11 Feb 04

Gauss-Markov set up with correlated errors

(The General Gauss-Markov Model) GMM

$\underset{\sim}{Y} = \underset{\sim}{X}\,\underset{\sim}{\alpha} + \underset{\sim}{\epsilon}$

$\underset{\sim}{X}$ fixed, $E\underset{\sim}{\epsilon} = 0$, $\operatorname{var}\underset{\sim}{\epsilon} = \sigma^2 \underset{\sim}{G}$

$E\,\underset{\sim}{Y} \mid \underset{\sim}{X} = \underset{\sim}{X}\,\underset{\sim}{\alpha}$, $\operatorname{var}\underset{\sim}{Y} \mid \underset{\sim}{X} = \sigma^2 \underset{\sim}{G}$

$\underset{\sim}{Y}$, $\underset{\sim}{X}$, $\underset{\sim}{G}$ known

Nonsingular case of $\underset{\sim}{G}^{-1}$ existing

$\underset{\sim}{G}^{-\frac{1}{2}}$: $\underset{\sim}{G}^{-\frac{1}{2}}\underset{\sim}{G}^{-\frac{1}{2}} = \underset{\sim}{G}^{-1}$

$\underset{\sim}{G}$ is symmetric, positive definite ($\lambda_i > 0$)

SVD $\underset{\sim}{G} = \underset{\sim}{U}\,\underset{\sim}{\Lambda}\,\underset{\sim}{U}^T$

$\underset{\sim}{G}^{-\frac{1}{2}} = \underset{\sim}{U}\,\underset{\sim}{\Lambda}^{-\frac{1}{2}}\,\underset{\sim}{U}^T$

11 Feb 04

Consider the transformation

$$\underset{\sim}{\mathcal{Y}} = \underset{\sim}{G}^{-\frac{1}{2}} \underset{\sim}{Y}, \quad \underset{\sim}{\mathcal{X}} = \underset{\sim}{G}^{-\frac{1}{2}} \underset{\sim}{X}, \quad \underset{\sim}{\mathcal{E}} = \underset{\sim}{G}^{-\frac{1}{2}} \underset{\sim}{\epsilon}$$

$$\underset{\sim}{\mathcal{Y}} = \underset{\sim}{\mathcal{X}} \underset{\sim}{\alpha} + \underset{\sim}{\mathcal{E}}$$

Normal equations

$$\underset{\sim}{X}^T \underset{\sim}{G}^{-1} \underset{\sim}{X} \hat{\underset{\sim}{\alpha}} = \underset{\sim}{X}^T \underset{\sim}{G}^{-1} \underset{\sim}{Y}$$

Sum of squares minimized

$$(\underset{\sim}{Y} - \underset{\sim}{X} \underset{\sim}{\alpha})^T \underset{\sim}{G}^{-1} (\underset{\sim}{Y} - \underset{\sim}{X} \underset{\sim}{\alpha})$$

$\underset{\sim}{P}^T \underset{\sim}{\alpha}$ is estimable iff $\exists \ \underset{\sim}{L} \ni$

$$\underset{\sim}{X}^T \underset{\sim}{L} = \underset{\sim}{P}$$

$$E \ \underset{\sim}{P}^T \hat{\underset{\sim}{\alpha}} = \underset{\sim}{P}^T \underset{\sim}{\alpha}$$

$$\text{var } \underset{\sim}{P}^T \hat{\underset{\sim}{\alpha}} = \underset{\sim}{P}^T \underset{\sim}{C} \underset{\sim}{P} \sigma^2$$

$$\underset{\sim}{C} = (\underset{\sim}{X}^T \underset{\sim}{G}^{-1} \underset{\sim}{X})^{-}$$

11 Feb 04

Residuals

$$\hat{\mathbf{Y}} - \hat{\mathbf{X}}\hat{\boldsymbol{\alpha}} = \mathbf{G}^{-1/2}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\alpha}})$$

Residual SS (RSS)

$$(\hat{\mathbf{Y}} - \hat{\mathbf{X}}\hat{\boldsymbol{\alpha}})^T(\hat{\mathbf{Y}} - \hat{\mathbf{X}}\hat{\boldsymbol{\alpha}})$$

$$= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\alpha}})^T \mathbf{G}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\alpha}})$$

$$\hat{\sigma}^2 = RSS/r \qquad r = r(\mathbf{X})$$

Example

Weighted least squares (WLS)

$$\mathbf{G} = \mathbf{W}^{-1} \qquad \mathbf{W} = \text{diag}\{w_i\}$$

SS

$$\sum_{i=1}^{n} w_i (y_i - \mathbf{x}_i \boldsymbol{\alpha})^2$$

Material basic for generalized linear model.
Splus, R routines generally allow weights

The linear mixed effect model

$$\underset{n\times 1}{\tilde{Y}} = \underset{n\times p}{\tilde{X}}\tilde{\alpha} + \underset{n\times q}{\tilde{Z}}\tilde{b} + \tilde{e}$$

Data $\tilde{Y}, \tilde{X}, \tilde{Z}$

$\tilde{b}, \tilde{e}$ r.v.'s mean 0

$\text{var}\,\tilde{b} = \tilde{D}$, $\text{var}\,\tilde{e} = \tilde{R}$, $\text{cov}\{b, e\} = 0$

$\tilde{X}, \tilde{Z}$ fixed

$\tilde{D}$ and $\tilde{R}$ may be parametrized by θ

Example

Earthquake attenuation laws

$$\log A_{ij} = \alpha_i + \beta_i M_i - \log\left(\sqrt{d_{ij}^2 + \delta_i^2}\right) - \gamma_i\sqrt{d_{ij}^2 + \delta_i^2} + \epsilon_{ij}$$

i: earthquake
j: measurement within earthquake

$\alpha_i, \beta_i, \gamma_i, \delta_i, \epsilon_{ij}$ random

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \underset{\sim}{Z}\underset{\sim}{b} + \underset{\sim}{e}$$

$$E\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\alpha} \qquad E\underset{\sim}{Y}|\underset{\sim}{b} = \underset{\sim}{X}\underset{\sim}{\alpha} + \underset{\sim}{Z}\underset{\sim}{b}$$

$$\text{var}\,\underset{\sim}{Y} = \underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}^T + \underset{\sim}{R} = \underset{\sim}{V}$$

Harville (1977) JASA, (1976) Ann Statist.

(GMM) estimate of $\underset{\sim}{\alpha}$ satisfies

$$(\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X})\hat{\underset{\sim}{\alpha}} = \underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{Y}$$

$$\text{var}\,\hat{\underset{\sim}{\alpha}} = (\underset{\sim}{X}^T\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}$$

What about $\underset{\sim}{b}$? Let $\tilde{\underset{\sim}{b}}$ denote a predictor

Consider

$$\begin{bmatrix} \underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{Z} \\ \underset{\sim}{Z}^T\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{D}^{-1} + \underset{\sim}{Z}^T\underset{\sim}{R}^{-1}\underset{\sim}{Z} \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{\alpha}} \\ \tilde{\underset{\sim}{b}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}^T\underset{\sim}{R}^{-1}\underset{\sim}{Y} \\ \underset{\sim}{Z}^T\underset{\sim}{R}^{-1}\underset{\sim}{Y} \end{bmatrix} \quad (*)$$

assuming inverses exist and that $\underset{\sim}{D}$ and $\underset{\sim}{R}$ are known

(Having R is important.)

11 Feb. 04

The equations (*) result from minimizing, wrt α, b

$$(y - X\alpha - Zb)^T R^{-1} (y - X\alpha - Zb) + b^T D^{-1} b$$

(VI) 11 Feb 04

Suppose one wishes to estimate

$$\underset{\sim}{\lambda}_1^T \underset{\sim}{\alpha} + \underset{\sim}{\lambda}_2^T \underset{\sim}{\beta}$$

with $\underset{\sim}{\lambda}_1^T \underset{\sim}{\alpha}$ estimable, by a linear combination of $\underset{\sim}{Y}$, that is unbiased and has minimum mean squared error.

$$E\hat{\underset{\sim}{\theta}} = \underset{\sim}{\lambda}_1^T \underset{\sim}{\alpha}, \quad E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\lambda}_1 \underset{\sim}{\alpha} - \underset{\sim}{\lambda}_2^T \underset{\sim}{\beta})^2 \text{ is minimized}$$

The estimate is

$$\underset{\sim}{\lambda}_1^T \hat{\underset{\sim}{\alpha}} + \underset{\sim}{\lambda}_2^T \tilde{\underset{\sim}{\beta}}$$

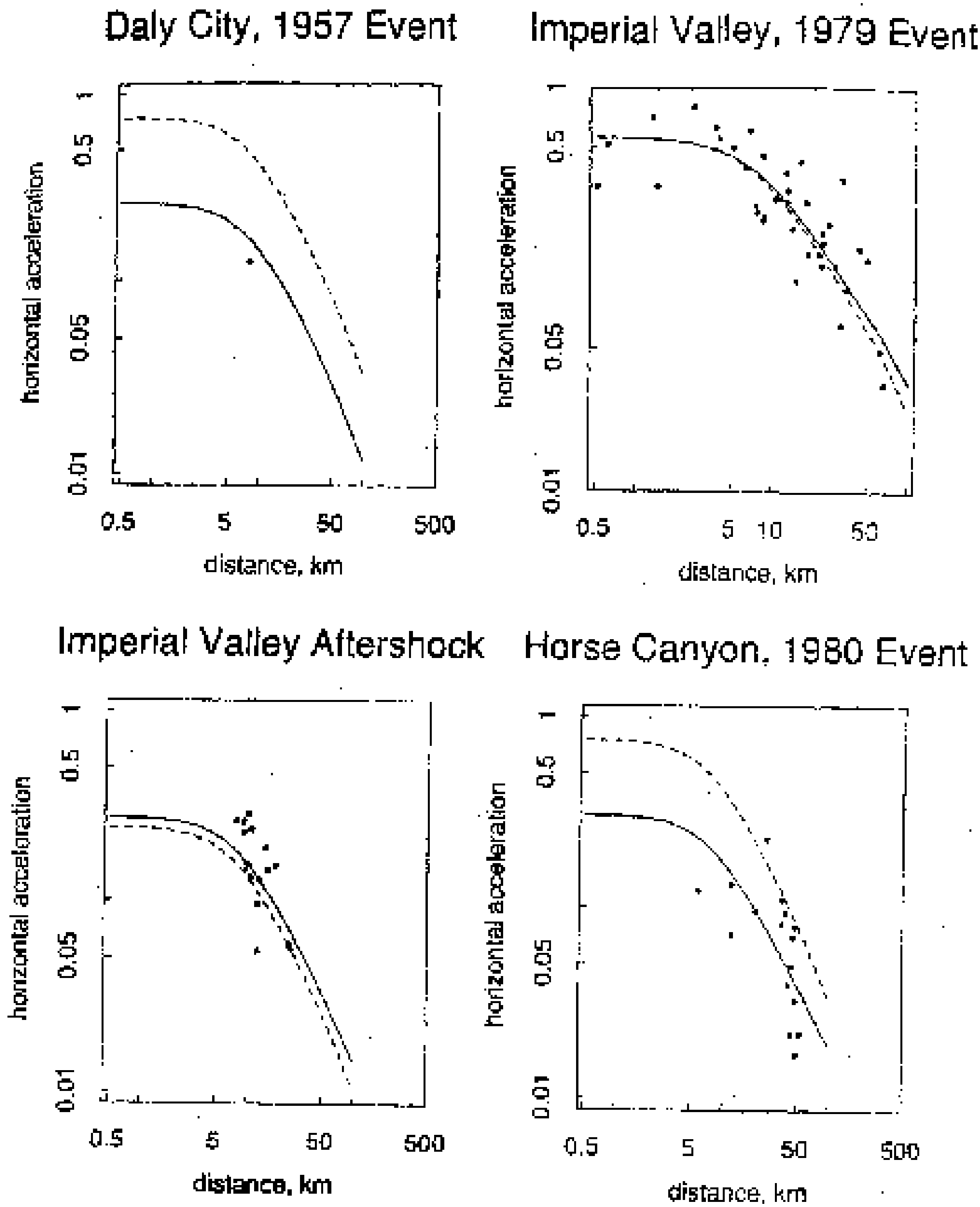

FIG. 1. For the four indicated earthquakes, observed maximum accelerations are plotted at corresponding distances from the event's epicenter. The solid curve is the "borrowed strength" estimate. The dashed one is that of Joyner and Boore (1981).

Development of (*)

Suppose $\underset{\sim}{\alpha} = \underset{\sim}{0}$

Consider estimating $\underset{\sim}{\lambda}^T \underset{\sim}{b}$ (1×g g×1) by $\underset{\sim}{L}^T \underset{\sim}{Y}$ (1×n n×1)

$$\underset{\sim}{L}^T \underset{\sim}{Y} = \underset{\sim}{L}^T \underset{\sim}{Z} \underset{\sim}{b} + \underset{\sim}{L}^T \underset{\sim}{e}$$

$$E\{\underset{\sim}{L}^T \underset{\sim}{Y}\} = \underset{\sim}{0}$$

$$E\{(\underset{\sim}{L}^T \underset{\sim}{Y} - \underset{\sim}{\lambda}^T \underset{\sim}{b})^2\}$$

$$= E\{[(\underset{\sim}{L}^T \underset{\sim}{Z} - \underset{\sim}{\lambda}^T)\underset{\sim}{b}]^2\} + E\{[\underset{\sim}{L}^T \underset{\sim}{e}]^2\}$$

$$= (\underset{\sim}{L}^T \underset{\sim}{Z} - \underset{\sim}{\lambda}^T)\underset{\sim}{D}(\underset{\sim}{Z}^T \underset{\sim}{L} - \underset{\sim}{\lambda}) + \underset{\sim}{L}^T \underset{\sim}{R} \underset{\sim}{L}$$

$$= \underset{\sim}{L}^T \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{Z}^T \underset{\sim}{L} - \underset{\sim}{\lambda}^T \underset{\sim}{D} \underset{\sim}{Z}^T \underset{\sim}{L} - \underset{\sim}{L}^T \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{\lambda} + \underset{\sim}{\lambda}^T \underset{\sim}{\lambda} + \underset{\sim}{L}^T \underset{\sim}{R} \underset{\sim}{L}$$

Differentiate wrt $\underset{\sim}{L}$

$$2 \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{Z}^T \underset{\sim}{L} - \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{\lambda} - \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{\lambda} + 2 \underset{\sim}{R} \underset{\sim}{L} \qquad (**)$$

Set to 0

$$(\underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{Z}^T + \underset{\sim}{R})\underset{\sim}{L} = \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{\lambda}$$

$$\underset{\sim}{L} = (\underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{Z}^T + \underset{\sim}{R})^{-1} \underset{\sim}{Z} \underset{\sim}{D} \underset{\sim}{\lambda}$$

VI

11 Feb. 04

Lemma

$$(Z D Z^T + R)^{-1} Z D = R^{-1} Z (D^{-1} + Z^T R^{-1} Z)^{-1}$$

Proof. Need to show that

$$Z D (D^{-1} + Z^T R^{-1} Z) = (Z D Z^T + R) R^{-1} Z$$

$$LHS = Z + Z D Z^T R^{-1} Z$$

$$RHS = Z D Z^T R^{-1} Z + Z$$

Proof complete

So $$L^T Y = (D^{-1} + Z^T R^{-1} Z)^{-1} Z^T R^{-1} Y$$

$$= \tilde{b}$$

as in (*)

Differentiating (**) gives

$$2 Z D Z^T + 2 R > 0$$

so finding minimum

Estimating $\underset{\sim}{R}$, $\underset{\sim}{D}$

Parametrize them e.g. write

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \sum_{i=1}^{I} \underset{\sim}{Z}_i^{\,n\times q_i} \underset{\sim}{b}_i + \underset{\sim}{e}$$

$$\text{var}\, \underset{\sim}{e} = \sigma^2 \underset{\sim}{I}_n = \underset{\sim}{R}$$

$$\text{var}\, \underset{\sim}{b}_i = \sigma_i^2 \underset{\sim}{I}_{q_i}$$

The $\underset{\sim}{b}_i$ are uncorrelated

$$\underset{\sim}{D} = \text{diag}\{\sigma_i^2 \underset{\sim}{I}_{q_i}\}$$

Estimate σ_i^2 by $\underset{\sim}{b}_i^T \underset{\sim}{b}_i / q_i$ (biased)

Estimate σ^2 by $\|\underset{\sim}{Y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}} - \underset{\sim}{Z}\hat{\underset{\sim}{b}}\|^2/n$

11 Feb. 04

IIII

Uses

1. Looking for distribution of $\underset{\sim}{b}$ to use in MLE
   lme() assumes normal
2. Feeding $\underset{\sim}{\tilde{b}}$ into further computations
3. Predictions

11 08 04

(IX)

Prediction

Future value $\lambda_1^T \alpha + \lambda_2^T b_0 = \Lambda$

$E \hat{\Lambda} = \lambda_1^T \alpha$

$\text{var}\, \hat{\Lambda} = \lambda_1^T D \lambda_1$

$\hat{\Lambda} = \lambda_1^T \hat{\alpha}$

$E \hat{\Lambda} = \lambda_1^T \alpha$

$\text{var}\, \hat{\Lambda} = (\lambda_1^T V^{-1} \lambda_1)^{-}$