix} n & \Sigma X_1 & \Sigma X_2 & & \Sigma X_k \\ \Sigma X_1 & \Sigma X_1^2 & \Sigma X_1X_2 & \dots & \Sigma X_1X_k \\ \Sigma X_2 & \Sigma X_1X_2 & \Sigma X_2^2 & \dots & \Sigma X_2X_k \\ \vdots & & \vdots & & \vdots \\ \Sigma X_k & \Sigma X_1X_k & \Sigma X_2X_k & \dots & \Sigma X_k^2 \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{1i}Y_i \\ \Sigma X_{2i}Y_i \\ \vdots \\ \Sigma X_{ki}Y_i \end{bmatrix}$$

Matrix calculations for General Regression : Numerical Example - NWK7.20

Mathematician salaries. X_1= Index of publication quality, X_2=years of experience, X_3=success in getting grant support.

$$X=\begin{bmatrix} 1 & 33.2 & 3.5 & 9.0 \\ 1 & 40.3 & 5.3 & 20.0 \\ 1 & 38.7 & 5.1 & 18.0 \\ 1 & 46.8 & 5.8 & 33.0 \\ 1 & 41.4 & 4.2 & 31.0 \\ 1 & 37.5 & 6.0 & 13.0 \\ 1 & 39.0 & 6.8 & 25.0 \\ 1 & 40.7 & 5.5 & 30.0 \\ 1 & 30.1 & 3.1 & 5.0 \\ 1 & 52.9 & 7.2 & 47.0 \\ 1 & 38.2 & 4.5 & 25.0 \\ 1 & 31.8 & 4.9 & 11.0 \\ 1 & 43.3 & 8.0 & 23.0 \\ 1 & 44.1 & 6.5 & 35.0 \\ 1 & 42.8 & 6.6 & 39.0 \\ 1 & 33.6 & 3.7 & 21.0 \\ 1 & 34.2 & 6.2 & 7.0 \\ 1 & 48.0 & 7.0 & 40.0 \\ 1 & 38.0 & 4.0 & 35.0 \\ 1 & 35.9 & 4.5 & 23.0 \\ 1 & 40.4 & 5.9 & 33.0 \\ 1 & 36.8 & 5.6 & 27.0 \\ 1 & 45.2 & 4.8 & 34.0 \\ 1 & 35.1 & 3.9 & 15.1 \end{bmatrix} \qquad Y=\begin{bmatrix} 6.1 \\ 6.4 \\ 7.4 \\ 6.7 \\ 7.5 \\ 5.9 \\ 6.0 \\ 4.0 \\ 5.8 \\ 8.3 \\ 5.0 \\ 6.4 \\ 7.4 \\ 7.0 \\ 5.0 \\ 4.4 \\ 5.5 \\ 7.0 \\ 6.0 \\ 3.5 \\ 4.9 \\ 4.3 \\ 8.0 \\ 5.0 \end{bmatrix}$$

Raw data matrices (X and Y) and the intermediate calculations (X′X, X′Y & Y′Y).

$$X'X = \begin{bmatrix} n & \sum_{i=1}^{n} X_1 & \sum_{i=1}^{n} X_2 & \sum_{i=1}^{n} X_3 \\ \sum_{i=1}^{n} X_1 & \sum_{i=1}^{n} X_1^2 & \sum_{i=1}^{n} X_1X_2 & \sum_{i=1}^{n} X_1X_3 \\ \sum_{i=1}^{n} X_2 & \sum_{i=1}^{n} X_1X_2 & \sum_{i=1}^{n} X_2^2 & \sum_{i=1}^{n} X_2X_3 \\ \sum_{i=1}^{n} X_3 & \sum_{i=1}^{n} X_1X_3 & \sum_{i=1}^{n} X_2X_3 & \sum_{i=1}^{n} X_3^2 \end{bmatrix}$$

$$= \begin{bmatrix} 24 & 948 & 128.6 & 599 \\ 948 & 38135.26 & 5188.17 & 24873.7 \\ 128.6 & 5188.17 & 727.44 & 3365.3 \\ 599 & 24873.7 & 3365.3 & 17847 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum_{i=1}^{n} Y \\ \sum_{i=1}^{n} X_1Y \\ \sum_{i=1}^{n} X_2Y \\ \sum_{i=1}^{n} X_3Y \end{bmatrix} = \begin{bmatrix} 143.5 \\ 5767.77 \\ 782.49 \\ 3671.9 \end{bmatrix}$$

$$Y'Y = \begin{bmatrix} \sum_{i=1}^{n} Y^2 \end{bmatrix} = \begin{bmatrix} 899.49 \end{bmatrix}$$

the normal equations derived are;

$$
\begin{array}{lllll}
nb_0 & + \Sigma X_1 b_1 & + \Sigma X_2 b_2 & + \Sigma X_k b_k & = \Sigma Y_i \\
\Sigma X_1 b_0 & + \Sigma X_1^2 b_1 & + \Sigma X_1 X_2 b_2 & + \Sigma X_1 X_k b_k & = \Sigma X_1 Y_i \\
\Sigma X_2 b_0 & + \Sigma X_1 X_2 b_1 & + \Sigma X_2^2 b_2 & + \Sigma X_2 X_k b_k & = \Sigma X_2 Y_i \\
\Sigma X_3 b_0 & + \Sigma X_1 X_3 b_1 & + \Sigma X_2 X_3 b_2 & + \Sigma X_3^2 b_3 & = \Sigma X_3 Y_i
\end{array}
$$

which can be factored out to an equation of matrices

$$
\begin{bmatrix} n & \Sigma X_1 & \Sigma X_2 & \Sigma X_3 \\ \Sigma X_1 & \Sigma X_1^2 & \Sigma X_1 X_2 & \Sigma X_1 X_3 \\ \Sigma X_2 & \Sigma X_1 X_2 & \Sigma X_2^2 & \Sigma X_2 X_3 \\ \Sigma X_3 & \Sigma X_1 X_3 & \Sigma X_2 X_3 & \Sigma X_3^2 \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{1i} Y_i \\ \Sigma X_{2i} Y_i \\ \Sigma X_{3i} Y_i \end{bmatrix}
$$

Analysis starts with the X′X inverse

$$(X'X)^{-1} = \begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 5.3478 & -0.1958 & 0.1486 & 0.0654 \\ -0.1958 & 0.008422 & -0.01215 & -0.002874 \\ 0.1486 & -0.01215 & 0.05088 & 0.002356 \\ 0.06541 & -0.002874 & 0.002356 & 0.001422 \end{bmatrix}$$

$$B = (X'X)^{-1}(X'Y) = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -4.511385879 \\ 0.3743477585 \\ -0.276494134 \\ -0.112439513 \end{bmatrix}$$

Analysis of Variance

USSTotal = $Y'Y$

USSRegression = $B'X'Y$

SSE = $Y'Y - B'X'Y$ = UCSSTotal – UCSSReg

The ANOVA table calculated with matrix formulas is

Source	d.f.	SS
Regression	p-1=3	$B'X'Y - CF$ =21.241
Error	n – p-1=20	$Y'Y - B'X'Y$ = 17.845
Total	n – 1=23	$Y'Y - CF$ = 39.086

where the correction factor is calculated as usual, $CF = \frac{(\Sigma Y)^2}{n} = n\bar{Y}^2$.

The F test of the model is a joint test of the regression coefficients,

$$H_0: \beta_1 = \beta_2 = \beta_3 = ... = \beta_{p-1} = \mathbf{0}$$

$$F = \frac{MSRegression}{MSError} = \frac{(B'X'Y - CF)/dfReg}{(Y'Y - B'X'Y)/dfError}$$

where

$$E(MSR) = \sigma^2 + \frac{\beta_1^2\ \Sigma(X_{1i}-\bar{X}_1)^2 + \beta_2^2\ \Sigma(X_{2i}-\bar{X}_2)^2 + 2\beta_1\beta_2\ \Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2)}{2}$$

NOTE: E(MSR) departs from σ^2 as β_k increases in magnitude (+ of –) or as any X_{ki} increases in distance from $\bar{X}_k$. The F test is a joint test for all β_k jointly equal 0.

To test any β_k individually, we can still use $t = \frac{(b_k - 0)}{s_{b_k}}$

where s_{b_k} is obtained from the VARIANCE – COVARIANCE matrix (below).

The confidence interval, for any β_k, is given by

$$P(b_k - t_{1-\frac{\alpha}{2},n-p}\, s_{b_k} \le \beta_k \le b_k + t_{1-\frac{\alpha}{2},n-p}\, s_{b_k}) = 1\text{-}\alpha$$

and the Bonferroni joint confidence interval for several β_k parameters is given by

$$P(b_k - t_{1-\frac{\alpha}{2g},n-p}\, s_{b_k} \le \beta_k \le b_k + t_{1-\frac{\alpha}{2g},n-p}\, s_{b_k}) = 1\text{-}\alpha$$

where “g" is the number of parameters

The VARIANCE – COVARIANCE matrix is calculated as from the $X'X^{-1}$ matrix.

$$(X'X)^{-1} = MSE \begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix}$$

$$= \begin{bmatrix} MSE*c_{00} & MSE*c_{01} & MSE*c_{02} & MSE*c_{03} \\ MSE*c_{10} & MSE*c_{11} & MSE*c_{12} & MSE*c_{13} \\ MSE*c_{20} & MSE*c_{21} & MSE*c_{22} & MSE*c_{23} \\ MSE*c_{30} & MSE*c_{31} & MSE*c_{32} & MSE*c_{33} \end{bmatrix} = \begin{bmatrix} s^2_{b_{00}} & s_{b_{01}} & s_{b_{02}} & s_{b_{03}} \\ s_{b_{10}} & s^2_{b_{11}} & s_{b_{12}} & s_{b_{13}} \\ s_{b_{20}} & s_{b_{21}} & s^2_{b_{22}} & s_{b_{23}} \\ s_{b_{30}} & s_{b_{31}} & s_{b_{32}} & s^2_{b_{33}} \end{bmatrix}$$

$$= \begin{bmatrix} Var(b_1) & Cov(b_1 b_2) & \dots & Cov(b_1 b_n) \\ Cov(b_2 b_1) & Var(b_2) & \dots & Cov(b_2 b_n) \\ \vdots & & & \\ Cov(b_n b_1) & Cov(b_n b_2) & \dots & Var(b_n) \end{bmatrix} = \text{Var-Cov matrix}$$

where the c_{ij} values are called Gaussian multipliers. The VARIANCE – COVARIANCE matrix is then calculated from this matrix by multiplying by the MSError.

These are unbiased estimates of σ^2_b

The individual values then provide the variances and covariances such that

$MSE*c_{00}$ = Variance of b_0 = $VAR(b_0)$

$MSE*c_{11}$ = Variance of b_1 = $VAR(b_1)$, so $s_{b_1} = \sqrt{MSE*c_{11}}$

$MSE*c_{01}$ = $MSE*c_{10}$ = Covariance of b_0 and b_1 = $COV(b_0,b_1)$

Prediction of mean response

For simple linear regression we got $\hat{Y}$ and its CI for some X_h

For multiple regression, we need an X_h for each X_j

Given a vector of $\underline{X}_h = \begin{bmatrix} 1 \\ X_{h1} \\ X_{h2} \\ \vdots \\ X_{h,p\text{-}1} \end{bmatrix}$

$E(Y_h) = \underline{X}_h\beta$

$\hat{Y}_h = \underline{X}_h B$

The variance estimates for mean responses are given by
$MSE{*}X(X'X)^{-1}X'$

for individual observations, add one MSE $MSE+MSE{*}X(X'X)^{-1}X'$

$$P(\hat{Y}_h - t_{1-\frac{\alpha}{2},n-p}\, s_{\hat{Y}_h} \leq E(\hat{Y}_h) \leq \hat{Y}_h + t_{1-\frac{\alpha}{2},n-p}\, s_{\hat{Y}_h}) = 1\text{-}\alpha$$

simultaneous estimates of several mean responses can employ either the

Working-Hotelling approach

$\hat{Y}_h \pm W s_{b_{\hat{Y}_h}}$ where $W^2 = pF_{1-\alpha;p,n-p}$

or the Bonferroni approach

$\hat{Y}_h \pm B s_{b_{\hat{Y}_h}}$ where $B = t_{1-\frac{\alpha}{2g};n-p}$

for individual observations the prediction is the same $\hat{Y}_h = \underline{X}'_h B$

and the variance is one MSE larger than for the mean

$$MSE + MSE*X(X'X)^{-1}X'$$

and for the mean of a new sample of size m, the variance is

$$\frac{MSE}{m} + MSE*X(X'X)^{-1}X'$$

As with the SLR, confidence intervals of g new observations can be done with

Scheffé limits

$$\hat{Y}_h \pm S s_{b_{\hat{Y}_h}} \quad \text{where } S^2 = gF_{1-\alpha;g,n-p}$$

or the Bonferroni approach

$$\hat{Y}_h \pm B s_{b_{\hat{Y}_h}} \quad \text{where } B = t_{1-\frac{\alpha}{2g};n-p}$$

Coefficient of Multiple Determination - the proportion of the SSTotal (usually corrected) accounted for by the Regression line (SSReg).

Models <u>with</u> an intercept

$$R^2 = \frac{SS_R(b_1,b_2,\ldots, b_k|b_0)}{S_{YY}} = \frac{SSRegression}{SSTotal\ (corrected)} = \frac{B'X'Y-CF}{Y'Y-CF} = \frac{SS_{YY} - SS_E}{S_{YY}} = 1-\frac{SS_E}{S_{YY}}$$

$$= \frac{21.24}{39.09} = 0.5434$$

Recalling the expressions for S_R we have

$$R^2 = \frac{b'X'Y - n\overline{Y}^2}{Y'Y - n\overline{Y}^2} = \frac{\Sigma(\hat{Y}_i - \overline{Y})^2}{\Sigma(Y_i - \overline{Y})^2} = 1 - \frac{SS_E}{SSTotal\ (UNCorrected)}$$

Some Properties of R^2 : Same as for SLR

1) $0 \leq R^2 \leq 1$

2) $R^2 = 1.0$ iff $\hat{Y}_i = Y_i$ for all i (perfect prediction, SSE = 0)

3) $R^2 = r^2_{XY}$ for simple linear regression

4) $R^2 = r^2_{\hat{Y}Y}$ for all models with intercepts

5) $R^2 \neq 1.0$ when there are different repeated values of Y_i at some value of X_i (no matter how well the model fits)

6) $R^2_{SubModel} \leq R^2_{FullModel}$

New independent variables added to a model will increase R^2. The R^2 for the full model could be EQUAL, but never less than the R^2 for the submodel

F test for Lack of Fit

$$E(Y) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_{p\text{-}1} X_{p\text{-}1,i}$$

To get true repeats in multiple regression, EVERY independent variable must remain the same from one observation to another

This can be calculated with either full and reduced model

New problems associated with MULTIPLE REGRESSION

The SLR fitted only a single slope, so we needed only one SSReg to describe it. With various slopes, we will need some other sums of squares to describe the various fitted slopes.

we will actually see 2 types of SS

non-problems associated with Multiple regression

a) All previous definitions and notation apply.

b) The assumptions are basically the same (more X_i, each measured without error).

Many of the tests of hypothesis will be discussed in terms of the General Linear Test, with appropriate Full and Reduced models.

This test does not really change with multiple regression, We still have the same table,

Model	d.f	SS	MS	F
Reduced (Error)	n-p	SSE_{Red}		
Full (Error)	n-p-q	SSE_{Full}		
Difference	q	SS_{Diff}	MS_{Diff}	$\frac{MS_{Diff}}{MSE_{Full}}$
Full (Error)	n-p-q	SSE_{Full}	MSE_{Full}	

Model	d.f	SS	MS	F
Full (SSReg)	p+q	SSR_{Red}		
Reduced (SSReg)	p	SSR_{Full}		
Difference	q	SS_{Diff}	MS_{Diff}	$\frac{MS_{Diff}}{MSE_{Full}}$
Full (Error)	n-p-q	SSE_{Full}	MSE_{Full}	