Least squares estimation of the linear regression coefficients, when the observed values of the dependent and independent variables have error and repeated measurements can be made on the variables

Alice Hsuan

BU-456-M March, 1973

Abstract

Least squares estimation of the linear regression coefficients, when the observed values of the dependent and independent variables have error and repeated measurements can be made on the variables.

The regression model is

$$E\left(y \mid x_1, \ldots x_k\right) = \alpha + \sum_{i=1}^{k} \beta_i x_i$$

and the observations are k+1-tuples

$$\left(\left\{X_{1ij},\ j=1,\ldots,r_{1i}\right\},\ \left\{X_{2ij},\ j=1,\ldots,r_{2i}\right\}\cdots\left\{X_{kij},\ j=1,\ldots,r_{ki}\right\},\right.$$

$$\left.\left\{Y_{ij},\ j=1,\ldots,s_i\right\}\right) \qquad i = 1,\ldots,n$$

where $\quad X_{\ell ij} = x_{\ell i} + \delta_{\ell ij} \qquad \delta_{\ell ij} \sim \mathrm{NIID}\left(0,\ \sigma^2_{\ell i}\right) \qquad \ell = 1,\ldots,k$

$$Y_{ij} = y_i + \eta_{ij} \qquad \eta_{ij} \sim \mathrm{NIID}\left(0,\ \tau^2_i\right) .$$

For the case k=1, $r_i=r$, $s_i=s$, $\sigma_i^2=\sigma^2$, $\tau_i^2=\tau^2$ $\forall$ i = 1,...,n and σ^2, τ^2 are known. The following results are given

$$\hat{\beta} = \frac{a_o b_2 - a_2 b_o + \sqrt{(a_2 b_o - a_o b_2)^2 + a_1^2 b_o b_2}}{a_1 b_o}$$

where $(a_2 b_o - a_o b_2)^2 + a_1^2 b_o b_2$

$$= r^2 s^2 \left\{ \left[r\tau^2 \sum_i (\bar{X}_{i.} - \bar{X}_{..})^2 - s\sigma^2 \sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2 \right]^2 + 4rs\sigma^2\tau^2 \left[\sum_i (\bar{X}_{i.} - \bar{X}_{..})(\bar{Y}_{i.} - \bar{Y}_{..}) \right]^2 \right\}$$

$$a_1 b_o = 2rs\left[s\sigma^2 \sum_i (\bar{X}_{i.} - \bar{X}_{..})(\bar{Y}_{i.} - \bar{Y}_{..}) \right]$$

$$a_o b_2 - a_2 b_o = rs\left[s\sigma^2 \sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2 - r\tau^2 \sum_i (\bar{X}_{i.} - \bar{X}_{..})^2 \right]$$

$$\hat{\alpha} = \bar{Y}_{..} - \hat{\beta}\ \bar{X}_{..}$$

$$\hat{x}_i = \frac{1}{\tau^2 r + \sigma^2 \hat{\beta}^2 s} \left(\tau^2 X_{i.} + \hat{\beta}\sigma^2 Y_{i.} + s\hat{\beta}^2\sigma^2 \bar{X}_{..} - s\hat{\beta}\sigma^2 \bar{Y}_{...} \right)$$

For more general cases, a method for estimating $\hat{\alpha}$, $\hat{\beta}$ and $\hat{x}_i$ is given.

Least squares estimation of the linear regression coefficients, when the observed values of the dependent and independent variables have error and repeated measurements can be made on the variables

Alice Hsuan

BU-456-M March, 1973

Basic Formulation

For the simplest case, we have only one independent variable. The regression model is $y = \alpha + \beta x$ and the observations are

$$\left(\left\{X_{ij} \quad j=1\ldots r_i\right\}, \left\{Y_{ij} \quad j=1\ldots s_i\right\}\right) \qquad i = 1,\ldots,n$$

where

$$X_{ij} = x_i + \delta_{ij} \qquad \delta_{ij} \sim \mathrm{NIID}\left(0, \sigma_i^2\right)$$

$$Y_{ij} = y_i + \eta_{ij} \qquad \eta_{ij} \sim \mathrm{NIID}\left(0, \tau_i^2\right)$$

δ_{ij} and η_{ij} are independent and

σ_i^2, τ_i^2 are known.

Let
$$X' = \left(X_{11},\ldots,X_{1r_1}, X_{21},\ldots,X_{2r_2} \cdots X_{n1},\ldots,X_{nr_n}\right)$$
$$Y' = \left(Y_{11},\ldots,Y_{1s_1}, Y_{21},\ldots,Y_{2s_2} \cdots Y_{n1},\ldots,Y_{ns_n}\right)$$
$$x' = \left(\underbrace{x_1,\ldots,x_1}_{r_1}, \underbrace{x_2,\ldots,x_2}_{r_2} \cdots \underbrace{x_n,\ldots,x_n}_{r_n}\right.$$
$$y' = \left(\underbrace{y_1,\ldots,y_1}_{s_1}, \underbrace{y_2,\ldots,y_2}_{s_2} \cdots \underbrace{y_n,\ldots,y_n}_{s_n}\right)$$

$$Z = \begin{pmatrix} X \\ \mathbf{Y} \end{pmatrix}, \qquad z = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$V = \mathrm{Cov}\left(Z\right) = \begin{pmatrix} \sigma_1^2 & & & & & & & & & & & \\ & \ddots & & & & & & & & & & \\ & & \sigma_1^2 & & & & & & & 0 & & \\ & & & \ddots & & & & & & & & \\ & & & & \sigma_n^2 & & & & & & & \\ & & & & & \ddots & & & & & & \\ & & & & & & \sigma_n^2 & & & & & \\ & & & & & & & \ddots & & & & \\ & & & & & & & & \tau_1^2 & & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & & & \tau_1^2 & \\ & & 0 & & & & & & & & & \ddots \\ & & & & & & & & & & & \tau_n^2 \\ & & & & & & & & & & & \ddots \quad \tau_n^2 \end{pmatrix}$$

The least squares solutions $\hat{\alpha}$, $\hat{\beta}$ are obtained by minimizing

$$\ell = \sum_{i=1}^{n} \left\{ \frac{1}{\sigma_i^2} \sum_{j=1}^{r_i} \left(X_{ij} - x_i\right)^2 + \frac{1}{\tau_i^2} \sum_{k=1}^{s_i} \left(Y_{ik} - y_i\right)^2 \right\}$$

$$= \left(Z - z\right)' \; V^{-1} \; \left(Z - z\right)$$

$$= \|Z - z\|^2_{V^{-1}}$$

Since $z' = (x' \;\; y') = (\underbrace{x_1, \ldots, x_1}_{r_1}, \;\cdots\; \underbrace{x_n, \ldots, x_n}_{r_n}, \; \underbrace{\alpha+\beta x_1, \ldots, \alpha+\beta x_1}_{s_1} \;\cdots$

$\underbrace{\alpha+\beta x_n, \ldots, \alpha+\beta x_n}_{s_n})$

$= (0, \ldots, 0, \ldots, 0, \ldots, 0, \; \alpha, \ldots, \alpha, \ldots, \alpha, \ldots, \alpha)$

$+ \; (x_1, \ldots, x_1, \ldots, x_n, \ldots, x_n, \; \beta x_1, \ldots, \beta x_1, \ldots, \beta x_n, \ldots, \beta x_n)$

$= \alpha \, w' + \mathbf{U}'$ where $w' = (\underset{-}{0} \;\; \underset{-}{1})$ $\mathbf{U}' = (x_1, \ldots, x_1, \ldots, x_n, \ldots, x_n,$

$\beta x_1, \ldots, \beta x_1, \ldots, \beta x_n, \ldots, \beta x_n)$, z must be a linear combination of two vectors, one lies on M_β and the other on N, where $M_\beta = \{\mathbf{U};\ x_i \in R,\ i=1,\dots,n,\ \beta \in R\}$, $N = \{\alpha w;\ \alpha \in R\}$. Clearly, M_β and N are two disjoint subspaces of R^{2n} with dimensions n and 1 respectively.

Our goal is to minimize $\ell = \|Z - z\|^2_{V^{-1}} = \|Z - \mathbf{U} - \alpha w\|^2_{V^{-1}}$ subject to $\mathbf{U} \in M_\beta$, $w = \binom{0}{\underline{1}}$. Since we can rewrite ℓ as

$$\|Z - \mathbf{U} - \alpha w\|^2_{V^{-1}}$$

$$= \|P_{M_\beta \oplus N}(Z - U - \alpha w)\|^2_{V^{-1}} + \|P_{(M_\beta \oplus N)^\perp}(Z - U - \alpha w\|^2_{V^{-1}}$$

$$= \|P_{M_\beta \oplus N}(Z - U - \alpha w\|^2_{V^{-1}} + \|P_{(M_\beta \oplus N)^\perp}(Z)\|^2_{V^{-1}}$$

where $(M_\beta \oplus N)^\perp$ is the orthogonal completement of $M_\beta \oplus N$ with inner product with respect to V^{-1}. $P_*(Z)$ is the orthogonal projection (w.r.t. V^{-1}) of Z to the space (*). The minimum will be obtained, if

(1) $U + \alpha w = P_{M_\beta \oplus N}(Z)$ and

(2) $\|P_{(M_\beta \oplus N)^\perp}(Z)\|^2_{V^{-1}}$ is minimized.

Let $E = (e_1, e_2, \dots, e_n, e_{n+1})$ be an orthonormal basis for $M_\beta \oplus N^{(Z)}$ w.r.t. V^{-1}. i.e., $e_i' V^{-1} e_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ $\quad i, j = 1, \dots, n+1$

then $P_{M_\beta \oplus N}(Z) = EE'V^{-1}Z$

We need only to find $(e_1, e_2, \dots, e_n, e_{n+1})$ in order to get $P_{M_\beta \oplus N}(Z)$.

It is easy to see that for $i = 1,\dots,n$

$$e_i' = \Big(\underbrace{0,\dots,0}_{\sum_{j=1}^{i-1} r_j}, \underbrace{a_i,\dots,a_i}_{r_i}, \underbrace{0,\dots,0}_{\sum_{j=i+1}^{n} r_j}, \underbrace{0,\dots,0}_{\sum_{j=1}^{i-1} s_j}; \underbrace{\beta a_i,\dots,\beta a_i}_{s_i}, \underbrace{0,\dots,0}_{\sum_{j=i+1}^{n} s_j}\Big)$$

where

$$a_i^2 = \frac{\sigma_i^2\tau_i^2}{\tau_i^2 r_i+\sigma_i^2\beta^2 s_i} \quad .$$

Since $e_1,\dots,e_n$, w is a basis of $M_\beta \oplus N$, we can get e_{n+1} by the Gram-Schmitt orthogonalization method

$$e_{n+1} = \frac{w-(e_1,w)_{V^{-1}}e_1 - \cdots - (e_n,w)_{V^{-1}}e_n}{\|w-(e_1,w)_{V^{-1}}e_1 - \cdots - (e_n,w)_{V^{-1}}e_n\|_{V^{-1}}}$$

$$(e_i,w)_{V^{-1}} = e_i'V^{-1}w = \frac{s_i\beta a_i}{\tau_i^2}$$

$$\left(\sum_{i=1}^{n}(e_i,w)e_i\right)' = \Bigg(\underbrace{\frac{s_1\beta a_1}{\tau_1^2}a_1,\dots\frac{s_1\beta a_1}{\tau_1^2}a_1}_{r_1},\dots \underbrace{\frac{s_n\beta a_n}{\tau_n^2}a_n,\dots\frac{s_n\beta a_n}{\tau_n^2}a_n}_{r_n},\dots \underbrace{\frac{s_1\beta a_1}{\tau_1^2}\beta a_1,\dots\frac{s_1\beta a_1}{\tau_1^2}\beta a_1}_{s_1},\dots$$

$$\underbrace{\frac{s_n\beta a_n}{\tau_n^2}\beta a_n,\dots\frac{s_n\beta a_n}{\tau_n^2}\beta a_n}_{s_n}\Bigg) .$$

Let $d_i = 1 - \dfrac{s_i\beta a_i}{\tau_i^2}\beta a_i = \dfrac{\tau_i^2 r_i}{\tau_i^2 r_i+\sigma_i^2\beta^2 s_i}$

$$c_i = -\frac{s_i\beta a_i}{\tau_i^2}a_i = -\frac{s_i\beta\sigma_i^2}{\tau_i^2 r_i+\sigma_i^2\beta^2 s_i},$$

then

$$e'_{n+1} = \frac{c_1,\ldots,c_1,\ldots,c_n,\ldots,c_n,\ d_1,\ldots,d_1,\ldots,d_n,\ldots,d_n)}{\|(c_1,\ldots,c_1,\ldots,c_n,\ldots,c_n\ ,\ d_1,\ldots,d_1,\ldots,d_n,\ldots,d_n)\|_{V^{-1}}}$$

$$\|(c_1,\ldots,c_1,\ldots,c_n,\ldots,c_n,\ d_1,\ldots,d_1,\ldots,d_n,\ldots,d_n)\|^2_{V^{-1}}$$

$$= \sum_i \frac{r_i s_i}{\tau_i^2 r_i+\sigma_i^2\beta^2 s_i} \overset{\text{call}}{=} K^2$$

$$e'_{n+1} = \frac{1}{\sqrt{\sum_{i=1}^{n}\frac{r_i s_i}{\tau_i^2 r_i+\sigma_i^2\beta^2 s_i}}}\Bigg(\underbrace{-\frac{s_1\beta\sigma_1^2}{\tau_1^2 r_1+\sigma_1^2\beta^2 s_1},\ldots,-\frac{s_1\beta\sigma_1^2}{\tau_1^2 r_1+\sigma_1^2\beta^2 s_1}}_{r_1},\ldots,$$

$$\underbrace{-\frac{s_n\beta\sigma_n^2}{\tau_n^2 r_n+\sigma_n^2\beta^2 s_n},\ldots,-\frac{s_n\beta\sigma_n^2}{\tau_n^2 r_n+\sigma_n^2\beta^2 s_n}}_{r_n},\ldots,$$

$$\underbrace{\frac{\tau_1^2 r_1}{\tau_1^2 r_1+\sigma_1^2\beta^2 s_1},\ldots,\frac{\tau_1^2 r_1}{\tau_1^2 r_1+\sigma_1^2\beta^2 s_1}}_{s_1},\ldots,$$

$$\underbrace{\frac{\tau_n^2 r_n}{\tau_n^2 r_n+\sigma_n^2\beta^2 s_n},\ldots,\frac{\tau_n^2 r_n}{\tau_n^2 r_n+\sigma_n^2\beta^2 s_n}}_{s_n}\Bigg)$$

$$E' = \begin{bmatrix} a_1 \cdots a_1 & 0 \cdots\cdots 0 & \beta a_1 \cdots \beta a_1 & 0 \cdots\cdots 0 \\ \vdots & & & \\ 0 \cdots\cdots 0 & a_n \cdots a_n & 0 \cdots\cdots 0 & \beta a_n \cdots \beta a_n \\ \frac{C_1}{K} \cdots \frac{C_1}{K} & \frac{C_n}{K} \cdots \frac{C_n}{K} & \frac{d_1}{K} \cdots \frac{d_1}{K} & \frac{d_n}{K} \cdots \frac{d_n}{K} \end{bmatrix}$$

$$EE'V^{-1} = \begin{bmatrix} \boxed{\frac{a_1^2}{\sigma_1^2}} & & \boxed{\frac{\beta a_1^2}{\tau_1^2}} & \\ & \ddots & & \ddots \\ & \boxed{\frac{a_n^2}{\sigma_n^2}} & & \boxed{\frac{\beta a_n^2}{\tau_n^2}} \\ \boxed{\frac{\beta a_1^2}{\sigma_1^2}} & & \boxed{\frac{\beta a_1^2}{\tau_1^2}} & \\ & \ddots & & \ddots \\ & \boxed{\frac{\beta a_n^2}{\sigma_n^2}} & & \boxed{\frac{\beta a_n^2}{\tau_n^2}} \end{bmatrix}$$

$$+ \frac{1}{K^2} \begin{bmatrix} \boxed{\frac{C_1^2}{\sigma_1^2}} & \boxed{\frac{C_1 C_2}{\sigma_2^2}} & \cdots & \boxed{\frac{C_1 C_n}{\sigma_n^2}} & \boxed{\frac{C_1 d_1}{\tau_1^2}} & \cdots & \boxed{\frac{C_1 d_n}{\tau_n^2}} \\ \boxed{\frac{C_1 C_2}{\sigma_1^2}} & \ddots & & & & & \\ \vdots & & \ddots & & & & \\ \boxed{\frac{C_1 C_n}{\sigma_1^2}} & & & \boxed{\frac{C_n^2}{\sigma_n^2}} & & & \\ \boxed{\frac{C_1 d_1}{\sigma_1^2}} & & & & \boxed{\frac{d_1^2}{\tau_1^2}} & & \\ \vdots & & & & & \ddots & \\ \boxed{\frac{C_1 d_n}{\sigma_1^2}} & & & & & & \boxed{\frac{d_n^2}{\tau_n^2}} \end{bmatrix}$$

where $\frac{a_i^2}{\sigma_i^2} = \frac{\tau_i^2}{\tau_i^2 r_i + \sigma_i^2 \beta^2 s_i}$ $\qquad$ $\frac{a_i^2}{\tau_i^2} = \frac{\sigma_i^2}{\tau_i^2 r_i + \sigma_i^2 \beta^2 s_i}$

$$\frac{c_i^2}{\sigma_i^2} = \frac{s_i^2 \beta^2 \sigma_i^2}{(\tau_i^2 r_i + \sigma_i^2 \beta^2 s_i)^2} \qquad \frac{d_i^2}{\tau_i^2} = \frac{\tau_i^2 r_i^2}{(\tau_i^2 r_i + \sigma_i^2 \beta^2 s_i)^2} \quad .$$

Also, $P_{M_\beta \oplus N}(Z) = U + \alpha w = EE'V^{-1}Z$.

Hence, we can get the M.L.E. for $\hat{\alpha}$.

In order to get the M.L.E. for $\hat{\beta}$, we have to find $\min \| P_{(M_\beta \oplus N)^{\perp}}(Z) \|^2_{V^{-1}}$.

Since, $\| P_{(M_\beta \oplus N)^{\perp}}(Z) \|^2_{V^{-1}} = \|Z\|^2_{V^{-1}} - \| P_{M_\beta \oplus N}(Z) \|^2_{V^{-1}}$ so this is equivalent to

maximizing $\| P_{M_\beta \oplus N}(Z) \|^2_{V^{-1}}$ by using $P_{M_\beta \oplus N}(Z) = EE'V^{-1}Z$ we have

$$\begin{aligned} \| P_{M_\beta \oplus N}(Z) \|^2_{V^{-1}} &= \left(EE'V^{-1}Z \right)' V^{-1} \left(EE'V^{-1}Z \right) \\ &= Z'V^{-1}EE'V^{-1}EE'V^{-1}Z \\ &= Z'V^{-1}EE'V^{-1}Z \end{aligned}$$

Hence, we need only to maximize $Z'V^{-1}EE'V^{-1}Z$ w.r.t. β .

A special case

It is too messy to calculate all these by hand. But it is quite easy to do so by using a computer program to get it. In order to have some results, I worked out the simplest case.

When $r_i = r$, $s_i = s$, $\sigma_i^2 = \sigma^2$, $\tau_i^2 = \tau^2$ $\quad \forall i = 1,\ldots,n$

$$a^2 = a_i^2 = \frac{\sigma^2\tau^2}{\tau^2 r+\sigma^2\beta^2 s}$$

$$c^2 = c_i^2 = -\frac{s\ \beta\ \sigma^2}{\tau^2 r+\sigma^2\beta^2 s}$$

$$d = d_i = \frac{\tau^2 r}{\tau^2 r+\sigma^2\beta^2 s}$$

$$K^2 = \frac{nrs}{\tau^2 r+\sigma^2\beta^2 s}$$

$$EE'V^{-1} = \frac{1}{\tau^2 r+\sigma^2\beta^2 s}\left\{\left(\begin{array}{ccc|ccc} \boxed{\tau^2} & & & \boxed{\beta\sigma^2} & & \\ & \ddots & & & \ddots & \\ & & \boxed{\tau^2} & & & \boxed{\beta\sigma^2} \\ \hline \boxed{\beta\tau^2} & & & \boxed{\beta^2\sigma^2} & & \\ & \ddots & & & \ddots & \\ & & \boxed{\beta\tau^2} & & & \boxed{\beta^2\sigma^2} \end{array}\right) + \frac{1}{n}\left(\begin{array}{c|c} \dfrac{s\beta^2\sigma^2}{r} & -\beta\sigma^2 \\ \hline -\beta\tau^2 & \dfrac{\tau^2 r}{s} \end{array}\right)\right\}$$

$$P_{M_\beta \oplus N}(Z) = EE'V^{-1}Z$$

$$= \frac{1}{\tau^2 r+\sigma^2\beta^2 s}\left\{\begin{pmatrix} \tau^2 X_{1.} + \beta\sigma^2 Y_{1.} \\ \vdots \\ \tau^2 X_{n.} + \beta\sigma^2 Y_{n.} \\ \vdots \\ \hline \beta\tau^2 X_{1.} + \beta^2\sigma^2 Y_{1.} \\ \vdots \\ \beta\tau^2 X_{n.} + \beta^2\sigma^2 Y_{n.} \\ \vdots \end{pmatrix} + \frac{1}{n}\begin{pmatrix} \frac{s\beta^2\sigma^2}{r} X_{..} - \beta\sigma^2 Y_{..} \\ \hline -\beta\tau^2 X_{..} + \frac{\tau^2 r}{s} Y_{..} \end{pmatrix}\right\}$$

$$= \begin{pmatrix} 0 \\ \hat{\alpha} \end{pmatrix} + \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_n \\ \vdots \\ \hat{\beta}\hat{x}_1 \\ \vdots \\ \hat{\beta}\hat{x}_n \\ \vdots \end{pmatrix}$$

Hence

$$\hat{\alpha} = \frac{1}{n(\tau^2 r+\sigma^2\hat{\beta}^2 s)}\left(-\hat{\beta}\tau^2 X_{..} + \frac{\tau^2 r}{s} Y_{..} - \frac{s\hat{\beta}^3\sigma^2}{r} X_{..} + \hat{\beta}^2\sigma^2 Y_{..}\right)$$

$$= \frac{\tau^2 r+\sigma^2\hat{\beta}^2 s}{n(\tau^2 r+\sigma^2\hat{\beta}^2 s)}\left(\frac{1}{s} Y_{..} - \frac{\hat{\beta}^2}{r} X_{..}\right)$$

$$= \bar{Y}_{..} - \hat{\beta}\bar{X}_{..} \qquad \text{as expected}$$

$$\hat{x}_i = \frac{1}{\tau^2 r+\sigma\alpha\hat{\beta}^2 s}\left(\tau^2 X_{i.} + \hat{\beta}\sigma^2 Y_{i.} + s\hat{\beta}^2\sigma^2\bar{X}_{..} - s\hat{\beta}\sigma^2\bar{Y}_{..}\right)$$

To find $\hat{\beta}$, we have to maximize $Z'V^{-1}EE'V^{-1}Z$

let $f_{(\beta)} = Z'V^{-1}EE'V^{-1}Z$

$$= \frac{1}{\tau^2 r+\sigma^2\beta^2 s}\left\{\sum_i \frac{1}{\sigma^2}\left(\tau^2 X_{i.} + \beta\sigma^2 Y_{i.}\right) X_{i.} + \sum_i \frac{1}{\tau^2}\left(\beta\tau^2 X_{i.} + \beta^2\sigma^2 Y_{..}\right)\right.$$

$$\left. + \frac{1}{n}\left[\left(\frac{s\beta^2\sigma^2}{r} X_{..} - \beta\sigma^2 Y_{..}\right)\frac{X_{..}}{\sigma^2} - \left(\beta\tau^2 X_{..} - \frac{\tau^2 r}{s} Y_{..}\right)\frac{Y_{..}}{\tau^2}\right]\right\}$$

$$= \frac{1}{\tau^2 r+\sigma^2\beta^2 s}\left\{\sum_i\left(\frac{\tau^2}{\sigma^2} X_{i.}^2 \quad 2\beta X_{i.} Y_{i.} + \beta X_{i.} Y_{i.} + \frac{\beta^2\sigma^2}{\tau^2} Y_{i.}^2\right) + \frac{s\beta^2}{rn} X_{..}^2 - \frac{2\beta}{n} X_{..} Y_{..} + \frac{r}{sn} Y_{..}^2\right\}.$$

Write $f(\beta)$ as the following form:

$$f(\beta) = \frac{a_0\beta^2 + a_1\beta + a_2}{b_0\beta^2 + b_2}$$

where $b_0 = \sigma^2 s \qquad b_2 = \tau^2 r$

$$a_0 = \sum_i \frac{s^2\sigma^2}{\tau^2}\bar{Y}_{i.}^2 + nrs\bar{X}_{..}^2$$

$$a_1 = 2rs\sum_i \bar{X}_{i.}\bar{Y}_{i.} - 2nrs\bar{X}_{..}\bar{Y}_{..} = 2rs\sum_i(\bar{X}_{i.} - \bar{X}_{..})(\bar{Y}_{i.} - \bar{Y}_{..})$$

$$a_2 = \sum_i \frac{r^2\tau^2}{\sigma^2}\bar{X}_{i.}^2 + nrs\bar{Y}_{..}^2$$

$$f'(\beta) = \frac{(b_0\beta^2+b_2)(2a_0\beta+a_1) - (a_0\beta^2+a_1\beta+a_2)(2b_0\beta)}{(b_0\beta^2+b_2)^2}$$

set
$$f'(\hat{\beta}) = 0$$

then $\left(b_0\hat{\beta}^2+b_2\right)\left(2a_0\hat{\beta}+a_1\right) - \left(a_0\hat{\beta}^2+a_1\hat{\beta}+a_2\right)\left(2b_0\hat{\beta}\right) = 0$

$$a_1b_0\hat{\beta}^2 + 2\left(a_2b_0 - a_0b_2\right)\hat{\beta} - a_1b_2 = 0$$

$$\hat{\beta} = \frac{a_0b_2 - a_2b_0 \pm \sqrt{(a_2b_0-a_0b_2)^2+a_1^2b_0b_2}}{a_1b_0} .$$

Since we have two roots for $f'(\beta) = 0$, we have to find out which one maximizes $f(\beta)$. It is too messy to compute the 2^{nd} order derative to find this out. It is sufficient to find some behavior of this function which indicates the root which maximizes $f(\beta)$.

$$f(\beta) = \frac{a_0\beta^2+a_1\beta+a_2}{b_0\beta^2+ b_2} .$$

In our case, $a_0 > 0$, $a_2 > 0$, $b_0 > 0$, $b_2 > 0$. a_1 can be any real number if $a_1 = 0$, $\hat{\beta} = 0$ not what we are interested in. Assume $a_1 \neq 0$

$$\lim_{\beta\to+\infty} f(\beta) = \frac{a_0}{b_0} \qquad \lim_{\beta\to-\infty} f(\beta) = \frac{a_0}{b_0} .$$

Also, $f(\beta) = \frac{a_0}{b_0}$ has a root when $a_1 \neq 0$

$$\frac{a_0\beta^2+a_1\beta+a_2}{b_0\beta^2+b_2} = \frac{a_0}{b_0}$$

the root is $\beta = \frac{a_0b_2-a_2b_0}{a_1b_0}$

if $a_1 > 0$

when β is a large positive number,

$$f(\beta) \approx \frac{a_0\beta^2+a_1\beta}{b_0\beta^2} = \frac{a_0}{b_0} + \frac{a_1}{b_0\beta} > \frac{a_0}{b_0}$$

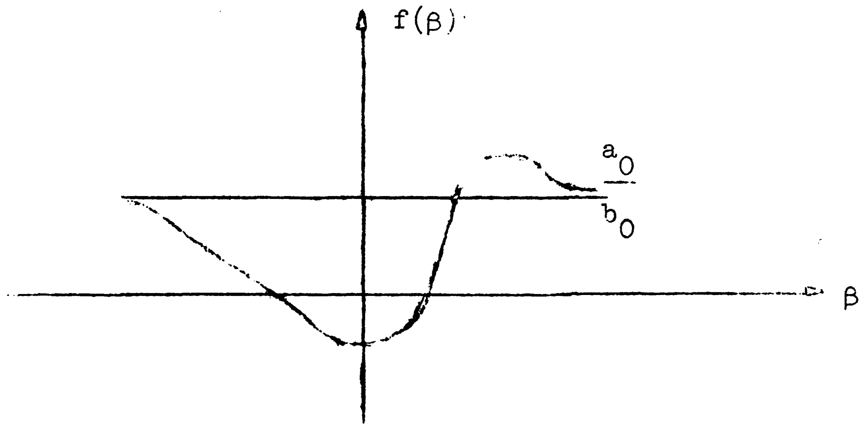
$f(\beta)$ decreasing to $\frac{a_0}{b_0}$

when β is a large negative number,

$$f(\beta) \approx \frac{a_0\beta^2+a_1\beta}{b_0\beta^2} = \frac{a_0}{b_0} + \frac{a_1}{b_0\beta} < \frac{a_0}{b_0}$$

$f(\beta)$ increasing to $\frac{a_0}{b_0}$.

Also, since $f(\beta)$ is continuous, $f(\beta)$ looks like



Hence, the larger root of $f'(\beta) = 0$ achieves the maximum. For the case $a_1 < 0$ we can use a similar argument and then the smaller root of $f'(\beta)$ achieves the maximum.

In both cases, we have

$$\hat{\beta} = \frac{a_0b_2-a_2b_0 + \sqrt{(a_2b_0-a_0b_2)^2+a_1^2b_0b_2}}{a_1b_0}$$

where

$$\left(a_2b_0-a_0b_2\right)^2 + a_1^2b_0b_2 = \left(r^2s\tau^2\sum_i \bar{X}_{i.}^2 + nrs^2\sigma^2\bar{Y}_{..}^2 - rs^2\sigma^2\Sigma\bar{Y}_{i.}^2 - nr^2s\tau^2\bar{X}_{..}^2\right)^2$$

$$+ 4r^3s^3\sigma^2\tau^2\left[\sum_i(\bar{X}_{i.}-\bar{X}_{..})(\bar{Y}_{i.}-\bar{Y}_{..})\right]^2$$

$$= r^2s^2\left\{\left[r\tau^2\sum_i(\bar{X}_{i.}-\bar{X}_{..})^2 - s\sigma^2\sum_i(\bar{Y}_{i.}-\bar{Y}_{..})^2\right]^2 + 4rs\sigma^2\tau^2\left[\sum_i(\bar{X}_{i.}-\bar{X}_{..})(\bar{Y}_{i.}-\bar{Y}_{..})\right]^2\right\}$$

$$a_1b_0 = 2rs\left[s\sigma^2\sum_i(\bar{X}_{i.}-\bar{X}_{..})(\bar{Y}_{i.}-\bar{Y}_{..})\right]$$

$$a_0b_2 - a_2b_0 = rs\left[s\sigma^2\sum_i(\bar{Y}_{i.}-\bar{Y}_{..})^2 - r\tau^2\sum_i(\bar{X}_{i.}-\bar{X}_{..})^2\right] .$$

For more general cases, when the regression model is

$$y_i = \alpha + \beta x_i + \epsilon_i \qquad \epsilon_i \sim N(0, \lambda_i^2)$$

ϵ_i, δ_{ij}, η_{ij} are independent and σ_i^2, τ_i^2, λ_i^2 are known

$$Y_{ij} = y_i + \eta_{ij} = \alpha + \beta x_i + \epsilon_i + \eta_{ij}$$

$$\mathrm{Var}\left(Y_{ij}\right) = \lambda_i^2 + \tau_i^2$$

$$\mathrm{Cov}\left(Y_{ij}, Y_{ik}\right) = \mathrm{Cov}\left(\epsilon_i + \eta_{ij} \quad \epsilon_i + \eta_{ik}\right) = \mathrm{Var}\left(\epsilon_i\right) = \lambda_i^2 .$$

The covariance matrix V is positive definite, and

$$V = \left[\begin{array}{c|c} \begin{matrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} \boxed{\begin{matrix} \lambda_1^2+\tau_1^2 & \lambda_1^2 \\ \lambda_1^2 & \lambda_1^2+\tau_1^2 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_n^2+\tau_n^2 & \lambda_n^2 \\ \lambda_n^2 & \lambda_n^2+\tau_n^2 \end{matrix}} \end{matrix} \end{array}\right]$$

Since V is invertable, we can still use the method given above to get the M.L.E. for $\hat{\alpha}$, $\hat{\beta}$ and $\hat{x}_i$.

When we have more than one independent variable, the method can still be used, only the dimension of the vector $Z' = (X_1, \ldots X_K, Y)$ increases rapidly.

A similar result was reported by C. Villegas (Ann. Math. Statist. 32, 1048-1062).

Illustration of some areas of application (by D. S. Robson)

A standard method of assaying dog serum for distemper antibody entails the use of an egg-adapted strain of distemper virus which produces visible signs of infection in the embryonated chicken egg. A fixed dose of virus mixed with serial dilutions of immune dog serum can thus be inoculated into chicken eggs to determine the serum dilution level at which the given dose of virus is neutralized in 50 percent of the eggs. On the logarithmic scale this end-point is referred to as the titer of the serum (with respect to the fixed dose of egg-adapted virus).

Laboratory incubation of large numbers of inoculated chicken eggs is both awkward and expensive, and a new assay technique has been developed around a strain of distemper virus adapted to dog kidney tissue cultured cells. A sample of tissue culture placed in a small hemispherical well (less than a cm. diameter) in a plastic plate now replaces the embryonated chicken egg; when a well of tissue culture is inoculated with a dose of serum-virus mixture and incubated for a few hours at controled temperature, infection of the cell culture becomes visibly apparent unless serum antibody was present in sufficient amount to neutralize the fixed dose of virus. The serum dilution level at which 50 percent of the wells show infection is then taken as the end-point which quantifies the antibody level of the original serum, and on the logarithmic scale is again referred to as the serum titer.

Since both the host and pathogen differ in these two systems then the same immune dog serum tested in the two systems will produce two different titers. Moreover, there are several sources of error variation within each system so that when the same serum sample is repeatedly and independently tested within the same system there is variation among the resulting titers. Thus, if x is the true (expected value) titer of the serum in the egg system and y is the true titer in the tissue culture system then independent replicates of the i'th serum produce observations $X_{ij} = x_i + \delta_{ij}$ and $Y_{ij} = y_i + \eta_{ij}$. Calibration of the new method in terms of the old requires estimation of the regression of y on x.

Regression problems of this type also arise in other contexts where the principle of repeated measurements is replaced by the more general principle sampling error estimation. Suppose, for example, that x and y are unknown parameters of a stochastic process and are identifiable with respect to a given experimental or sampling design, so that data are available to produce estimators X and Y and also estimators of the (conditional) covariance matrix

$$V = \begin{bmatrix} E\left[(X-x)^2 | x,y\right] & E\left[(X-x)(Y-y) | x,y\right] \\ E\left[(X-x)(Y-y | x,y)\right] & E\left[(Y-y)^2 | x,y\right] \end{bmatrix}$$

If this process is observed under n different conditions with unknown parameter values (x_i, y_i), $i = 1,\ldots,n$, there may be interest in examining the relationship between x and y. For example, suppose the process under consideration is a pure death process with an annual survival rate $S_i = \exp(-x_i - y_i)$ reflecting two distinct causes of death in year i. In this case the primary purpose of the investigation may be to examine the relationship between x and y when only correlated estimators (X_i, Y_i) of (x_i, y_i) are available, as when (x_i, y_i) are estimated from tag-recapture experiments.