Danmarks Statistik
MODELGRUPPEN
Edith Madsen

Arbejdspapir*

18. april 1998

Estimation af systemer med singulær kovarians

Resumé:

Har man et singulært system, og ønsker man at estimere parametrene i systemet, er spørgsmålet, hvordan gør man? Det man oftest gør er, at udelade en relation i systemet og så estimere i de resterende relationer. Dette resultat blev vist af Barten i 1969. Det kan da også som regel bruges, men som alt andet har det sine begrænsninger. I en artikel af Dhrymes fra 1994 gives en anden metode, som på mange måder er mere generel, og denne beskrives kort i papiret.

Har man det dårligt med lineær algebra og likelihoodfunktioner, kan man nøjes med at læse indledningen.

Filnavn: edm18498.tex

Nøgleord: Ikke-lineær, systemestimation, singulær, Barten

1 Indledning

Er man i den situation at skulle estimere et system af budgetandele har man et singulært system, da budgetandelene jo summerer til 1. Den klassiske metode til estimation af et sådant system er, at man udelader en relation og så estimrer det regulære system bestående af de resterende relationer. I Bartens artiklen fra 1969 vises, hvorfor man kan gøre sådan. Det vises nemlig, at parameterestimaterne er ufhængigige af, hvilken relation der udelades. I Bartens artikel er dette resultat netop udledt på baggrund af en estimation af budgetandele, men resultatet kan bruges i andre sammenhænge. I det følgende gennemgås først beviset for ovenstående resultat som i Barten (1969), blandt andet for at gøre det klart, hvilke forudsætninger det faktisk bygger på. Derefter lidt om anvendelserne af det, herunder også hvornår man ikke kan anvende det. Til sidst et afsnit om forslag til en anden fremgangsmåde.

2 Bartens resultat

Lad $\varepsilon_1, ..., \varepsilon_T$ være en stokastisk proces hvor $\varepsilon_t \sim iidN_n(0, \Omega)$. Antag at $v'\varepsilon_t = \varepsilon_{1t} + ... + \varepsilon_{nt} = 0$ for alle $t = 1, ..., T$, hvor v er en n-vektor af et-taller. Fordelingen af ε_t ´erne er da singulær. Lad os nu udelade den sidste komponent af ε_t og derved få en stokastisk proces $\tilde{\varepsilon}_1, ..., \tilde{\varepsilon}_T$ med en regulær fordeling, nemlig $\tilde{\varepsilon}_t \sim iidN_{n-1}(0, \tilde{\Omega})$, hvor $\tilde{\Omega}$ er delmatricen af Ω bestående af de $n-1$ første rækker og søjler i Ω. $\tilde{\Omega}$ har fuld rang. Likelihoodfunktionen for fordelingen af $\tilde{\varepsilon}_1, ..., \tilde{\varepsilon}_T$ er:

$$L\left(\tilde{\varepsilon}_1, ..., \tilde{\varepsilon}_T, \tilde{\Omega}\right) = (2\pi)^{-\frac{1}{2}T(n-1)} \left|\tilde{\Omega}\right|^{-\frac{1}{2}T} \exp\left(-\frac{1}{2}\sum_{t=1}^{T} \tilde{\varepsilon}_t' \tilde{\Omega}^{-1} \tilde{\varepsilon}_t\right) \tag{1}$$

Det kan nu vises, at udtrykket på højesiden i (1) ikke afhænger af, hvilken komponent der er udeladt. Beviset bygger udelukkende på antagelsen $v'\varepsilon_t = 0$, og at der dermed gælder: $\Omega v = E\left(\varepsilon_t \varepsilon_t'\right) v = E\left(\varepsilon_t \varepsilon_t' v\right) = 0$. Altså at v er en egenvektor hørende til egenværdien 0.

Definer en $n \times n$ matrix E_n af fuld rang ved:

$$\begin{aligned} E_n &= I_n - i_n v' - v i_n' \\ &= \begin{bmatrix} & & & -1 \\ & I_{n-1} & & \vdots \\ & & & -1 \\ -1 & \cdots & -1 & -1 \end{bmatrix} \end{aligned}$$

hvor i_n er den n´te søjle i enhedsmatricen I_n. Ved udvikling efter sidste søjle ses, at $|E_n| = -n$. Ved at benytte $\Omega v = 0$ fås:

$$E_n \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} E_n = \Omega + \frac{1}{n} v v' \equiv \Sigma \tag{2}$$

Matricen på højre side i (2) har egenværdien 1 hørende til egenvektoren v, idet $\Sigma v = \left(\Omega + \frac{1}{n} v v'\right) v = v$. Σ har fuld rang, og determinanten bliver:

$$|\Sigma| = \left| \Omega + \frac{1}{n} v v' \right| = \left| E_n \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} E_n \right| = |E_n|^2 \left| \tilde{\Omega} \right| \frac{1}{n} = n \left| \tilde{\Omega} \right| \tag{3}$$

Desuden gælder jvf. konstruktionen af $\tilde{\varepsilon}_t$og antagelsen $v' \varepsilon_t = 0$:

$$E_n \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix} = (I_n - i_n v' - v i_n') \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ -\varepsilon_{nt} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \varepsilon_t$$

Da E_n har fuld rang gælder:

$$\begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix} = E_n^{-1} \varepsilon_t$$

Dermed kan vi skrive:

$$\begin{aligned}
\tilde{\varepsilon}_t' \tilde{\Omega}^{-1} \tilde{\varepsilon}_t &= \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix}' \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \end{bmatrix} \\
&= \varepsilon_t' E_n^{-1} \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix}^{-1} E_n^{-1} \varepsilon_t \\
&= \varepsilon_t' \left(\begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} E_n \right)^{-1} E_n^{-1} \varepsilon_t \\
&= \varepsilon_t' \left(E_n \begin{bmatrix} \tilde{\Omega} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} E_n \right)^{-1} \varepsilon_t \\
&= \varepsilon_t' \left(\Omega + \frac{1}{n} v v' \right)^{-1} \varepsilon_t
\end{aligned}$$

Højesiden i (1) kan nu skrives som:

$$\begin{aligned}
& n^{\frac{1}{2}T} (2\pi)^{-\frac{1}{2}T(n-1)} \left| \Omega + \frac{1}{n} v v' \right|^{-\frac{1}{2}T} \exp \left(-\frac{1}{2} \sum_{t=1}^{T} \varepsilon_t' \left(\Omega + \frac{1}{n} v v' \right)^{-1} \varepsilon_t \right) \\
\equiv \;& h(\varepsilon_1, ..., \varepsilon_T, \Omega)
\end{aligned}$$

og det ønskede er hermed vist.

Estimatoren for $\tilde{\Omega}$ er som sædvanlig givet ved :

$$\tilde{\Omega} = \frac{1}{T} \sum_{t=1}^{T} \tilde{\varepsilon}_t \tilde{\varepsilon}_t'$$

Dermed bliver den koncentrerede log-likelihoodfunktion:

$$\log L\left(\tilde{\varepsilon}_1,...,\tilde{\varepsilon}_T,\tilde{\Omega}\right) = -\frac{1}{2}T\left(n-1\right)\log\left(2\pi\right) - \frac{1}{2}T\log\left|\tilde{\Omega}\right|^{-}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}\tilde{\varepsilon}_t'\tilde{\Omega}^{-1}\tilde{\varepsilon}_t\right)$$

3 Estimation ved udeladelse af en relation

Antag nu at vi har to processer $Y_t(n\times 1)$ og $X_t\,(p\times 1)$ for $t=1,...,T$, hvorom der gælder:

$$Y_t = f\left(X_t,\beta\right)+\varepsilon_t \tag{4}$$

hvor $\varepsilon_t \sim N_n\left(0,\Omega\right)$ og $f\left(X_t,\cdot\right)$ er en C^2-funktion. Antag også at $v\varepsilon_t=0$, så vi er i situationen fra foregående afsnit. Hvis β i (4) er identificeret, kan Bartens metode benyttes, dvs. maksimering af den partielle likelihoodfunktion baseret på $n-1$ af relationerne i (4). Dette er som tidligere nævnt situationen ved estimation af et budgetandelssystem. Metoden blev f.eks. anvendt ved estimation af DLU, det dynamiske lineære udgiftssytem, se modelgruppepapir EDM 4. februar 1997.

Er funktionen f i (4) lineær bliver modellen:

$$Y_t = \beta' X_t + \varepsilon_t \tag{5}$$

og maksimering af den partielle likelihoodfunktion baseret på $n-1$ af relationerne er givet ved OLS. Bemærk at i modellen i (5) er højresidevariablene de samme i alle relationer. Er højresidevariablene ikke de samme i alle relationer, kan systemet ikke opskrives på formen (5) . I dette tilfælde er OLS ikke det samme som maksimum-likelihood, og estimerer man systemet bestående af $n-1$ af relationerne ved OLS, kan man ikke være sikker på, at parameteresimaterne er uafhængige af, hvilken relation der udelades. Resutatet fra foregående afsnit holder ikke.

I Dhrymes (1994) behandles problemet med singulære systemer, med mulighed for at have forskellige højresidevariabler.

4 Andre metoder

I Dhrymes (1994) gives som sagt en anden metode til at estimere et singulært system, når strukturen er som antaget i tidligere afsnit, dvs. summen af de normalfordelte restled er 0. Metoden går ud på at transformere de singulært normalfordelte restled til normerede uafhængige normalfordelte variable. I artiklen vises det, at dette giver fornuftige estimatorer, når modellen er lineær i parametrene. Desværre har udgifssystemer sjældent denne egenskab, og jeg ved ikke om noget tilsvarende fører til noget fornuftigt i sådanne ikke-lineære systemer. Men her en gennemgang af, hvad jeg havde tænkt mig.

Lad $\varepsilon_t \sim iidN_n(0, \Omega)$ og antag som før $\Omega v = 0$, dvs. Ω har rang $n-1$. Da Ω er positiv semidefinit findes en $n \times n$ matrix Q bestående af n egenvektorer for Ω, og en $n \times n$ diagonalmatrix Λ med de tilhørende egenværdier så $\Omega Q = Q\Lambda$. Da vi ved, at v er en egenvektor hørende til egenværdien 0 gælder:

$$\begin{aligned} Q &= \begin{bmatrix} Q_1 & v \end{bmatrix} \\ \Lambda &= \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

hvor Q_1 er en $n \times (n-1)$ matrix af fuld rang $(n-1)$, som kan vælges ortonormalt, dvs. med $Q_1' Q_1 = I_{n-1}$, og Λ_1 er en $(n-1) \times (n-1)$ diagonalmatrix af egenværdierne forskellige fra 0. Dermed gælder:

$$Q_1' \Omega Q_1 = Q_1' Q_1 \Lambda_1 = I_{n-1} \Lambda_1 = \Lambda_1 \tag{6}$$

Sæt $C = \Lambda_1^{-\frac{1}{2}} Q_1'$ med $u_t = C\varepsilon_t$ fås ved at anvende (6):

$$cov\,(u_t) = cov\,(C\varepsilon_t) = C\Omega C' = \Lambda_1^{-\frac{1}{2}} Q_1' \Omega Q_1 \Lambda_1^{-\frac{1}{2}} = \Lambda_1^{-\frac{1}{2}} \Lambda_1 \Lambda_1^{-\frac{1}{2}} = I_{n-1}$$

Dette svarer til at vælge enhedsvektorer som basisvektorer i underrummet $\mathbb{R}^{n-1}$, som er det underrum ε_t´erne faktisk udspænder.

Vi kan nu opskrive en likelihoodfunktion for fordelingen af $u_1, ..., u_T$, som jo er $u_t \sim iddN(0, I_{n-1})$, og den kommer kun til at afhænge af u_t´erne gennem følgende udtryk:

$$\begin{aligned} \sum_{t=1}^{T} u_t' u_t &= \sum_{t=1}^{T} \varepsilon_t' C' C \varepsilon_t = \sum_{t=1}^{T} \varepsilon_t' Q_1 \Lambda_1^{-\frac{1}{2}} \Lambda_1^{-\frac{1}{2}} Q_1' \varepsilon_t \\ &= \sum_{t=1}^{T} \varepsilon_t' Q_1 \Lambda_1^{-1} Q_1' \varepsilon_t \end{aligned} \tag{7}$$

Antag nu som før at vi har to processer $Y_t (n \times 1)$ og $X_t\,(p \times 1)$ for $t = 1, ..., T$, hvorom der gælder:

$$Y_t = f\,(X_t, \beta) + \varepsilon_t \tag{8}$$

hvor $\varepsilon_t \sim N_n\,(0, \Omega)$ og $f\,(X_t, \cdot)$ er en C^2 funktion og $v' \varepsilon_t = 0$. Desuden gives,der mulighed for at have lineære restriktioner på β, givet ved $A\beta = R$, hvor A og R er passende matricer. I et budgetandelssystem vil en af restriktionerne ofte være, at konstanterne i de n relationer summerer til 1.

Følgende algoritme kan muligvis bruges, når modellen i (8) estimeres:

1) Vælg startværdier $\hat{\beta}$ for β.

2) Udregn $\hat{\varepsilon}_t = Y_t - f\left(X_{t,}\hat{\beta}\right)$

3) Sæt $\hat{\Omega} = \sum_{t=1}^{T} \hat{\varepsilon}_t' \hat{\varepsilon}_t$, og løs egenværdiproblemet, dvs. find tilhørende $\hat{Q}_1$ og $\hat{\Lambda}_1$.

4) Minimer $\sum \varepsilon_t \hat{Q}_1 \hat{\Lambda}_1^{-1} \hat{Q}_1' \varepsilon_t'$ m.h.t. β u. b. $A\beta = R$.

5) Sæt $\hat{\beta}$ lig det fundne β i 4) og gå tilbage til 2).

I artiklen vises, at dette kan lade sig gøre i tilfældet, hvor f er en lineær funktioner. Spørgsmålet er, hvad der sker, når f er ikke-lineær.

Litteraturliste

Barten, A.P. (1969), "Maximum likelihood estimation of a complete system of demand equations", *European Economic Rewiew*, Vol.1, pp.7-73

Dhrymes, P.J. (1994), "Autoregressive errors in singular systems of equations", *Econometric Theory,* Vol. 10 pp.254-285