Algoritmos MC2E para la resolución de Modelos de Ecuaciones Simultaneas mediante la descomposición QR

Jose Juan López Espín (Un. Miguel Hernández)

Antonio M. Vidal (Un. Politécnica de Valencia)
Domingo Giménez (Un. de Murcia)
SEIO 10

Septiembre de 2010

Esquema

Introducción

- Los Modelos de Ecuaciones Simultáneas (M.E.S.) nacen como herramienta econométrica.
- Limitados por su gran necesidad computacional.
- Actualmente han comenzado a usarse en otras disciplinas.
- No conocemos la existencia de software en paralelo para la resolución de M.E.S. (aunque sí en secuencial).
- Solo se conocen estudios sobre algoritmos paralelos para M.E.S. de estimadores de información completa.
- La utilización de la descomposición QR puede reducir el coste en secuencial de la resolución de un M.E.S.
- Se han desarrollado metaheurísticas que obtienen un M.E.S satisfactorio a partir de un conjunto de datos de variables.

¿Que es un M.E.S.?

El esquema de un modelo con N ecuaciones, N variables endógenas y K variables exógenas en forma matricial es:

$$BY^T + \Gamma X^T + u^T = 0$$

siendo:

$$Y = (y_1 \dots y_N) \ , \ X = (x_1 \dots x_K) \ , \ u = (u_1 \dots u_N)$$
$$B = \begin{pmatrix} \beta_{1,1} & \dots & \beta_{1,N} \\ & \dots & \\ \beta_{N,1} & \dots & \beta_{N,N} \end{pmatrix}, \ \Gamma = \begin{pmatrix} \gamma_{1,1} & \dots & \gamma_{1,K} \\ & \dots & \\ \gamma_{N,1} & \dots & \gamma_{N,K} \end{pmatrix}$$

matrices densas con valores reales, siendo $\beta_{i,i} = -1 \ \forall \ i = 1, ..., N$.
El modelo estructural se puede expresar también en forma reducida:
$Y = X\Pi + v$, con $\Pi^T = -B^{-1}\Gamma \ , \ v^T = -B^{-1}u^T$.

¿Que es un M.E.S.?

Si se descompone B de la forma $B = diag(-1, ..., -1) + \tilde{B}$, la expresión anterior queda de la siguiente forma:

$$Y^T = \tilde{B} Y^T + \Gamma X^T + u^T$$

O en ecuaciones:

$$\begin{gathered}
y_1 = \gamma_{1,1} x_1 + \ldots + \gamma_{1,K} x_K + \beta_{1,2} y_2 + \beta_{1,3} y_3 + \ldots + \beta_{1,N} y_N + u_1 \\
y_2 = \gamma_{2,1} x_1 + \ldots + \gamma_{2,K} x_K + \beta_{2,1} y_1 + \beta_{2,3} y_3 + \ldots + \beta_{2,N} y_N + u_2 \\
\ldots \\
y_N = \gamma_{N,1} x_1 + \ldots + \gamma_{N,K} x_K + \beta_{N,1} y_1 + \ldots + \beta_{N,N-1} y_{N-1} + u_N
\end{gathered}$$

donde $x_1, x_2, \ldots, x_K$ son variables exógenas, $y_1, y_2, \ldots, y_N$ son variables endógenas, y $u_1, u_2, \ldots, u_N$ son variables de ruido blanco. Todas ellas son vectores de dimensión d, siendo d el tamaño de la muestra.

Mínimos Cuadrados en Dos Etapas (MC2E)

Objetivo: Estimar los valores de $\tilde{B}$ y Γ en la expresión:

$$Y^T = \tilde{B}Y^T + \Gamma X^T + u^T$$

Problema: No es posible usar Mínimos Cuadrados (Y correlac. con u).
Solución: Se sustituyen las variables endogenas en cada ecuación (Y_i) por otras variables llamadas *proxy* ($\hat{Y}_i$).
La matriz de datos asociada a la i-th equation es denotada por $Z_i = [X_i|\hat{Y}_i]$ y puede ser expresada usando una matriz de selección S_i de forma que $Z_i = [X|\hat{Y}]S_i$, donde $S_i \in \mathbb{R}^{(K+N)\times(k_i+n_i-1)}$.
Una vez sustituidas las variables *proxy*, $B_{i,:}$ y $\Gamma_{i,:}$ pueden ser estimadas resolviendo $min||y_i - Z_i\eta_i||$, donde y_i es la variable endogena principal is y $\eta_i^T = [\Gamma_{i,:}|B_{i,:}]S_i \in \mathbb{R}^{k_i+n_i-1}$.

Mínimos Cuadrados en Dos Etapas (MC2E)

El siguiente algoritmo muestra un esquema del estimador MC2E:

Algoritmo 1 Algoritmo $MC2E_{bas}$

Entrada: $X \in \mathbb{R}^{d\times K}$, $Y \in \mathbb{R}^{d\times N}$ y $S_i \in \mathbb{R}^{(K+N)\times(k_i+n_i-1)}$ $\forall i = 1, ..., N$
Salida: $B \in \mathbb{R}^{N\times N} y \Gamma \in \mathbb{R}^{N\times K}$

1: Estimar las variables proxy $\hat{Y} = X(X^TX)^{-1}X^TY$
2: Sustituir Y por $\hat{Y}$ resultando $Y^T = \tilde{B}\hat{Y}^T + \Gamma X^T + u^T$
3: **Para** i=1...N **Hacer**
4:     **Si** la ecuación i-ésima está identificada **Entonces**
5:         $Z_i = [X|\hat{Y}]S_i$
6:         $\hat{\eta}_i = (Z_i^TZ_i)^{-1}Z_i^Ty_i$
7:     **Fin si**
8: **Fin Para**

MC2E mediante reflexiones de Householder

- Hacer la descomposición QR de X ($X = Q \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$ mediante reflexiones de Householder).
- Calcular $\hat{Y}$

$$\hat{Y} = X(X^TX)^{-1}X^TY = QR(R^TQ^TQR)^{-1}R^TQ^TY =$$
$$Q \begin{pmatrix} R_1 \\ 0 \end{pmatrix} (R_1^{-1}R_1^{-T})[R_1^T|0]Q^TY = Q \begin{pmatrix} Id_K \\ 0 \end{pmatrix} [Id_K|0]Q^TY$$
$$= Q \begin{pmatrix} Id_K & 0 \\ 0 & 0 \end{pmatrix} Q^TY = Q \begin{pmatrix} Id_K & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \end{pmatrix} = Q \begin{pmatrix} \tilde{Y}_1 \\ 0 \end{pmatrix}$$

- En cada ecuación i:
  - Hacer la descompsición QR de la matriz $Z_i = Q_iR_i$ mediante reflexiones de Householder.
  - Calcular $\hat{\eta}_i = (Z_i^TZ_i)^{-1}Z_i^Ty_i = [R_{i,1}^{-1}|0]\tilde{y}_i = R_{i,1}^{-1}\tilde{y}_{i,1}$

Algoritmo QR para MC2E

Algoritmo 2 Algoritmo $MC2E_{QR}$

Entrada: $X \in \mathbb{R}^{d \times K}$, $Y \in \mathbb{R}^{d \times N}$ y $S_i \in \mathbb{R}^{(K+N)\times(k_i+n_i-1)}$ $\forall i = 1, ..., N$

Salida: $B \in \mathbb{R}^{N \times N}$ y $\Gamma \in \mathbb{R}^{N \times K}$

1: Utilizando reflexiones de Householder obtener Q, R e $\tilde{Y}$ tal que $X = QR$ (QRD de X) e $\tilde{Y} = Q^T Y$ $\{$coste $\rightarrow \frac{4}{3}K^2(3d - K) + 2NK(2d - K)\}$

2: $\hat{Y} = Q \begin{pmatrix} \tilde{Y}_1 \\ 0 \end{pmatrix}$ $\{$coste $\rightarrow 2NK^2\}$

3: **Para** i=1...N **Hacer**

4: **Si** la ecuación i-ésima está identificada **Entonces**

5: $Z_i = [X|\hat{Y}]S_i$

6: Utilizando reflexiones de Householder obtener Q_i, R_i e $\tilde{y}_i$ tal que $Z_i = Q_i R_i$ (QRD of Z_i) e $\tilde{y}_i = Q_i^T y_i$ $\{$coste $\rightarrow \frac{4}{3}(n_i + k_i - 1)^2(3d - (n_i + k_i - 1)) + 2(n_i + k_i - 1)(2d - (n_i + k_i - 1))\}$

7: $\hat{\eta}_i = R_{i,1}^{-1}\tilde{y}_{i,1}$ $\{$coste $\rightarrow (n_i + k_i - 1)^2\}$

8: **Fin si**

9: **Fin Para**

MC2E mediante rotaciones de Givens

Se detectan las siguientes deficiencias en el algoritmo anterior:

- No se aprovecha la descomposición QR de X en la descomposición de Z_i en cada ecuación.
- El cálculo de la matriz $\hat{Y}$ exige aplicar dos veces los reflectores de Householder a la matriz Y.

Se propone: Obtener la descomposición QR de cada Z_i mediante rotaciones de Givens haciendo ceros los elementos no nulos por debajo de la diagonal principal.

- Hacer la descomposición QR de $[X|\hat{Y}]$ de la cual se obtendrán las descomposiciones QR de las Z_i.
- Hacer la descomposición QR de X mediante reflexiones de Householder.
- Calcular $Q^T[X|\hat{Y}]=[Q^TX|Q^T\hat{Y}] = [R|Q^T\hat{Y}] = \begin{pmatrix} R_1 & \tilde{Y}_1 \\ 0 & 0 \end{pmatrix}$

MC2E mediante rotaciones de Givens II

Resolver la ecuación i-ésima $y_i = X_i b_i + Y_i \Gamma_i + \epsilon_i$ cuya matriz asociada es $Z_i = [X_i | \hat{Y}_i]$.

La matriz se puede expresar de la forma $Z_i = [X | \hat{Y}] S_i$ donde S_i es una matriz de selección.

$$Q^T Z_i = Q^T [X | \hat{Y}] S_i = [R | Q^T \hat{Y}] S_i = \begin{pmatrix} R_1 & \tilde{Y}_1 \\ 0 & 0 \end{pmatrix} S_i = \begin{pmatrix} \dot{R}_{i,1} & \tilde{Y}_{i,1} \\ 0 & 0 \end{pmatrix},$$

siendo $\dot{R}_{i,1}$ e $\tilde{Y}_{i,1}$ las K primeras filas de la matriz resultante de la multiplicación ($\dot{R}_{i,1} \in \mathbb{R}^{K \times k_i}$ and $\tilde{Y}_{i,1} \in \mathbb{R}^{K \times (n_i - 1)}$).

Es necesario obtener la descomposición QR de Z_i, por lo que hay que encontrar una matriz ortogonal que multiplicada a la anterior dé como resultado una matriz triangular superior.

$$\tilde{Q}_i^T = \left(\prod_{n=1}^{n_i - 1} \prod_{j=1}^{K-n} G^{(i)}_{K-j-k_i, K-j-k_i+1} \right) \left(\prod_{n=1}^{k_i} \prod_{j=1}^{\lambda_{i,n} - n} G^{(n)}_{\lambda_{i,n} - j, \lambda_{i,n} - j + 1} \right)$$

MC2E mediante rotaciones de Givens III

La siguiente figura muestra la secuencia de eliminación de elementos en el proceso de retriangularización mediante rotaciones de Givens de la matriz $\begin{pmatrix} \dot{R}_{i,1} & \tilde{Y}_{i,1} \\ 0 & 0 \end{pmatrix}$ con K=7, N=6, $n_i = 3$, $k_i = 4$ y $\lambda_i = (1, 3, 4, 6, 8, 9, 11)$.

La entrada i (i=1,...,7) indica el elemento reducido a cero en la i-ésima rotación, multiplicando a la izquierda por la matriz $\tilde{Q}_i^T = G_{6,7}^{(6)} G_{5,6}^{(5)} G_{6,7}^{(5)} G_{4,5}^{(4)} G_{5,6}^{(4)} G_{3,4}^{(3)} G_{2,3}^{(2)}$.

$$\left(\begin{array}{cccc|ccc} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & 1 & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & 2 & \bullet & \bullet & \bullet & \bullet \\ & & & 4 & \bullet & \bullet & \bullet \\ & & & 3 & 6 & \bullet & \bullet \\ & & & & 5 & 7 & \bullet \end{array}\right)$$

Algoritmo para MC2E mediante rotaciones de Givens

Algoritmo 3 Algoritmo $MC2E_G$

Entrada: $X \in \mathbb{R}^{d \times K}$, $Y \in \mathbb{R}^{d \times N}$ y $S_i \in \mathbb{R}^{(K+N)\times(k_i+n_i-1)}$ $\forall i = 1, ..., N$
Salida: $B \in \mathbb{R}^{N \times N}$ y $\Gamma \in \mathbb{R}^{N \times K}$

1: Utilizando reflexiones de Householder obtener Q, R e $\tilde{Y}$ tal que $X = QR$ (QRD de X) e $\tilde{Y} = Q^T Y$ $\{$coste $\rightarrow \frac{4}{3}K^2(3d-K) + 2NK(2d-K)\}$
2: $Z = [R_1|\tilde{Y}_1]$
3: **Para** i=1...N **Hacer**
4: **Si** la ecuación i-ésima está identificada **Entonces**
5: $[\dot{R}_{i,1}|\tilde{Y}_{i,1}] = ZS_i$
6: Utilizando reflexiones de Householder obtener $\tilde{Q}_i$, $R_{i,1}$ e $\tilde{y}_i$ tal que $[\dot{R}_{i,1}|\tilde{Y}_{i,1}] = \tilde{Q}_i^T R_{i,1}$ e $\tilde{y}_i = \tilde{Q}_i^T \tilde{Y}_i$ $\{$coste $C_i(\lambda_i, k_i, n_i) + 6\left(\sum_{j=1}^{k_i}(\lambda_{i,j} - j) + \sum_{j=1}^{n_i-1}(K - j - k_i)\right)\}$
7: $\hat{\eta}_i = R_{i,1}^{-1}\tilde{y}_{i,1}$ $\{$coste $\rightarrow (n_i + k_i - 1)^2\}$
8: **Fin si**
9: **Fin Para**

Objetivos del estudio experimental

Los experimentos realizados tienen dos objetivos:

- Comparar los algoritmos $MC2E_G$ y $MC2E_{QR}$ para diferentes tamaños del problema con el fin de determinar el más eficiente en cada caso y comparar los costes obtenidos experimentalmente con aquellos obtenidos previamente mediante el uso de los costes teóricos.
- Estudiar la paralelización de los algoritmos con diferente número de cores y diferente tamaños del problema con el fin de evaluar el speedup obtenido con respecto a la versión secuencial.

Estudio experimental

Tiempo de ejecución (en segundos), ratio de $MC2E_{QR}$ y $MC2E_G$ y tiempos teóricos estimados de $MC2E_{QR}$ y $MC2E_G$. $N = 400$. variando K y d.

Tam. problema		Tiempo (segs)			tiempo teórico (segs)	
K	d	$MC2E_{QR}$	$MC2E_G$	Ratio	$T_{MC2E_{QR}}$	T_{MC2E_G}
400	1000	46.40	23.75	1.95	38.47	24.01
	1500	79.27	24.06	3.29	57.70	24.27
	2000	101.14	24.17	4.18	76.93	24.53
	2500	135.24	24.38	5.55	96.17	24.79
	3000	159.31	24.52	6.50	115.40	25.05
800	1000	126.07	122.12	1.03	126.39	156.00
	1500	209.78	121.32	1.73	189.59	156.78
	2000	362.98	122.01	2.98	252.78	157.57
	2500	527.71	123.05	4.29	315.98	158.35
	3000	731.35	123.50	5.92	379.17	159.13

Estudio experimental

Tiempo de ejecución (en segundos), ratio de $MC2E_{QR}$ y $MC2E_G$ y tiempos teóricos estimados de $MC2E_{QR}$ y $MC2E_G$. $d = 2000$. variando N y K.

Tam. problema		Tiempo (segs)		
N	K	$MC2E_{QR}$	$MC2E_G$	Ratio
400	400	110.49	24.02	4.60
	600	219.16	65.28	3.36
	800	320.77	122.12	2.63
	1000	561.06	208.58	2.69
600	400	154.61	43.69	3.54
	600	332.69	121.15	2.75
	800	571.86	239.64	2.39
	1000	942.08	394.48	2.39
800	400	224.55	66.83	3.36
	600	469.06	185.74	2.53
	800	897.46	374.66	2.40
	1000	1439.45	632.23	2.28
1000	400	292.22	88.13	3.32
	600	581.92	258.95	2.25

Estudio experimental

Tiempo de ejecución (en segundos), ratio de $MC2E_{QR}$ y $MC2E_G$ y tiempos teóricos estimados de $MC2E_{QR}$ y $MC2E_G$. Variando N, K y d.

Problem size			Time (secs)			Theoretical Time (secs)	
N	K	d	$MC2E_{QR}$	$MC2E_G$	Ratio	$T_{MC2E_{QR}}$	T_{MC2E_G}
$N = K = d$							
400	400	400	15.68	24.62	0.64	15.39	23.69
600	600	600	79.10	121.69	0.65	79.39	122.25
800	800	800	256.24	373.81	0.69	246.74	379.91
$N \leq K = d$							
200	400	400	6.30	7.59	0.83	6.42	9.88
200	600	600	19.18	20.11	0.95	19.24	29.63
200	800	800	44.42	41.66	1.07	40.87	62.93
$N \geq K = d$							
1000	400	400	44.92	87.57	0.51	38.37	59.07
1000	600	600	110.38	210.11	0.53	132.75	204.40
1000	800	800	271.23	509.16	0.53	305.23	469.98

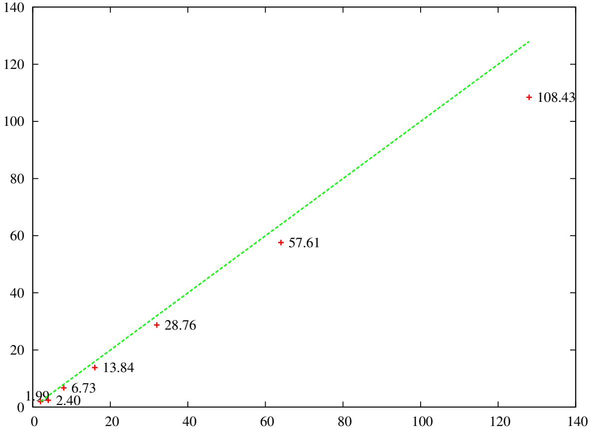
Estudio experimental

Tiempo de ejecución (en segundos) de $PMC2E_G$ y speedup. Variando N, K y d.

N :	400	400	400	400	600	600	600	600
K :	400	400	600	600	400	400	600	600
d :	1000	2000	1000	2000	1000	2000	1000	2000
p	Time (secs)							
1	23.71	24.03	64.01	64.96	43.19	43.51	120.9	121.39
4	6.18	6.29	16.75	16.77	11.28	11.34	31.04	31.22
16	1.70	1.79	4.60	4.68	3.01	3.09	8.87	9.11
32	1.10	1.12	2.51	2.63	1.75	1.85	4.60	4.80
64	0.65	0.73	1.45	1.59	1.06	1.13	2.64	2.76
128	0.85	0.86	1.25	1.44	1.05	1.22	1.74	1.83
Speedup								
4	3.84	3.82	3.82	3.87	3.83	3.84	3.89	3.89
16	13.95	13.42	13.92	13.88	14.35	14.08	13.63	13.32
32	21.55	21.46	25.50	24.70	24.68	23.52	26.28	25.29
64	36.48	32.92	44.14	40.86	40.75	38.50	45.80	43.98
128	27.89	27.94	51.21	45.11	41.13	35.66	69.48	66.33

Estudio experimental

Escalabilidad-speedup del algoritmo paralelo $PMC2E_G$



Conclusiones

- Se han desarrollado algoritmos para el estimador MC2E utilizando la descomposición matriciales QR.
- El algoritmo $MC2E_G$ que utiliza rotaciones de Givens mejora a $MC2E_H$ basado únicamente en reflexiones de Householder cuando el tamaño de d es grande. Cuando d se aproxima a K ocurre lo contrario.
- Se han desarrollado versiones en paralelo en memoria compartida obteniéndose speed-ups satisfactorios.