The General Linear Model

See ch 7

Scalar form

For $i = 1, \dots, n$

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i$$ where

β_j are unknown constants

x_{ij} are observable, known constants

$\varepsilon_1, \dots, \varepsilon_n$ are unobservable random variables with
$E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$, $Cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$
σ^2 unknown

Y_i are observable random variables

Matrix form

$$Y = X\beta + \varepsilon$$ where

β is a $(k+1) \times 1$ vector of ~~observable~~ unknown constants

X is an $n \times (k+1)$ matrix of observable constants

ε is an $n \times 1$ random vector with
$E(\varepsilon) = 0$, $cov(\varepsilon) = \sigma^2 I_n$
σ^2 unknown

Y is an $n \times 1$ observable random vector

β values are called REGRESSION COEFFICIENTS

Meaning of the regression coefficients

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i$$

$$E(Y_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$$ Really $E(Y_i \mid x_{i1}, \ldots x_{ik})$

Imagine a sub-population of Y values at each combination of x values. THIS IS THE population MEAN.

For $k=1$, a straight line

For $k>1$, a (hyper) plane

β_0 is the intercept

$$\frac{\partial E(Y)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j + \cdots + \beta_k x_k\right)$$

$$= \beta_j$$

HOLDING CONSTANT = CONTROLLING FOR

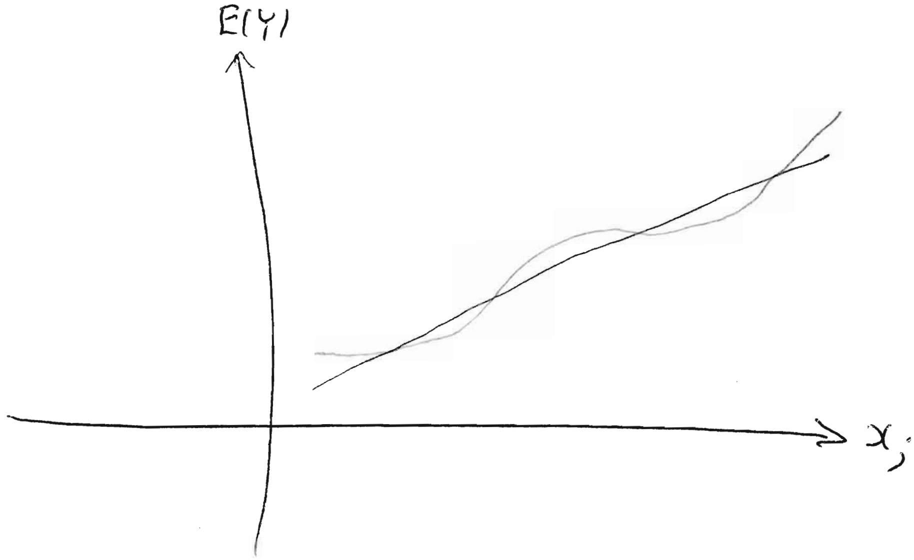
"Partial" regression coefficient

It's more: For fixed values of x_ℓ, $\ell \neq j$

$$E(Y) = \left(\beta_0 + \sum_{\ell \neq j} \beta_\ell x_\ell\right) + \beta_j x_j$$

STRAIGHT LINE

NOT REALLY TRUE, BUT OFTEN CLOSE ENOUGH TO BE USEFUL



"All models are wrong, but some are useful"

Sometimes not close – try to detect

Say "controlling for" the other variables

$\left.\begin{array}{lll} \beta_j > 0 & x_j \text{ positively related to } Y \\ \beta_j < 0 & x_j \text{ neg " " "} \\ \beta_j = 0 & \text{unrelated} \end{array}\right\}$ need tests

MODEST GOAL – we really mean x_j related to average Y.

The $\underset{\sim}{X}$ matrix holds the independent variables.

[redacted]

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & x_{31} & x_{32} & \cdots & x_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

n by $k+1$

↑

Sample size

Study $X'X$ You'll see it a lot

- Symmetric: $(X'X)' = X'X'' = X'X$
- Sums of Squares & cross-products (SSCP)

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & x_{31} & \cdots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2} \\ x_{13} & x_{23} & x_{33} & \cdots & x_{n3} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ \vdots & & \vdots & \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix} = \begin{bmatrix} n & \sum x_{i1} & \sum x_{i2} & \sum x_{i3} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \sum x_{i1}x_{i3} \\ \sum x_{i2} & \sum x_{i2}x_{i1} & \sum x_{i2}^2 & \sum x_{i2}x_{i3} \\ \sum x_{i3} & \sum x_{i3}x_{i1} & \sum x_{i3}x_{i2} & \sum x_{i3}^2 \end{bmatrix}$$

$(k+1) \times (k+1)$

- At least non-negative definite:

$$a'X'Xa = (Xa)'Xa = Z'Z = \sum_{i=1}^{n} z_i^2 \geq 0$$ save

$\underset{n\times(k+1)}{X}$ $\underset{(k+1)\times 1}{a}$

- Suppose the columns (independent variables) are linearly independent. This means the IVs are not redundant. No IV is a linear combination of the others (6 quiz marks ~~sum~~ the quiz ave)

<u>Def</u> $(X'X)a = 0 \Rightarrow a = 0$ save

<u>Theorem</u> $(X'X)^{-1}$ exists if and only if the columns of X are linearly independent

<u>Proof</u>: First linear ind $\Rightarrow$ existence of inverse

Suppose cols of X are linearly independent. Since $a'X'Xa \geq 0$, Showing $a'X'Xa = 0 \Rightarrow a = 0$ will show positive definite.

Let $a'X'Xa = 0 \Rightarrow Xa = 0 \Rightarrow a = 0$ by linear independence

Hence $X'X$ is pos def $\Rightarrow$ eigenvalues all pos

$\Rightarrow (X'X)^{-1}$ exists.

Now $(X'X)^{-1}$ hence l.i.

Assume $(X'X)^{-1}$ exists. Show $Xa = 0 \Rightarrow a = 0$. Let $Xa = 0$

$\Rightarrow X'Xa = X'0 = 0 \Rightarrow a = 0$ ▥

Estimation by LEAST SQUARES

Estimate $\beta_0, \beta_1, \ldots \beta_k$ by picking values that get the observed Y_i values as close as possible to their expected values.

Minimize $\sum_{i=1}^{n} (Y_i - E(Y_i))^2 = Q$

$$Q = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2$$

$$= (Y - X\beta)'(Y - X\beta)$$

Choose a plane that minimizes the sum of squared vertical distances from the points to the plane.

(5.7)

Try differentiating, setting $k+1$ derivatives to zero. Cleaner in matrix notation but...

$$Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_j x_{ij} - \cdots - \beta_k x_{ik})^2$$

$$\frac{\partial Q}{\partial \beta_0} = \sum_{i=1}^{n} 2 (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik}) \cdot 1 \overset{\text{set}}{=} 0$$

$$\frac{\partial Q}{\partial \beta_1} = \sum_{i=1}^{n} 2 (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik}) \cdot x_{i1} \overset{\text{set}}{=} 0$$

$$\frac{\partial Q}{\partial \beta_2} = \sum_{i=1}^{n} 2 (y_i - \beta_0 - \cdots - \beta_k x_{ik}) x_{i2} \overset{\text{set}}{=} 0$$

$$\vdots$$

$$\frac{\partial Q}{\partial \beta_j} = \sum 2 (\qquad\qquad) x_{ij} \overset{\text{set}}{=} 0$$

$$\vdots \qquad\qquad \vdots$$

$$\frac{\partial Q}{\partial \beta_k} = \sum 2 (\qquad\qquad) x_{ik} = 0$$

Divide by 2, multiply through, rearrange to get the NORMAL EQUATIONS

Normal Equations

$$\sum y_i = n\beta_0 + \beta_1 \sum x_{i1} + \cdots + \beta_k \sum x_{ik}$$

$$\sum x_{i1} y_i = \beta_0 \sum x_{i1} + \beta_1 \sum x_{i1}^2 + \beta_2 \sum x_{i1} x_{i2} + \cdots + \beta_k \sum x_{i1} x_{ik}$$

$$\sum x_{i2} y_i = \beta_0 \sum x_{i2} + \beta_1 \sum x_{i1} x_{i2} + \beta_2 \sum x_{i2}^2 + \cdots + \beta_k \sum x_{i2} x_{ik}$$

$$\vdots$$

$$\underbrace{\sum x_{ik} y_i}_{X'y} = \underbrace{\beta_0 \sum x_{ik} + \beta_1 \sum x_{i1} x_{ik} + \beta_2 \sum x_{i2} x_{ik} + \cdots + \beta_k \sum x_{ik}^2}_{(X'X)\beta}$$

So in matrix form, the normal equations are

$$(X'X)\beta = X'y$$

$$\Rightarrow (X'X)^{-1} X'X \beta = (X'X)^{-1} X'y$$

$$\Rightarrow \hat{\beta} = (X'X)^{-1} X'y$$

& put a hat on it

Is it really a minimum?

Least Squares

Minimize $Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2$ over $\underset{\sim}{\beta}$

$$= \sum_{i=1}^{n} (y_i - \underset{\sim}{x}_i'\beta)^2 = (Y - X\beta)'(Y - X\beta)$$

$$= (\underbrace{Y - X\hat{\beta}}_{A} + \underbrace{X\hat{\beta} - X\beta}_{B})'(\underbrace{Y - X\hat{\beta}}_{A} + \underbrace{X\hat{\beta} - X\beta}_{B}) = A'A + A'B + B'A + B'B$$

Look at $A'B = (Y' - \hat{\beta}'X')(X\hat{\beta} - X\beta)$

$$= Y'X\hat{\beta} - Y'X\beta - \hat{\beta}'X'X\hat{\beta} + \hat{\beta}'X'X\beta$$

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

$$= Y'X(X'X)^{-1}X'Y - Y'X\beta - Y'X\underbrace{(X'X)^{-1}X'X}_{I}\hat{\beta} + Y'X\underbrace{(X'X)^{-1}X'X}_{I}\beta$$

$$= Y'X(X'X)^{-1}X'Y - \cancel{Y'X\beta} - Y'X(X'X)^{-1}X'Y + \cancel{Y'X\beta} = 0$$

So $B'A = 0' = 0$, and

$$Q = (Y - X\hat{\beta})'(Y - X\hat{\beta}) + [X(\hat{\beta} - \beta)]'X(\hat{\beta} - \beta)$$

$$= \underbrace{(Y - X\hat{\beta})'(Y - X\beta)}_{SSE} + (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)$$

Since $X'X$ is positive definite (see reverse side), the second term is non-negative, and zero only when $\hat{\beta} - \beta = 0 \iff \beta = \hat{\beta}$. This minimizes the function Q over all β.

$$a'X'Xa = \underset{1\times n}{(Xa)'}\underset{n\times 1}{(Xa)} \qquad \text{SS}, \ \geq 0$$

Suppose $= 0$.

$$\Longleftrightarrow \quad Xa = 0 \Longrightarrow X'Xa = X'0 = 0$$

$$\Longrightarrow (X'X)^{-1}XX'a = (X'X)^{-1}0 = 0$$

$$\Longrightarrow a = 0 \qquad \text{So } X'X \text{ is P.D.}$$

Now Least squares with R: Unit 6