Modeling Integrated Data

John M. Abowd
April 2005

© John M. Abowd 2005, all rights reserved

Outline

- Application of the grouping algorithm
- Further discussion of the methods
- Discussion of software

© John M. Abowd 2005, all rights reserved

Characteristics of the Groups

Table 4: Results of Applying the Grouping Algorithm to the Pooled Data Set

	Largest Group	*Second Largest Group*	*Average of All Other Groups*	*Total of All Groups*
Observations	285,402,315	90	4.3	287,241,891
Persons	64,441,382	38	1.5	68,329,212
Firms	3,200,067	8	1.1	3,662,974
Groups	1	1	430,529	430,531
Estimable Effects	67,641,448	45		71,992,185

Notes: The "pooled" data are comprised of annual observations from California, Florida, Illinois, Maryland, Minnesota, North Carolina, and Texas over the period 1986-2000. No single state contributed observations for all years. See Table 1. *Sources:* Author's calculations using the LEHD Program data base.

© John M. Abowd 2005, all rights reserved

Estimation by Direct Solution of Least Squares

- Once the grouping algorithm has identified all estimable effects, we solve for the least squares estimates by direct minimization of the sum of squared residuals.
- This method, widely used in animal breeding and genetics research, produces a unique solution for all estimable effects.

© John M. Abowd 2005, all rights reserved

Least Squares Conjugate Gradient Algorithm

- The matrix Δ is chosen to precondition the normal equations.
- The data matrices and parameter vectors are redefined as shown.

$$\Delta = \text{diagonal elements of} \begin{bmatrix} X'X & X'D & X'F \\ D'X & D'D & D'F \\ F'X & F'D & F'F \end{bmatrix}$$

$$y = \left[X \mid D \mid F\right]\Delta^{-1/2}\Delta^{1/2}\begin{bmatrix} \beta \\ \theta \\ \psi \end{bmatrix} + \varepsilon \equiv Z\delta + \varepsilon$$

$$Z \equiv \left[X \mid D \mid F\right]\Delta^{-1/2} \text{ and } \delta \equiv \Delta^{1/2}\begin{bmatrix} \beta \\ \theta \\ \psi \end{bmatrix}$$

© John M. Abowd 2005, all rights reserved

LSCG (II)

- The goal is to find δ to solve the least squares problem shown.
- The gradient vector g figures prominently in the equations.
- The initial conditions for the algorithm are shown.
  - e is the vector of residuals.
  - d is the direction of the search.

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}}\left[\left(y - Z\delta\right)'\left(y - Z\delta\right)\right]$$

$$0 = \frac{1}{2}\frac{\partial\left(y - Z\delta\right)'\left(y - Z\delta\right)}{\partial\delta} = Z'\left(y - Z\delta\right) \equiv g$$

$$\tau_{-1} = 0$$
$$d_{-1} = 0$$
$$\delta_0 = 0$$
$$e_0 = y - Z\delta_0$$
$$g_0 = Z'e_0 = Z'y - Z'Z\delta_0$$
$$d_0 = g_0$$
$$\rho_0 = g_0'g_0$$
$$\lambda_0 = 0$$

© John M. Abowd 2005, all rights reserved

LSCG (III)

- The loop shown has the following features:
  - The search direction *d* is the current gradient plus a fraction of the old direction.
  - The parameter vector δ is updated by moving a positive amount in the current direction.
  - The gradient, *g*, and residuals, *e*, are updated.
  - The original parameters are recovered from the preconditioning matrix.

$$\text{For } \ell = 0,1,2,3,\dots$$

$$d_\ell = g_\ell + \tau_{\ell-1} d_{\ell-1}$$

$$q_\ell = Z d_\ell$$

$$\lambda_\ell = \rho_\ell / \left(q_\ell ' q_\ell \right)$$

$$\delta_{\ell+1} = \delta_\ell + \lambda_\ell d_\ell$$

$$e_{\ell+1} = e_\ell - \lambda_\ell q_\ell$$

$$g_{\ell+1} = Z' e_{\ell+1}$$

$$\begin{bmatrix} \beta_{\ell+1} \\ \theta_{\ell+1} \\ \psi_{\ell+1} \end{bmatrix} = \Delta^{-1/2} \delta_{\ell+1}$$

© John M. Abowd 2005, all rights reserved

LSCG (IV)

- Verify that the residuals are uncorrelated with the three components of the model.
  - Yes: the LS estimates are calculated as shown.
  - No: certain constants in the loop are updated and the next parameter vector is calculated.

$$\begin{bmatrix} X\beta_{\ell+1} & D\theta_{\ell+1} & F\psi_{\ell+1} \end{bmatrix}' e_{\ell+1} \begin{cases} < \begin{bmatrix} c \\ c \\ c \end{bmatrix}, \text{stop } \hat{\delta} = \delta_{\ell+1} \\ \text{else, continue} \end{cases}$$

$$\rho_{\ell+1} = \left(g_{\ell+1}' g_{\ell+1} \right)$$

$$\tau_{\ell} = \rho_{\ell+1} / \rho_{\ell}$$

$$\begin{bmatrix} \hat{\beta} \\ \hat{\theta} \\ \hat{\psi} \end{bmatrix} = \Delta^{-1/2} \hat{\delta}$$

$$S = \left(y - Z\hat{\delta} \right)' \left(y - Z\hat{\delta} \right)$$

© John M. Abowd 2005, all rights reserved

Mixed Effects Assumptions

$$\Lambda = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \Sigma_N \end{bmatrix} \quad \mathrm{E}\left[\begin{matrix} \theta \\ \psi \end{matrix} \middle| X \right] = 0 \quad \mathrm{V}\left[\begin{matrix} \theta \\ \psi \end{matrix} \middle| X \right] = \Omega$$

- The assumptions above specify the complete error structure with the firm and person effects random.
- For maximum likelihood or restricted maximum likelihood estimation assume joint normality.

© John M. Abowd 2005, all rights reserved

Estimation by Mixed Effects Methods

$$\begin{bmatrix} X'\Lambda^{-1}X & X'\Lambda^{-1}\left[D \mid F\right] \\ \begin{bmatrix} D' \\ -- \\ F' \end{bmatrix}\Lambda^{-1}X & \begin{bmatrix} D' \\ -- \\ F' \end{bmatrix}\Lambda^{-1}\left[D \mid F\right]+\Omega^{-1} \end{bmatrix}\begin{bmatrix} \beta \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} X'\Lambda^{-1}y \\ \begin{bmatrix} D' \\ -- \\ F' \end{bmatrix}\Lambda^{-1}y \end{bmatrix}$$

- Solve the mixed effects equations
- Techniques: Bayesian EM, Restricted ML

© John M. Abowd 2005, all rights reserved

Relation Between Fixed and Mixed Effects Models

$$\Lambda = \sigma_{\varepsilon}^{2} I_{N*} \qquad \left|\Omega\right| \to \infty$$

- Under the conditions shown above, the ME and estimators of all parameters approaches the FE estimator

© John M. Abowd 2005, all rights reserved

Correlated Random Effects vs. Orthogonal Design

$$X'D = 0 \text{ orthogonal personal characteristics and person-effect design}$$
$$X'F = 0 \text{ orthogonal personal characteristics and firm-effect design}$$
$$D'F = 0 \text{ orthogonal person-effect and firm-effect designs}$$

- Orthogonal design means that characteristics, person design, firm design are orthogonal.
- Uncorrelated random effects means that Ω is diagonal.

© John M. Abowd 2005, all rights reserved

Software

- SAS: proc mixed
- ASREML
- aML
- SPSS: Linear Mixed Models
- STATA: xtreg, gllamm, xtmixed
- R: the lme() function
- S+: linear mixed models
- Gauss
- Matlab
- Genstat: REML

© John M. Abowd 2005, all rights reserved