$$\min \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_p X_{ip})^2 = \min f(\beta_0, \beta_1, \ldots, \beta_p)$$

> differentiate wrt β pars; set $= 0$

$$\partial/\partial\beta_0 \; f(\beta_0, \ldots, \beta_p) = \sum_{i=1}^{n} 2(Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_p X_{ip})(-1) \overset{set}{=} 0$$

$$\partial/\partial\beta_1 \; f(\beta_0, \ldots, \beta_p) = \sum_{i=1}^{n} 2(Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_p X_{ip})(-X_{i1}) \overset{set}{=} 0$$

$$\vdots$$

$$\frac{\partial}{\partial\beta_p} f(\beta_0, \ldots, \beta_p) = \sum_{i=1}^{n} 2(Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_p X_{ip})(-X_{ip}) \text{ set } = 0$$

$\Rightarrow$ GIVES ESTIMATING EQUATIONS

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip})$$

$$\sum_{i=1}^{n} X_{i1} Y_i = \sum_{i=1}^{n} X_{i1} (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip})$$

$$\vdots$$

$$\sum_{i=1}^{n} X_{ip} Y_i = \sum_{i=1}^{n} X_{ip} (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip})$$

BIOSTATISTICS 140.653
LEAST SQUARES DERIVATION
FOR MULTIPLE LINEAR REG.

MATRIX ESTIMATING EQUATIONS

- $$\sum_{i=1}^{n} Y_i = \underbrace{(1 \; \cdots \; 1)}_{n} \underset{\sim}{Y}$$

  $$\sum_{i=1}^{n} X_{i1} Y_i = (x_{11} \; \cdots \; x_{n1}) \underset{\sim}{Y}$$

  $$\vdots$$

  $$\sum_{i=1}^{n} X_{ip} Y_i = (x_{1p} \; \cdots \; x_{np}) \underset{\sim}{Y}$$

- $$\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip} = (1 \;\; x_{i1} \; \cdots \; x_{ip}) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

  $$= (1 \;\; x_{i1} \; \cdots \; x_{ip}) \underset{\sim}{\beta}$$

- $\sum_{i=1}^{n} (\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip})$

$$= (1 \; 1 \; \dots \; 1) \begin{bmatrix} (1 \;\; X_{11} \; \cdots \; X_{1p}) \, \underset{\sim}{\beta} \\ (1 \;\; X_{21} \; \cdots \; X_{2p}) \, \underset{\sim}{\beta} \\ \vdots \\ (1 \;\; X_{n1} \; \cdots \; X_{np}) \, \underset{\sim}{\beta} \end{bmatrix}$$

$$= (1 \; 1 \cdots \; 1) X \underset{\sim}{\beta}$$

- SIMILARLY

$$\sum_{i=1}^{n} X_{i1} (\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}) = (X_{11} \cdots X_{n1}) X \underset{\sim}{\beta}$$

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

$$\sum_{i=1}^{n} X_{ip} (\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}) = (X_{1p} \cdots X_{np}) X \underset{\sim}{\beta}$$

- SUMMARIZING, EQUATIONS (= <u>NORMAL</u> EQUATIONS) ARE

$$\boxed{X^T \underset{\sim}{Y} = X^T X \underset{\sim}{\beta}}$$