Generalized Least Squares Estimation of Panel Data Models

Paolo Foschi

Dip. di Matematica, Bologna University

The 2-Way Random Effect Model

Intro
● The 2-Way Random Effect Model
Spherical disturbances
Non-spherical disturbances
Time-correlated errors

- The 2-way random effect model is given by

$$\boldsymbol{Y} = \boldsymbol{\iota}_T\, \boldsymbol{\iota}_N^T \alpha + \sum_i \boldsymbol{X}_i \beta_i + \boldsymbol{U}, \qquad\qquad T \times N$$

  with

$$\boldsymbol{U} = \boldsymbol{\lambda} \boldsymbol{\iota}_N^T + \boldsymbol{\iota}_T \boldsymbol{\mu}^T + \boldsymbol{V},$$

  where $\boldsymbol{\iota}_n : n \times 1$ is a vector of all ones, $\boldsymbol{\lambda} : T \times 1$ and $\boldsymbol{\mu} : N \times 1$.
- $\boldsymbol{\lambda} \sim (\boldsymbol{0}, \boldsymbol{\Psi}_\lambda)$, $\boldsymbol{\mu} \sim (\boldsymbol{0}, \boldsymbol{\Psi}_\mu)$ and $\mathrm{Vec}(\boldsymbol{V}) \sim (\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}_{NT})$
- $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are independent
- $\boldsymbol{\Omega} \equiv \mathrm{Cov}(\mathrm{Vec}(\boldsymbol{U})) = \boldsymbol{J}_N \otimes \boldsymbol{\Psi}_\lambda + \boldsymbol{\Psi}_\mu \otimes \boldsymbol{J}_T + \sigma_v^2 \boldsymbol{I}_{NT}$
- where $\boldsymbol{J}_n = \boldsymbol{\iota}_n \boldsymbol{\iota}_n^T$, i.e. a matrix of all ones
- It is a GLM: $\quad \boldsymbol{y} = \boldsymbol{\iota}_{NT} \alpha + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \qquad \boldsymbol{u} \sim (\boldsymbol{0}, \boldsymbol{\Omega})$

State of the art

Spherically Distributed Random Effects:

- $\boldsymbol{U} = \boldsymbol{\lambda}\boldsymbol{\iota}_N^T + \boldsymbol{\iota}_T\boldsymbol{\mu}^T + \boldsymbol{V}$,
- $\mathrm{Cov}(\boldsymbol{\lambda}) = \sigma_\lambda^2\boldsymbol{I}$, $\mathrm{Cov}(\boldsymbol{\mu}) = \sigma_\mu^2\boldsymbol{I}$ and $\mathrm{Cov}(\boldsymbol{v}) = \sigma_v^2\boldsymbol{I}_{NT}$
- $\boldsymbol{\Omega} = \sigma_\lambda^2(\boldsymbol{J} \otimes \boldsymbol{I}) + \sigma_\mu^2(\boldsymbol{I} \otimes \boldsymbol{J}) + \sigma_v^2(\boldsymbol{I} \otimes \boldsymbol{I})$
- Eigendecomposition: $\boldsymbol{\Omega} = \lambda_1\boldsymbol{Q}_1 + \lambda_2\boldsymbol{Q}_2 + \lambda_3\boldsymbol{Q}_3 + \lambda_4\boldsymbol{Q}_4$

Eigenvalue	Multiplicity	Projector
$\lambda_1 = \sigma_v^2$	$(N-1)(T-1)$	$\boldsymbol{Q}_1 = (\boldsymbol{E}_N \otimes \boldsymbol{E}_T)$
$\lambda_2 = (T\sigma_\mu^2 + \sigma_v^2)$	$N-1$	$\boldsymbol{Q}_2 = (\boldsymbol{E}_N \otimes \bar{\boldsymbol{J}}_T)$
$\lambda_3 = (N\sigma_\lambda^2 + \sigma_v^2)$	$T-1$	$\boldsymbol{Q}_3 = (\bar{\boldsymbol{J}}_N \otimes \boldsymbol{E}_T)$
$\lambda_4 = (\sigma_v^2 + N\sigma_\lambda^2 + T\sigma_\mu^2)$	1	$\boldsymbol{Q}_4 = (\bar{\boldsymbol{J}}_N \otimes \bar{\boldsymbol{J}}_T)$

- where $\bar{\boldsymbol{J}}_N = \boldsymbol{\iota}_N\boldsymbol{\iota}_N^T/N$ and $\boldsymbol{E}_N = \boldsymbol{I}_N - \bar{\boldsymbol{J}}_N$

GLS estimation

- The GLS estimator for $\boldsymbol{\beta}$ is computed as
  - OLS on $\sigma_v \boldsymbol{\Omega}^{-1/2}\boldsymbol{y} = \sigma_v \boldsymbol{\Omega}^{-1/2}\boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{v}, \quad \boldsymbol{v} \sim (\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}),$ or
  - GLS on

$$\begin{pmatrix} \boldsymbol{Q}_1\boldsymbol{y} \\ \boldsymbol{Q}_2\boldsymbol{y} \\ \boldsymbol{Q}_3\boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q}_1\boldsymbol{X} \\ \boldsymbol{Q}_2\boldsymbol{X} \\ \boldsymbol{Q}_3\boldsymbol{X} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \tilde{\boldsymbol{u}}_1 \\ \tilde{\boldsymbol{u}}_2 \\ \tilde{\boldsymbol{u}}_3 \end{pmatrix}, \qquad \begin{pmatrix} \lambda_1\boldsymbol{Q}_1 & 0 & 0 \\ 0 & \lambda_2\boldsymbol{Q}_2 & 0 \\ 0 & 0 & \lambda_3\boldsymbol{Q}_3 \end{pmatrix}$$

$$\boldsymbol{Q}_4\boldsymbol{y} = \boldsymbol{\iota}_{NT}\alpha + \boldsymbol{Q}_4\boldsymbol{X}\boldsymbol{\beta} + \tilde{\boldsymbol{u}}, \qquad \lambda_4\boldsymbol{Q}_4$$

- The latter method allows for the direct estimation of λ_i (and, thus, of σ_v^2, σ_λ^2 and σ_μ^2)
- The nr. of observations is now $4NT$, this approach is not optimal
- Strategy: find the basis of the eigenspaces
- Find $(\boldsymbol{w}_N \ \boldsymbol{W}_N)$ orthogonal s.t. $\bar{\boldsymbol{J}}_N = \boldsymbol{w}_N\boldsymbol{w}_N^T$ and $\boldsymbol{E}_N = \boldsymbol{W}_N\boldsymbol{W}_N^T$.

Eigendecomposition of Ω

Intro

Spherical disturbances
- State of the art
- GLS estimation
- Eigendecomposition of Ω
- GLS estimation

Non-spherical disturbances

Time-correlated errors

- $\bar{\boldsymbol{J}}_N = N^{-1}\boldsymbol{\iota}_N\boldsymbol{\iota}_N^T = \boldsymbol{w}_N\boldsymbol{w}_N^T$
- $\boldsymbol{E}_N = \boldsymbol{I}_N - \bar{\boldsymbol{J}}_N = \boldsymbol{W}_N\boldsymbol{W}_N^T$
- A possible solution:

$$\boldsymbol{W}_N\boldsymbol{D}_N = \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 \\ 1 & -1 & -1 & \cdots & -1 \\ 0 & 2 & -1 & \cdots & -1 \\ 0 & 0 & 3 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & N-1 \end{pmatrix} \quad \boldsymbol{D}_N = \mathrm{diag}(\sqrt{i^2+i}, i = 1, \ldots, N-1)$$

- $\boldsymbol{Q}_i = \boldsymbol{P}_i\boldsymbol{P}_i^T$,
where $\boldsymbol{P}_1 = \boldsymbol{W}_N \otimes \boldsymbol{W}_T$, $\boldsymbol{P}_2 = \boldsymbol{W}_N \otimes \boldsymbol{w}_T$, $\boldsymbol{P}_3 = \boldsymbol{w}_N \otimes \boldsymbol{W}_T$ and $\boldsymbol{P}_4 = \boldsymbol{w}_N \otimes \boldsymbol{w}_T$

GLS estimation

- The GLSE of the 2-way ECM is given by the GLSE of

$$\begin{pmatrix} \boldsymbol{P}_4^T\boldsymbol{y} \\ \boldsymbol{P}_1^T\boldsymbol{y} \\ \boldsymbol{P}_2^T\boldsymbol{y} \\ \boldsymbol{P}_3^T\boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \sqrt{NT} \\ 0 \\ 0 \\ 0 \end{pmatrix} \alpha + \begin{pmatrix} \boldsymbol{P}_4^T\boldsymbol{X} \\ \boldsymbol{P}_1^T\boldsymbol{X} \\ \boldsymbol{P}_2^T\boldsymbol{X} \\ \boldsymbol{P}_3^T\boldsymbol{X} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{u}_4 \\ \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \boldsymbol{u}_3 \end{pmatrix},$$

  with covariance matrix

$$\begin{array}{l} \\ 1 \\ (N-1)(T-1) \\ N-1 \\ T-1 \end{array} \begin{array}{cccc} 1 & (N-1)(T-1) & N-1 & T-1 \\ \lambda_4 & 0 & 0 & 0 \\ 0 & \lambda_1\boldsymbol{I} & 0 & 0 \\ 0 & 0 & \lambda_2\boldsymbol{I} & 0 \\ 0 & 0 & 0 & \lambda_3\boldsymbol{I} \end{array}$$

- α can be computed by back-substitution once having computed β
  The first equation can be dropped

Non spherical disturbances

- $\boldsymbol{U} = \boldsymbol{\lambda}\boldsymbol{\iota}_N^T + \boldsymbol{\iota}_T\boldsymbol{\mu}^T + \boldsymbol{V}, \quad \mathrm{Cov}\left(\begin{pmatrix}\boldsymbol{v}\\ \boldsymbol{\lambda}\\ \boldsymbol{\mu}\end{pmatrix}\right) = \begin{pmatrix}\sigma_v^2\boldsymbol{I}_{NT} & 0 & 0\\ 0 & \boldsymbol{\Psi}_\lambda & 0\\ 0 & 0 & \boldsymbol{\Psi}_\mu\end{pmatrix}$
- $\boldsymbol{\Omega} = (\boldsymbol{J} \otimes \boldsymbol{\Psi}_\lambda) + (\boldsymbol{\Psi}_\mu \otimes \boldsymbol{J}) + \sigma_v^2(\boldsymbol{I} \otimes \boldsymbol{I})$
- The 2-way ECM is equivalent to the GLM

$$\begin{pmatrix}\boldsymbol{P}_1^T\boldsymbol{y}\\ \boldsymbol{P}_2^T\boldsymbol{y}\\ \boldsymbol{P}_3^T\boldsymbol{y}\end{pmatrix} = \begin{pmatrix}\boldsymbol{P}_1^T\boldsymbol{X}\\ \boldsymbol{P}_2^T\boldsymbol{X}\\ \boldsymbol{P}_3^T\boldsymbol{X}\end{pmatrix}\boldsymbol{\beta} + \tilde{\boldsymbol{u}}, \qquad \mathrm{Cov}(\tilde{\boldsymbol{u}}) = \begin{pmatrix}\sigma_v^2\boldsymbol{I} & 0 & 0\\ 0 & \boldsymbol{\Omega}_2 & 0\\ 0 & 0 & \boldsymbol{\Omega}_3\end{pmatrix}$$

- $\boldsymbol{\Omega}_2 = T\boldsymbol{W}_N^T\boldsymbol{\Psi}_\mu\boldsymbol{W}_N + \sigma_v^2\boldsymbol{I}, \quad \boldsymbol{\Omega}_3 = N\boldsymbol{W}_T^T\boldsymbol{\Psi}_\lambda\boldsymbol{W}_T + \sigma_v^2\boldsymbol{I}$

GLLSP Approach

- Compute the QRDs $\boldsymbol{Q}_i^T \left(\boldsymbol{P}_i^T \boldsymbol{X} \quad \boldsymbol{P}_i^T \boldsymbol{y} \right) = \begin{pmatrix} \boldsymbol{R}_i & \tilde{\boldsymbol{y}}_i \\ 0 & \hat{\boldsymbol{y}}_i \end{pmatrix}$, for $i = 1, 2, 3$
- Compute the Cholesky decomps: $\boldsymbol{Q}_i^T \boldsymbol{\Omega}_i \boldsymbol{Q}_i = \boldsymbol{C}_i \boldsymbol{C}_i^T$, for $i = 2, 3$,
- Let $\boldsymbol{C}_i = \begin{pmatrix} \widetilde{\boldsymbol{C}}_i & \check{\boldsymbol{C}}_i \\ 0 & \widehat{\boldsymbol{C}}_i \end{pmatrix}$
- Solve the GLLSP: $\min_{\boldsymbol{\beta},\boldsymbol{\eta}} \|\boldsymbol{\eta}\|^2$, subject to

$$\begin{pmatrix} \tilde{\boldsymbol{y}}_1 \\ \tilde{\boldsymbol{y}}_2 \\ \tilde{\boldsymbol{y}}_3 \\ \hat{\boldsymbol{y}}_1 \\ \hat{\boldsymbol{y}}_2 \\ \hat{\boldsymbol{y}}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_2 \\ \boldsymbol{R}_3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{I} & & & 0 & & \\ & \widetilde{\boldsymbol{C}}_2 & & & \check{\boldsymbol{C}}_2 & \\ & & \widetilde{\boldsymbol{C}}_3 & & & \check{\boldsymbol{C}}_3 \\ & & & \boldsymbol{I} & & \\ & & & & \widehat{\boldsymbol{C}}_2 & \\ & & & & & \widehat{\boldsymbol{C}}_3 \end{pmatrix} \boldsymbol{\eta}$$

GLLSP Approach

- Back-substitution: $\check{\boldsymbol{y}}_i = \tilde{\boldsymbol{y}}_i - \check{\boldsymbol{C}}_i \hat{\boldsymbol{C}}_i^{-1} \hat{\boldsymbol{y}}_i$, for $i = 2, 3$
- Solve the GLLSP: $\min_{\boldsymbol{\beta}, \check{\boldsymbol{\eta}}} \|\check{\boldsymbol{\eta}}\|^2$, subject to

$$\begin{pmatrix} \check{\boldsymbol{y}}_1 \\ \check{\boldsymbol{y}}_2 \\ \check{\boldsymbol{y}}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_2 \\ \boldsymbol{R}_3 \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{I} & & \\ & \widetilde{\boldsymbol{C}}_2 & \\ & & \widetilde{\boldsymbol{C}}_3 \end{pmatrix} \check{\boldsymbol{\eta}}$$

- Solved by using an (Updating) GQRD
- Complexity:
    - QRDs of $\boldsymbol{P}_i^T \boldsymbol{X}$: $O(NT\,k^2)$,
    - Cholesky of $\boldsymbol{Q}_i^T \boldsymbol{\Omega}_i \boldsymbol{Q}_i$: $O(N^3 + T^3)$
    - GQRD: $O(k^3)$
- Notice that in several models the Cholesky factors of $\boldsymbol{\Psi}_\lambda$ and/or $\boldsymbol{\Psi}_\mu$ are known. Updating QRDs can be used to factorize $\boldsymbol{\Omega}_i$

Time-correlated errors

- $\boldsymbol{U} = \boldsymbol{\lambda}\boldsymbol{\iota}_N^T + \boldsymbol{\iota}_T\boldsymbol{\mu}^T + \boldsymbol{V}, \quad \mathrm{Cov}\left(\begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix}\right) = \begin{pmatrix} \boldsymbol{I}_N \otimes \boldsymbol{\Psi}_v & 0 & 0 \\ 0 & \boldsymbol{\Psi}_\lambda & 0 \\ 0 & 0 & \boldsymbol{\Psi}_\mu \end{pmatrix}$
- $\boldsymbol{\Omega} = (\boldsymbol{J} \otimes \boldsymbol{\Psi}_\lambda) + (\boldsymbol{\Psi}_\mu \otimes \boldsymbol{J}) + (\boldsymbol{I} \otimes \boldsymbol{\Psi}_v)$
- The 2-way ECM is equivalent to the GLM

$$\begin{pmatrix} \boldsymbol{P}_1^T\boldsymbol{y} \\ \boldsymbol{P}_2^T\boldsymbol{y} \\ \boldsymbol{P}_3^T\boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_1^T\boldsymbol{X} \\ \boldsymbol{P}_2^T\boldsymbol{X} \\ \boldsymbol{P}_3^T\boldsymbol{X} \end{pmatrix}\boldsymbol{\beta} + \tilde{\boldsymbol{u}}, \quad \mathrm{Cov}(\tilde{\boldsymbol{u}}) = \begin{pmatrix} \boldsymbol{I} \otimes \boldsymbol{\Omega}_1 & 0 & \boldsymbol{I} \otimes \boldsymbol{\omega}_{13} \\ 0 & \boldsymbol{\Omega}_2 & 0 \\ \boldsymbol{I} \otimes \boldsymbol{\omega}_{13}^T & 0 & \boldsymbol{\Omega}_3 \end{pmatrix}$$

- $\boldsymbol{\Omega}_1 = \boldsymbol{W}_T^T\boldsymbol{\Psi}_b\boldsymbol{W}_T, \quad \boldsymbol{\omega}_{13} = \boldsymbol{W}_T^T\boldsymbol{\Psi}_\mu\boldsymbol{w}_T$
  $\boldsymbol{\Omega}_2 = T\boldsymbol{W}_N^T\boldsymbol{\Psi}_\mu\boldsymbol{W}_N + \boldsymbol{w}_T^T\boldsymbol{\Psi}_v\boldsymbol{w}_T\boldsymbol{I}, \quad \boldsymbol{\Omega}_3 = \boldsymbol{W}_T^T(N\boldsymbol{\Psi}_\lambda + \boldsymbol{\Psi}_v)\boldsymbol{W}_T$

Computational aspects

- Need to compute the GQRD of

$$\begin{pmatrix} \boldsymbol{P}_1^T \boldsymbol{X} \\ \boldsymbol{P}_2^T \boldsymbol{X} \\ \boldsymbol{P}_3^T \boldsymbol{X} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \boldsymbol{I} \otimes \boldsymbol{C}_1 & 0 & \boldsymbol{I} \otimes \boldsymbol{c}_{12} \\ 0 & \boldsymbol{C}_2 & 0 \\ 0 & 0 & \boldsymbol{C}_3 \end{pmatrix}$$

- $\boldsymbol{\Omega}_i = \boldsymbol{C}_i \boldsymbol{C}_i^T$ are Cholesky decompositions.
- Done efficiently by splitting $\boldsymbol{P}_1^T \boldsymbol{X}$ in blocks of T observations.
  1. For each block the corresponding GQRD is computed
  2. All the GQRDs are combined by means of updating techniques
- Computationally efficient when the model is reesimated for different covariance matrices