Review

X : design matrix

y : output

b : estimated model parameters

also

Φ : design matrix of "features"

where

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_2) & \cdots & \phi_p(x_1) \\ \vdots & & & \\ \phi_1(x_n) & \phi_2(x_n) & \cdots & \phi_p(x_n) \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix}$$

Least squares estimate (soln to both LS & ML)

$$b = (X'X)^{-1}X'y \qquad \text{(error in book)}$$

Estimated fitted values & residuals

$$\hat{y} = Xb \Rightarrow \hat{Y} = HY \quad \text{wher } H = X(X'X)^{-1}X'$$

residuals

$$e = (I - H)Y$$

and

$$\text{cov}(e) = (I-H)\,\text{cov}(Y)\,(I-H)'$$

where if $\varepsilon \sim N(0, \sigma^2 I)$, $\text{cov}(Y) = \sigma^2 I$

and

$$\text{cov}(e) = \sigma^2(I-H)$$

when we don't know σ^2 we replace it with MSE to yield

$$s^2\{e\} = MSE(I-H)$$

$$\text{rank}\left(I - \tfrac{1}{n}J\right) = \overbrace{\text{rank}(I)}^{n} - \overbrace{\text{rank}\left(\tfrac{1}{n}J\right)}^{1}$$

iff $I - \frac{1}{n}J$ sym & idempotent

$$\left[I - \tfrac{1}{n}J\right]\left[I - \tfrac{1}{n}J\right] = I - \tfrac{2}{n}J + \tfrac{1}{n^2}JJ$$

$$= I - \tfrac{2}{n}J + \tfrac{1}{n^2}\begin{bmatrix} n & n & n & n \\ \vdots & & & \vdots \\ n & n & n & n \end{bmatrix}$$

$$= I - \tfrac{1}{n}J \quad \blacksquare$$

$$\text{rank}\left(H - \tfrac{1}{n}J\right) = \overbrace{\text{rank}(H)}^{p} - \overbrace{\text{rank}\left(\tfrac{1}{n}J\right)}^{1}$$

iff $H - \frac{1}{n}J$ sym & idempotent

$$\left[H - \tfrac{1}{n}J\right]\left[H - \tfrac{1}{n}J\right] = \cancel{H - \tfrac{2}{n}HJ - \tfrac{1}{n^2}JJ}$$

$$= \cancel{H - \tfrac{2}{n}HJ - \tfrac{1}{n}J}$$

$$= H - \tfrac{1}{n}HJ - \tfrac{1}{n}JH + \tfrac{1}{n^2}JJ$$

$$= H + \tfrac{1}{n}J - \tfrac{1}{n}(HJ + JH)$$

$$= H + \tfrac{1}{n}J - \tfrac{2}{n}HJ$$

$$(J'H')' = (JH)' = H'J' = HJ$$

~~symmetric & idempotent~~

$\cancel{\Rightarrow J = H'H'J}$

ANOVA Sums of squares and Mean squares

$$SSTO = Y'\left[I - \tfrac{1}{n}J\right]Y \qquad df = \text{rank}\left(I - \tfrac{1}{n}J\right) = n-1$$

$$SSE = Y'(I-H)Y \qquad df = \text{rank}(I-H) = n-p \text{ (if design matrix full col. rank)}$$

$$SSR = Y'\left(H - \tfrac{1}{n}J\right)Y \qquad df = \text{rank}\left(H - \tfrac{1}{n}J\right) = p-1$$

Where we define the mean squares as before (by dividing by df)

$$MSR = SSR/p-1$$
$$MSE = SSE/n-p$$

$$E[MSE] = \sigma^2$$
$$E[MSR] = \sigma^2 + f(b)$$

positive function of model params

ANOVA table for general linear regression model:

	SS	df	MS
Regression	SSR	$p-1$	$MSR = SSR/_{p-1}$
Error	SSE	$n-p$	$MSE = SSE/_{n-p}$
Total	SSTO	$n-1$	

F-test for regression relation

To choose between the alternatives

$H_0: \beta_1 = \beta_2 = \cdots \beta_{p-1} = 0$

H_a: not all β_k $(k = 1, \dots, p-1)$ equal zero

we use the test statistic

$$F^* = MSR/_{MSE}$$

The decision rule to control type I error at α is

If $F^* \leq F(1-\alpha; p-1, n-p)$ conclude H_0

elseif $F^* > F(1-\alpha; p-1, n-p)$ " H_a

Standard normal

Cochran's Thm

$$\frac{b_k-\beta_k}{s\{b_k\}} = \frac{b_k-\beta_k}{\sigma\{b_k\}} \div \frac{s\{b_k\}}{\sigma\{b_k\}}$$

$$\frac{\chi^2(n-p)}{n-p}$$

Note! $$\frac{s^2\{b_k\}}{\sigma^2\{b_k\}} = \frac{MSE\ (X'X)^{-1}_{kk}}{\sigma^2\ (X'X)^{-1}_{kk}} = \frac{SSE/n-p}{\sigma^2}$$

so $$\frac{b_k-\beta_k}{s\{b_k\}} \sim \frac{z}{\sqrt{\frac{\chi^2(n-p)}{n-p}}}$$ independent as well, also by Cochran's Thm.

i.e.

$$\frac{b_k-\beta_k}{s\{b_k\}} \sim t(n-p)$$

Inferences about regression params

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

$$\mathrm{cov}[b] = \mathrm{cov}[(X'X)^{-1}X'Y]$$
$$= (X'X)^{-1}X'\,\mathrm{cov}[Y]X(X'X)^{-1}$$
$$= \sigma^2(X'X)^{-1}$$

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

<u>Interval estimation of β_k</u>

We can derive

$$\frac{b_k - \beta_k}{s\{b_k\}} \sim t(n-p) \quad , k = 0, 1, \dots, p-1$$

Hence the confidence limits for β_k with $1-\alpha$ confidence coefficient are:

$$b_k \pm t(1-\alpha/2; n-p)\, s\{b_k\}$$

<u>Tests for β_k</u>

What does this mean?

To test

$$H_0: \beta_k = 0$$
$$H_a: \beta_k \neq 0$$

we may use

$$t^* = \frac{b_k}{s\{b_k\}}$$

and the decision rule

if $|t^*| \leq t(1-\alpha/2; n-p)$ conclude H_0
esle H_a

Bonferroni joint confidence intervals can be used to test ~~whether~~ ~~multiple~~ multiple coeffs simultanously.

Interval Estimation of $E\{Y_h\}$

The mean response to be estimated is

$$E\{Y_h\} = X_h' \beta$$

the estimate of this quantity is

$$\hat{Y}_h = X_h' b$$

This estimator is unbiased

$$E[\hat{Y}_h] = E[X_h' b] = X_h' \beta = E\{Y_h\}$$

The variance of this estimator is

$$\sigma^2\{\hat{Y}_h\} = X_h' \sigma^2\{b\} X_h$$

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

when σ^2 unknown we replace it as usual with MSE

$$\Rightarrow s^2\{\hat{Y}_h\} = MSE\left(X_h' (X'X)^{-1} X_h \right)$$

The confidence limits $(1-\alpha)$ for $E\{Y_h\}$ are

$$\hat{Y}_h \pm t(1-\alpha/2; n-p)\, s\{\hat{Y}_h\}$$

Prediction of New Observation $Y_{h(new)}$

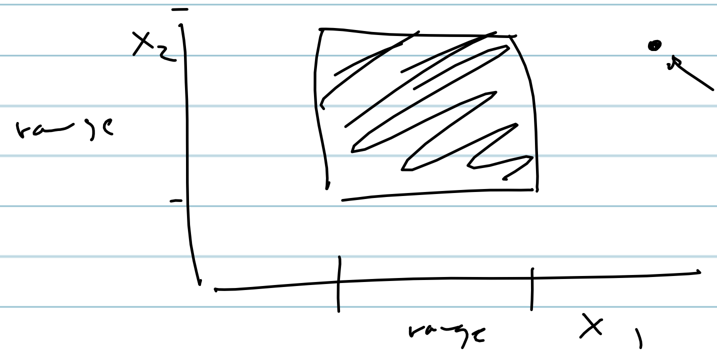
Same as in 1-d case, extra variance

$$\hat{Y}_h \pm t(1-\alpha/2; n-p)\, s\{pred\}$$

where

$$s^2\{pred\} = MSE + s^2\{\hat{Y}_h\}$$
$$= MSE\left(1 + X_h'(X'X)^{-1} X_h\right)$$

Note Hidden Extrapolations



in range for x_2, not for x_1

Meaning of response surface
(take partial derivatives of soln w.r.t. params)

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n$$