Matrix Computations and Statistics

Institut d'informatique Université de Neuchâtel, Switzerland

- Erricos Kontoghiorghes (Numerical Linear Algebra, Computational Statistics & Econometrics, Parallelism)
- Paolo Foschi (Numerical methods for estimating econometric models)
- Cristian Gatu (Algorithms for solving statistical model selection problems)
- Petko Ianev (Parallel matrix algorithms and econometrics)

Linear models

- Ordinary Linear Model (OLM): $y = Ax + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_m)$

- General Linear Model (GLM): $y = Ax + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 \Omega)$

- Seemingly Unrelated Regression (SUR):

  $\mathsf{vec}(Y) = \left(\bigoplus_{i=1}^{G} X_i\right)\mathsf{vec}(\{\beta_i\}_G) + \mathsf{vec}(U), \quad \mathsf{vec}(U) \sim N(0, \Sigma \otimes I_T)$

- Simultaneous Equations Models:

  $\mathsf{vec}(Y) = \left(\bigoplus_{i=1}^{G} W S_i\right)\mathsf{vec}(\{\delta_i\}_G) + \mathsf{vec}(U), \quad \mathsf{vec}(U) \sim N(0, \Sigma \otimes I_T)$

  $W \equiv \begin{pmatrix} X & Y \end{pmatrix}$ and S_i is a selection matrix.

Linear models

- Least Squares (LS) solutions.
- Modified linear models (up– and down–dating).
- Linear equality constraints.
- All possible regression models.
- Multiple-row diagnostics
- Cross-validation.

Computing the QRD of a long dense matrix

- Compute the QRD: $Q^T A = \begin{pmatrix} R \\ 0 \end{pmatrix}$

- Block-generalization of the Givens sequences.

- Partition $A \in \Re^{m\times n}$ as $A^T = \begin{pmatrix} A_1^T & A_2^T & \ldots & A_\mu^T \end{pmatrix}$

- $A_i \in \Re^{n\times n}$ $(i = 1, \ldots, \mu)$ and $m = \mu n$.

Block-updating and Bitonic QRD strategies

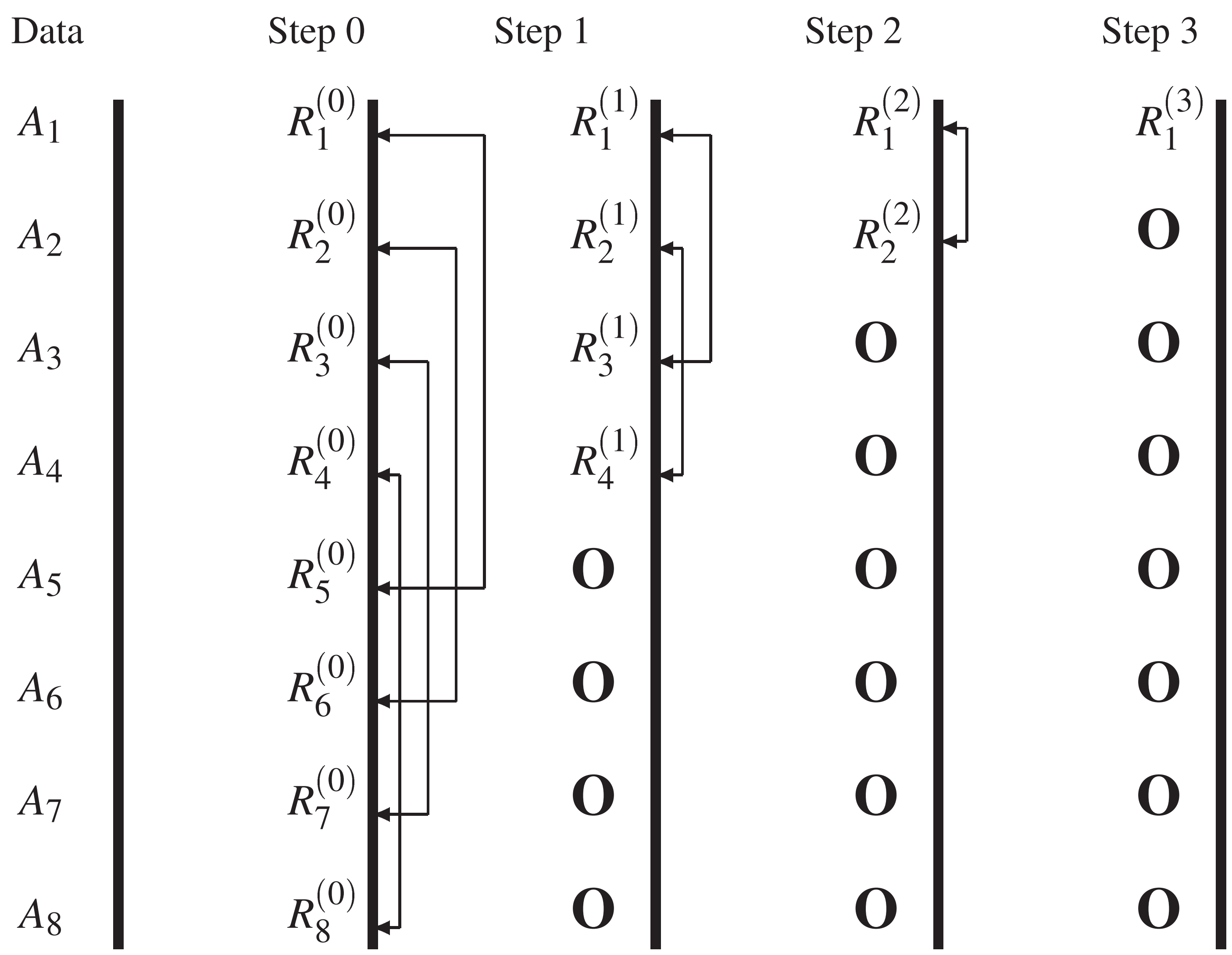
1: Compute the QRD $A_\mu = Q_\mu^T R_\mu$
2: **for** $i = \mu - 1, \ldots, 2, 1$ **do**
3: compute $\begin{pmatrix} R_i \\ 0 \end{pmatrix} \leftarrow Q_i^T \begin{pmatrix} A_i \\ R_{i+1} \end{pmatrix}$
4: **end for**

1: Compute simultaneously the QRDs $A_j = Q_{j,0}^T R_j^{(0)} \quad (j = 1, \ldots, 2^g)$
2: **for** $i = 1, 2, \ldots, g$ **do**
3: **for** $j = 1, \ldots, 2^{(g-i)}$ **do-in-parallel**
4: Compute the QRDs $\begin{pmatrix} R_j^{(i)} \\ 0 \end{pmatrix} \leftarrow Q_{j,i}^T \begin{pmatrix} R_j^{(i-1)} \\ R_{j+2^{(g-i)}}^{(i-1)} \end{pmatrix}$
5: **end for**
6: **end for**

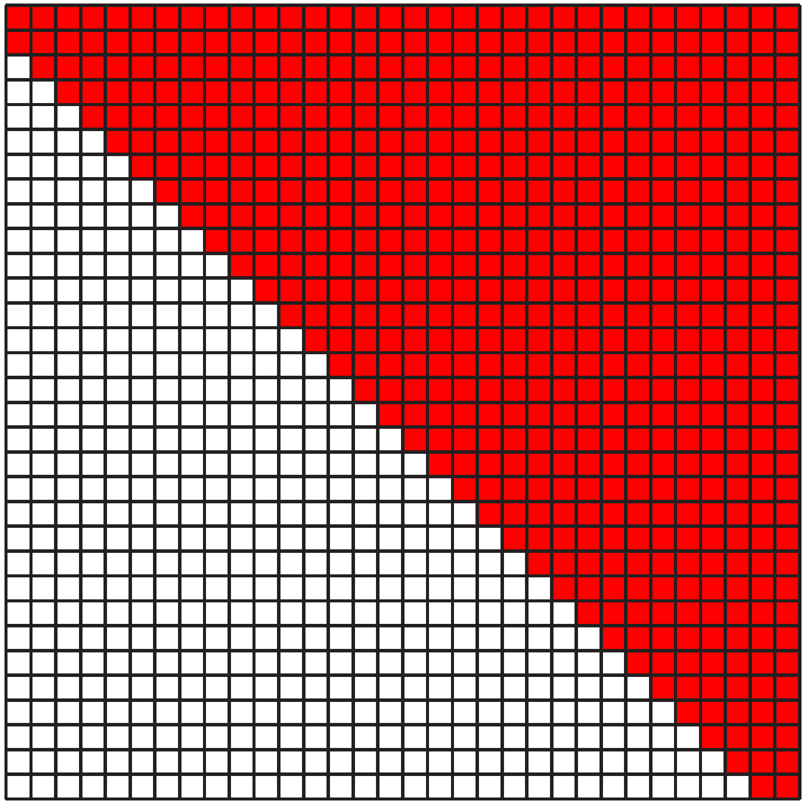
The Block-updating strategy

Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
A_1	A_1	A_1	A_1	A_1	A_1	A_1	R_1
A_2	A_2	A_2	A_2	A_2	A_2	R_2	$\mathbf{O}$
A_3	A_3	A_3	A_3	A_3	R_3	$\mathbf{O}$	$\mathbf{O}$
A_4	A_4	A_4	A_4	R_4	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$
A_5	A_5	A_5	R_5	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$
A_6	A_6	R_6	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$
A_7	R_7	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$
A_8	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$	$\mathbf{O}$

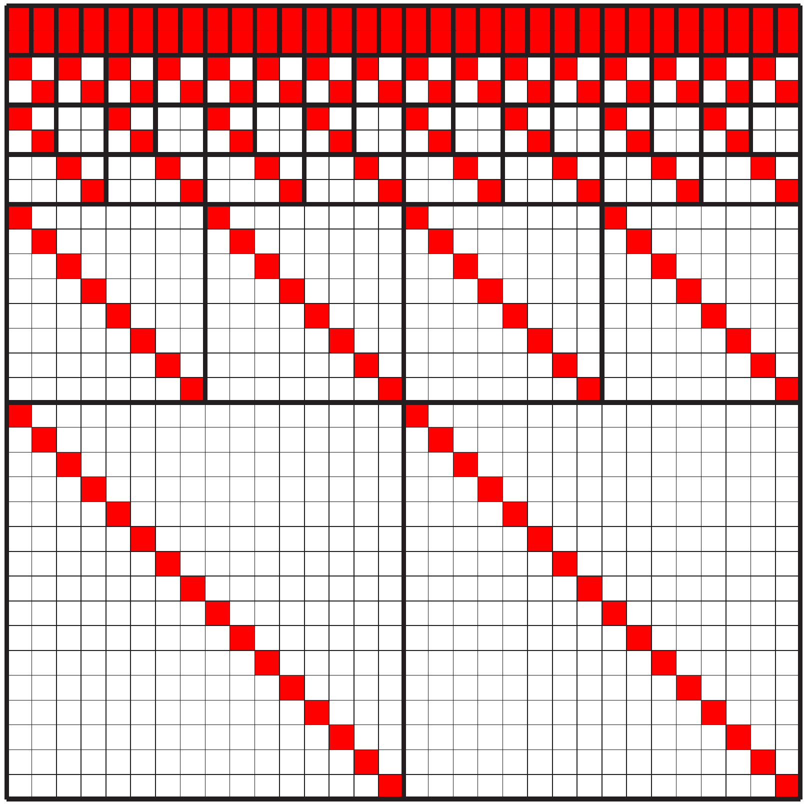
The Bitonic strategy



Structure of orthogonal matrices



Block-updating QRD algorithm



Bitonic QRD algorithm

Recursive structure of Q

- The structure of Q^T is given by:

$$S^{(g)} \stackrel{str.}{=} \begin{array}{cc} \begin{array}{cc} 2^{(g-1)}n & 2^{(g-1)}n \end{array} & \\ \begin{pmatrix} S^{(g-1)} & S^{(g-1)} \\ D^{(g-1)} & D^{(g-1)} \end{pmatrix} & \begin{array}{c} 2^{(g-1)}n \\ 2^{(g-1)}n \end{array} \end{array} \quad \text{for} \quad g \geq 1.$$

- $D^{(g)}$ is a $2^g n \times 2^g n$ block-diagonal matrix with each block having dimension $n \times n$.

- Q has the minimum fill-in $(g+1)2^g n^2$.

Block downdating the OLM: $y = Ax + \varepsilon$

- The LS solution is given by $x = R^{-1}u$, where the QRD of $\hat{A} = \begin{pmatrix} A & y \end{pmatrix} \in \Re^{m\times n}$ is:

$$Q^T\hat{A} = \begin{array}{l} \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix} \begin{matrix} n \\ m-n \end{matrix} \end{array}, \quad \text{with} \quad \hat{R} = \begin{pmatrix} R & u \\ 0 & s \end{pmatrix} \begin{matrix} n-1 \\ 1 \end{matrix}$$

- Delete a block of $m-k$ observations from the OLM:

$$\Pi^T\hat{A} \equiv \begin{pmatrix} A & y \end{pmatrix} \equiv \begin{pmatrix} \hat{A}_1 \\ \hat{A}_2 \end{pmatrix} \equiv \begin{pmatrix} A_1 & y_1 \\ A_2 & y_2 \end{pmatrix} \begin{matrix} m-k \\ k \end{matrix} \quad \text{and} \quad \Pi^T\varepsilon \equiv \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \begin{matrix} m-k \\ k \end{matrix}.$$

- Compute the LS solution of: $\quad y_2 = A_2x + \varepsilon_2, \quad \varepsilon_2 \sim N(0, \sigma^2 I_k).$

- Re-compute the QRD of $\hat{A}_2$:

$$Q_n^T\hat{A}_2 = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}, \quad \text{with} \quad \tilde{R} \equiv \begin{pmatrix} R_n & u_n \\ 0 & s_n \end{pmatrix} \begin{matrix} n-1 \\ 1 \end{matrix}$$

Block downdating LS

- Let the first $m-k$ columns of $\Pi^T Q^T$ be given by $\begin{pmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{pmatrix} \begin{matrix} n \\ m-k \\ k-n \end{matrix}$

- Compute the QRD: $H^T \begin{pmatrix} Q_2^T \\ Q_3^T \end{pmatrix} = \begin{pmatrix} Z \\ 0 \end{pmatrix} \begin{matrix} m-k \\ k-n \end{matrix}$

- Update the QRD: $G^T \begin{pmatrix} Q_1^T & \hat{R} \\ Z & 0 \end{pmatrix} = \begin{pmatrix} D & E \\ 0 & B \end{pmatrix} \begin{matrix} m-k \\ n \end{matrix}$ where $|D| = I_{m-k}$.

- $\tilde{Q}^T B = \tilde{R}$

The UGS-1

●	o	o	o	o	o	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
18	●	o	o	o	o	18	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
17	19	●	o	o	o	19	17	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
16	18	20	●	o	o	20	18	16	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
15	17	19	21	●	o	21	19	17	15	●	●	●	●	●	●	●	●	●	●	●	●	●	●
14	16	18	20	22	●	22	20	18	16	14	●	●	●	●	●	●	●	●	●	●	●	●	●
13	15	17	19	21	23	23	21	19	17	15	13	●	●	●	●	●	●	●	●	●	●	●	●
12	14	16	18	20	22		22	20	18	16	14	12	●	●	●	●	●	●	●	●	●	●	●
11	13	15	17	19	21			21	19	17	15	13	11	●	●	●	●	●	●	●	●	●	●
10	12	14	16	18	20				20	18	16	14	12	10	●	●	●	●	●	●	●	●	●
9	11	13	15	17	19					19	17	15	13	11	9	●	●	●	●	●	●	●	●
8	10	12	14	16	18						18	16	14	12	10	8	●	●	●	●	●	●	●
7	9	11	13	15	17							17	15	13	11	9	7	●	●	●	●	●	●
6	8	10	12	14	16								16	14	12	10	8	6	●	●	●	●	●
5	7	9	11	13	15									15	13	11	9	7	5	●	●	●	●
4	6	8	10	12	14										14	12	10	8	6	4	●	●	●
3	5	7	9	11	13											13	11	9	7	5	3	●	●
2	4	6	8	10	12												12	10	8	6	4	2	●
1	3	5	7	9	11													11	9	7	5	3	1
	2	4	6	8	10														10	8	6	4	2
		3	5	7	9															9	7	5	3
			4	6	8																8	6	4
				5	7																	7	5
					6																		6

Computing all possible subset regression models

- OLM: $y = Ax + \varepsilon$
- QRD: $Q^T A = \begin{pmatrix} R \\ 0 \end{pmatrix} \begin{matrix} n \\ m-n \end{matrix}$ and $Q^T y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \begin{matrix} n \\ m-n \end{matrix}$.
- $\text{RSS} = y_2^T y_2$.
- Let S denote a selection matrix, $A_{(S)} = AS$ and $x_{(S)} = S^T x$.
- Compute the $\text{RSS}_{(S)}$ of $y = A_{(S)} x_{(S)} + \varepsilon$ for all $2^n - 1$ possible S.
- QRD: $Q_{(S)}^T RS = \begin{pmatrix} R_{(S)} \\ 0 \end{pmatrix} \begin{matrix} k \\ n-k \end{matrix}$ and $Q_{(S)}^T y_1 = \begin{pmatrix} \tilde{y}_1 \\ \hat{y}_1 \end{pmatrix} \begin{matrix} k \\ n-k \end{matrix}$,
- $\text{RSS}_{(S)} = \text{RSS} + \hat{y}_1^T \hat{y}_1$.

Multiple-row diagnostics and cross-validation

- OLM: $y = Ax + \varepsilon$, where $(A\ y) \in \Re^{m\times(n+1)}$
- Identify the influential data $B \in \Re^{m^*\times(n+1)}$ where $m^* \ll m$.
- Investigate all possible subset of k observations (rows) taken from B ($k = 1,\ldots,\mu \le m^*$)
- The number of cases: $\sum_{k=1}^{\mu} \binom{m^*}{k}$.
- Cross validation: divide the m observations in two disjoint sets.
- The first set comprises m_c observations which are used to construct the model.
- The remaining $m_v = m - m_c$ observations comprise the second set and are used to validate the constructed model.
- There are $\binom{m}{m_v}$ possible cases.
- $m_c m_v$ cases can be generated by exchanging only one single observation from the initial model.
- Similar strategy used in all possible subset models.

GLM: $y = Ax + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2\Omega)$.

- GLLSP: $\underset{\upsilon,x}{\text{argmin}}\ \upsilon^T\upsilon$ subject to $y = Ax + C\upsilon$
- $\Omega = CC^T \quad C\upsilon = \varepsilon$ and $\upsilon \sim N(0, \sigma^2 I_m)$.
- GQRD of $\widetilde{A} \equiv (A \quad y)$ and C:

$$Q^T\widetilde{A} = \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix} \equiv \begin{array}{r} \\ n-1 \\ 1 \\ m-n \end{array} \begin{array}{c} \begin{array}{cc} n-1 & 1 \end{array} \\ \begin{pmatrix} R & \widetilde{y} \\ 0 & \eta \\ 0 & 0 \end{pmatrix} \end{array} \quad \text{and} \quad (Q^TC)P = L^T \equiv \begin{array}{c} \begin{array}{ccc} n-1 & 1 & m-n \end{array} \\ \begin{pmatrix} L_{11}^T & f & L_{12}^T \\ 0 & \delta & d^T \\ 0 & 0 & L_{22}^T \end{pmatrix} \end{array}$$

- $\underset{\upsilon_1,\nu,\upsilon_2,x}{\text{argmin}} \left(\|\upsilon_1\|^2 + \nu^2 + \|\upsilon_2\|^2\right)$ subject to $\begin{cases} \widetilde{y} = Rx + L_{11}^T\upsilon_1 + f\nu + L_{12}^T\upsilon_2 \\ \eta = \delta\nu + d^T\upsilon_2 \\ 0 = L_{22}^T\upsilon_2 \end{cases}$
- $\upsilon_2 = 0, \quad \nu = \eta/\delta$ and $Rx = \widetilde{y} - (\eta/\delta)f$.

Block-updating strategy

- GLLSP: $\underset{\upsilon,x}{\operatorname{argmin}}\ \upsilon^T\upsilon$ subject to $y = Ax + C\upsilon$

- $$Q_i^T\widetilde{A} = \begin{pmatrix} \widetilde{A}_i \\ 0 \end{pmatrix} \equiv \begin{array}{r} \\ m-m_i \\ m_i \end{array} \begin{array}{cc} n-1 & 1 \\ \multicolumn{2}{c}{\begin{pmatrix} A_i & y_i \\ 0 & 0 \end{pmatrix}} \end{array} \quad \text{and} \quad (Q_i^T C)P_i = \widetilde{L}_i^T \equiv \begin{array}{cc} m-m_i & m_i \\ \multicolumn{2}{c}{\begin{pmatrix} L_i^T & \widetilde{L}_{12}^T \\ 0 & \widetilde{L}_{22}^T \end{pmatrix}} \end{array}$$

- $$\underset{\upsilon_i,\widetilde{\upsilon}_i,x}{\operatorname{argmin}}\left(\|\upsilon_i\|^2 + \|\widetilde{\upsilon}_i\|^2\right) \text{ subject to } \begin{pmatrix} y_i \\ 0 \end{pmatrix} = \begin{pmatrix} A_i \\ 0 \end{pmatrix}\beta + \begin{pmatrix} L_i^T & \widetilde{L}_{12}^T \\ 0 & \widetilde{L}_{22}^T \end{pmatrix}\begin{pmatrix} \upsilon_i \\ \widetilde{\upsilon}_i \end{pmatrix}.$$

- $\upsilon^T P_i = (\upsilon_i^T\ \widetilde{\upsilon}_i^T)$, $\widetilde{L}_{22}^T\widetilde{\upsilon}_i = 0$ and $\widetilde{\upsilon}_i = 0$.

- Reduced size GLLSP: $\underset{\upsilon_i,x}{\operatorname{argmin}}\|\upsilon_i\|^2$ subject to $y_i = A_i x + L_i^T\upsilon_i$

- Use the block-updating strategy to solve smaller in size GLLSP at each step.

- The RQD of C is not (explicitly) computed.

Illustration GLLSP

- step 0: $$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} w + \begin{pmatrix} C_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} \\ 0 & C_{2,2} & C_{2,3} & C_{2,4} \\ 0 & 0 & C_{3,3} & C_{3,4} \\ 0 & 0 & 0 & C_{4,4} \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \\ \upsilon_4 \end{pmatrix}$$

- step 1: $$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} w + \begin{pmatrix} C_{1,1} & C_{1,2} & C_{1,3} \\ 0 & C_{2,2} & C_{2,3} \\ 0 & 0 & C_{3,3} \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \\ \upsilon_3 \end{pmatrix}$$

- step 2: $$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} w + \begin{pmatrix} C_{1,1} & C_{1,2} \\ 0 & C_{2,2} \end{pmatrix} \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \end{pmatrix}$$

- step 3: $y_1 = A_1 w + C_{1,1} \upsilon_1$

Augmented systems

- $$\begin{pmatrix} CC^T & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} q \\ b \end{pmatrix}.$$

- Equivalent formulation of the Quadratic Programming problem

  $$\underset{x}{\operatorname{argmin}} \frac{1}{2} x^T H x - q^T x \quad \text{subject to} \quad A^T x = b \qquad (H = CC^T).$$

- w is the Lagrangian vector associated with the constraints.

- If $b = 0$, the solution of the AS also solves the Generalized Linear Least squares problem (GLLSP)

  $$\underset{w,\upsilon}{\operatorname{argmin}} \upsilon^T \upsilon \quad \text{subject to} \quad q = Aw + C\upsilon$$

- x is the Lagrangian vector associated with the constraints and $\upsilon = B^T x$.

Seemingly Unrelated Regressions (SUR)

- $y_i = X_i\beta_i + u_i, \qquad i = 1,\ldots,G.$
- $y_i, u_i \in \Re^T, \quad X_i \in \Re^{T\times k_i}, \quad \beta_i \in \Re^{k_i}$ and $\mathrm{E}(u_i, u_j^T) = \sigma_{i,j} I_T.$
- $\begin{pmatrix} y_1 \\ \vdots \\ y_G \end{pmatrix} = \begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_G \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_G \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_G \end{pmatrix}$
- $\mathsf{vec}(Y) = \left(\bigoplus_{i=1}^G X_i\right)\mathsf{vec}(\{\beta_i\}_G) + \mathsf{vec}(U), \qquad \mathsf{vec}(U) \sim N(0, \Sigma \otimes I_T)$
- $\mathrm{E}\left(\mathsf{vec}(U)\,\mathsf{vec}(U)^T\right) = \Sigma \otimes I_T \equiv \begin{pmatrix} \sigma_{1,1}I_T & \sigma_{1,2}I_T & \ldots & \sigma_{1,G}I_T \\ \sigma_{2,1}I_T & \sigma_{2,2}I_T & \ldots & \sigma_{2,G}I_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{G,1}I_T & \sigma_{G,2}I_T & \ldots & \sigma_{G,G}I_T \end{pmatrix}$

GLLSP

- $\underset{V,\{\beta_i\}}{\operatorname{argmin}} \|V\|_F^2$ subject to $\mathsf{vec}(Y) = \left(\oplus_i X_i\right)\mathsf{vec}(\{\beta_i\}) + \mathsf{vec}(VC^T)$
- $\Sigma = CC^T, \quad C \in \Re^{G\times g}, \quad \operatorname{rank}(\Sigma) = g, \quad V \in \Re^{T\times g}, \quad VC^T = U, \quad \mathsf{vec}(V) \sim N(0, I_{gT})$.
- GQRD of $\oplus_i X_i$ and $(C \otimes I_T)$:

$$Q^T \oplus_i X_i = \begin{array}{c} \begin{pmatrix} \oplus_i R_i \\ 0 \end{pmatrix} \end{array} \begin{array}{l} K \\ GT-K \end{array} \quad \text{and} \quad Q^T\left(C\otimes I_T\right)P = \begin{array}{c} \begin{array}{cc} gT-q & q \end{array} \\ \begin{pmatrix} W_{11} & W_{12} \\ 0 & W_{22} \\ 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} K \\ q \\ GT-K-q \end{array}$$

- $\underset{\tilde{V},\hat{V},\{\beta_i\}}{\operatorname{argmin}} \|\tilde{V}\|^2 + \|\hat{V}\|^2$ s.t. $\begin{pmatrix} \mathrm{vec}(\{\tilde{y}_i\}) \\ \hat{y} \\ \check{y} \end{pmatrix} = \begin{pmatrix} \oplus_i R_i \\ 0 \\ 0 \end{pmatrix} \mathsf{vec}(\{\beta_i\}) + \begin{pmatrix} W_{11} & W_{12} \\ 0 & W_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{V} \\ \hat{V} \end{pmatrix}.$
- $P^T\mathsf{vec}(V) = \begin{pmatrix} \tilde{V} \\ \hat{V} \end{pmatrix}, \quad X_i = \begin{pmatrix} Q_{A,i} & Q_{B,i} \end{pmatrix} \begin{pmatrix} R_i \\ 0 \end{pmatrix}$ and $Q = \begin{pmatrix} \oplus_i Q_{A,i} & \oplus_i Q_{B,i} \end{pmatrix}$.

Common regressors

- $\text{vec}(Y) = \left(\oplus_{i=1}^{G} X^d S_i\right) \text{vec}(\{\beta_i\}_G) + \text{vec}(U)$

- $X^d \in \Re^{T \times K^d}$, $X_i = X^d S_i$, and $S_i \in \Re^{K^d \times k_i}$ is a selctor matrix.

- $Q_d^T X^d = \begin{array}{cc} K^d & \\ \begin{pmatrix} R^d \\ 0 \end{pmatrix} & \begin{array}{c} r \\ T-r \end{array} \end{array}$ with $Q_d = \begin{array}{c} \begin{array}{cc} r & T-r \end{array} \\ \begin{pmatrix} \widetilde{Q}_d & \widehat{Q}_d \end{pmatrix} \end{array}$

- Premultiply the SUR model by $Q_D = \begin{pmatrix} I_G \otimes \widetilde{Q}_d & I_G \otimes \widehat{Q}_d \end{pmatrix}$

- $\begin{pmatrix} \text{vec}(\widetilde{Y}) \\ \text{vec}(\widehat{Y}) \end{pmatrix} = \begin{pmatrix} \oplus_i R^d S_i \\ 0 \end{pmatrix} \text{vec}(\{\beta_i\}) + \begin{pmatrix} \text{vec}(\widetilde{U}) \\ \text{vec}(\widehat{U}) \end{pmatrix}$, $\text{vec}((\widetilde{U}\ \widehat{U})) \sim \begin{pmatrix} \Sigma \otimes I_r & 0 \\ 0 & \Sigma \otimes I_{T-r} \end{pmatrix}$

- $\text{vec}(\widetilde{Y}) = \left(\oplus_i R^d S_i\right) \text{vec}(\{\beta_i\}) + \text{vec}(\widetilde{U})$, $\text{vec}(\widetilde{U}) \sim (0, \Sigma \otimes I_r)$.

Special cases

- Triangular SUR models: $X_i = \begin{pmatrix} X_{i-1} & x_i \end{pmatrix}$
- Orthogonal regressors: $X_i^T X_j = 0$ for $i \neq j$.
- Covariance restrictions: $0 < \rho_{i,j} < 1$, where $\rho_{i,j} = \sigma_{i,j} / \sqrt{\sigma_{i,i}\sigma_{j,j}}$
- Vector Autoregressive disturbances: $U = ZUA + E$ where $\mathrm{vec}(E) \sim (0, \Sigma \otimes I_T)$.
- VAR models: X^d is block-Toeplitz.
- Unequal size (and missing) observations: $X_i \in \Re^{T_i \times k_i}$.
- Various structures of: Σ

Simultaneous equations models

- $y_i = X_i\beta_i + Y_i\gamma_i + u_i = W_i\delta_i + u_i \qquad i = 1,\ldots,G.$

- $\mathsf{vec}(Y) = \left(\bigoplus_{i=1}^{G} WS_i\right)\mathsf{vec}(\{\delta_i\}_G) + \mathsf{vec}(U) \qquad \mathsf{vec}(U) \sim (0, \Sigma \otimes I_T).$

- Transformed SEM: $\mathsf{vec}(X^TY) = \left(\oplus_i X^TWS_i\right)\mathrm{vec}(\{\delta_i\}) + \mathsf{vec}(X^TU)$

- $W \equiv \begin{pmatrix} X & Y \end{pmatrix} \in \Re^{T\times(K+G)}, \quad \mathsf{vec}(X^TU) \sim (0, \Sigma \otimes X^TX).$

- QRD of W: $Q^TW = \begin{array}{cc} \begin{array}{cc} K & G \end{array} & \\ \begin{pmatrix} R^{(1)} & R^{(2)} \\ 0 & R^{(3)} \end{pmatrix} & \begin{array}{l} K \\ T-K \end{array} \end{array}$, with $Q^T = \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix} \begin{array}{l} K \\ T-K \end{array}$

- TSEM: $\mathsf{vec}(R^{(2)}) = \left(\oplus_i RS_i\right)\mathrm{vec}(\{\delta_i\}) + \mathsf{vec}(\tilde{U}) \qquad \mathsf{vec}(\tilde{U}) \sim (0, \Sigma \otimes I_K).$

- $\underset{\{\delta_i\},V}{\mathrm{argmin}} \|V\|_F^2$ subject to $\mathsf{vec}(R^{(2)}) = \left(\oplus_i RS_i\right)\mathrm{vec}(\{\delta_i\}) + \mathsf{vec}(V\hat{C}^T)$