MODELOS MULTIVARIANTES NO ESTACIONARIOS

MODELOS MULTIVARIANTES NO ESTACIONARIOS

CLASE 5

1. TV-FAVAR
2. VAR Bayesiano

Time-varying FAVAR (TV-FAVAR)

- *Modelos FAVAR*: explotan la riqueza informacional de una gran BDD y manejan el problema de variables omitidas generalmente encontrado en modelos VAR estándar. (Bernanke et al. (2005))

- *Modelos pequeños con parámetros que varían en el tiempo y también varianzas que evolucionan en el tiempo*: explícitamente consideran los cambios en las fuentes y en el tamaño de los shocks y en su trasmisión en la economía. (Cogley y Sargent (2005), Sims y Zha (2006)).

- *TV-FAVAR*: Al unir los dos tipos de modelos se logran los beneficios de usar muchas variables y permitir un modelo con una estructura variante en el tiempo. (Baumeiter el al. (2010), Korobilis (2009), Del Negro y Otrok (2008), Liu y Mumtaz (2009), Mumtaz y Surico (2012); usan procedimientos bayesianos).

TV-FAVAR

- Eickmeier, Lemke, Marcellino (2011). “Classical time-varying FAVAR models – Estimation, forecasting and structural analysis”, CEPR.

- Modelo bastante flexible, que permite cambios suaves en: (a) los “loadings” de los factores; (b) los coeficientes autorregresivos del FAVAR; (c) en las relaciones contemporáneas entre los factores; (d) en la volatilidad de los shocks comunes.

- Datos: una gran base de datos para EEUU, más de 300 variables trimestrales macroeconómicas y financieras, observadas entre 1972 y 2007.

TV-FAVAR

- Estiman un TV-FAVAR en dos etapas.
- ***Primera etapa***: estiman los factores por componentes principales (PC).
- ***Segunda etapa***: estiman los "loadings" y los coeficientes variantes en el tiempo, las matrices autorregresivas del VAR de los factores y las varianzas y correlaciones que varían en el tiempo. Los factores estimados en la etapa anterior, se toman como dados:
  1. Las relaciones entre las variables observables y los factores se representan como un conjunto modelos de regresión univariante con parámetros que varían en el tiempo, que evolucionan como random-walks independientes.
  2. El modelo se estima convirtiendo cada ecuación en la forma de estados-espacios, estimando los hiperparámetros por máxima verosimilitud y aplicando filtro de Kalman para asegurar los senderos de los parámetros que varían en el tiempo. (Nyglom (1989)).

TV-FAVAR

- *Identificación*: matriz triangular inferior de relaciones contemporáneas, que deja a las ecuaciones del VAR condicionalmente independientes.

- *Especificación de la volatilidad*: es función (exponencialmente afín) de factores rezagados. Otros modelos la vinculan a un factor latente adicional.

- *Resultados*:
    1. Variación temporal sustancial en la varianza de los shocks y en sus mecanismos de transmisión
    2. La variación temporal es "sparce": cambios en unos pocos parámetros gobiernan la evolución del sistema y muchos de los parámetros son esencialmente constantes a lo largo del tiempo.

TV-FAVAR

- **Modelo**
- Sea $X_t{}' = (x_{1t}, \ldots, x_{Nt})$ un gran vector de N variables estacionarias de media cero, con t=1, …, T. Tanto N como T pueden tender a infinito.

- Cada elemento de X_t se asume que es la suma de una combinación lineal de G factores comunes $F_t{}' = (f_{1t}, f_{2t}, \ldots, f_{Gt})$ y un componente idiosincrático e_t

$$x_{it} = \Lambda_i{}' F_t + e_{it} \qquad i = 1, \ldots, N \qquad (1)$$

con $e_t = (e_{1t}, \ldots, e_{Nt})$.

- Se asume que los factores son otrogonales y están incorrelacionados con los errores idiosincráticos, que $E(e_t) = 0, E(e_t e_t{}') = R$, R es una matriz diagonal.

TV-FAVAR

- La dinámica de los factores se modela como un VAR(p):

$$F_t = B_1 F_{t-1} + \cdots + B_p F_{t-p} + w_t, \quad E(w_t) = 0, E(w_t w_t{}') = W \quad (2)$$

- (2) se puede interpretar como la forma reducida de un sistema del tipo:

$$PF_t = \mathcal{K}_1 F_{t-1} + \cdots + \mathcal{K}_p F_{t-p} + u_t, \quad E(u_t) = 0, E(u_t u_t{}') = S \quad (2)$$

donde P es triangular inferior con unos en la diagonal principal, S es una matriz diagonal.

- La relación de los parámetros de la forma reducida en (2) es $B_i = P^{-1}\mathcal{K}_i$ y $W = P^{-1}SP^{-1\prime}$.

- Se logra G ecuaciones condicionalmente independientes.

TV-FAVAR

- Si se relaja el supuesto de constancia en el tiempo:

$$x_{it}= \Lambda_{it}{}' F_t + e_{it} \qquad i = 1, \dots, N \qquad (3)$$

$$P_t F_t = \mathcal{K}_{1t} F_{t-1} + \cdots + \mathcal{K}_{pt} F_{t-p} + u_t, \quad E(u_t) = 0, E(u_t u_t{}')= S_t \quad (4)$$

donde P_t es triangular inferior y S_t es diagonal.

- Los componentes idiosincráticos siguen un proceso autorregresivo de primer orden:

$$e_{it} = \rho_i e_{it-1} + \xi_{it}\,, \qquad E(\xi_{it}) = 0, E({\xi_{it}}^2)= {\sigma_i}^2, \quad i = 1, \dots, N$$

Los elementos de ξ_t se suponen incorrelacionados contemporáneamente.

- Sea a_t el vector de los parámetros que varían en el tiempo $\{P_t, \mathcal{K}_{1t}, \dots, \mathcal{K}_{pt}, \Lambda_{1t}, \dots, \Lambda_{Nt}\}$, entonces se asume que varían lentamente, como random walks independientes: $a_t = a_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, Q)$

Q matriz diagonal y $(\xi_i, u_i, \varepsilon_t)$ están incorrelacionados contemporáneamente en el tiempo.

TV-FAVAR

- Molteni, Pappa. “Monetary Policy in Times of Fiscal Turbulence: A TVP-FAVAR Approach”.

- Los autores analizan cómo la respuesta de variables macroeconómicas a la política monetaria en la economía de EEUU depende de la política fiscal, usando un modelo FAVAR con coeficientes variables y volatilidad estocástica.

- La estructura variante en el tiempo del modelo, les permite determinar si los shocks fiscales expansivos o contractivos, identificados con un enfoque narrativo, afectan la trasmisión de los shocks de política monetaria.

TV-FAVAR

- Definen episodios de política fiscal expansiva o contractiva en base a variaciones exógenas grandes en los instrumentos fiscales (impuestos, gasto del gobierno y transferencias), tomando ventaja de la información de la evidencia narrativa desarrollada “fuera” del VAR. En esos períodos, determinan la función de impulso respuesta de un shock de política monetaria usando un TVP-FAVAR

- Concluyen que: (a) Cuando los shocks de política monetaria coinciden con épocas de expansión en los gastos del gobierno y en las transferencias temporarias, los efectos son más débiles.(b) Incrementos en las transferencias permanentes modifican menos los efectos sobre la demanda de los shocks de política monetaria. (c) Los shocks impositivos no afectan la propagación de los shocks de política monetaria.

VAR Bayesiano

- Debido a que los VARs frecuentemente requieren que se estime un gran número de parámetros, la sobreparametrización es un problema común, quedando demasiado pocas observaciones para estimar los parámetros del modelo.

- Una posible solución consiste en la reducción del set de parámetros a través de la imposición de restricciones a los mismos.

VAR Bayesiano (BVAR, Bayesian VAR en inglés)
Litterman (1986), Doan, Litterman, and Sims (1984), Sims and Zha (1998).

Las “priors” Bayesianas dan un método lógico y consistente para imponer restricciones a los parámetros.

VAR Bayesiano – conceptos previos

- El análisis Bayesiano requiere conocer las propiedades distribucionales de la “prior”, “likelihood” y “posterior”.

- En econometría Bayesiana, todo aquello que es incierto incluso el valor verdadero de un parámetro, puede pensarse como una variable aleatoria a la cual se le puede asignar una función de probabilidad.

- “**prior**” = información externa basada en la creencia de los investigadores sobre los parámetros de interés.

- “**likelihood**” = información de datos contenida en la función de distribución de probabilidad muestral.
- La combinación de la distribución prior a través del teorema de Bayes con la función de verosimilitud de los datos, resulta en la distribución “**posterior**”.

VAR Bayesiano – conceptos previos

- La econometría bayesiana realiza todas sus inferencias en base **al Teorema de Bayes**, que resulta de la probabilidad conjunta de dos eventos, A y B.

- Dados $\mathrm{P(A, B) = P(A/B)P(B)}$ y $\mathrm{P(B, A) = P(B/A)P(A)}$

  Igualando, resulta que: $\mathrm{P(A/B)} = \dfrac{\mathrm{P(B/A)P(A)}}{\mathrm{P(B)}}$

- Si y es un vector de datos y θ un vector de parámetros de un modelo que intenta explicar y, aplicando el Teorema de Bayes se obtiene: $\mathrm{P}(\theta/\mathrm{y}) = \dfrac{\mathrm{P}(\mathrm{y}/\theta)\mathrm{P}(\theta)}{\mathrm{P}(\mathrm{y})}$

- El investigador está interesado fundamentalmente en utilizar los datos para aprender sobre θ , por lo que el término P(y) se puede obviar: $\mathrm{P}(\theta/\mathrm{y}) \propto P(y/\theta)P(\theta)$

VAR Bayesiano – conceptos previos

Llamemos:

$\boldsymbol{\theta} = (\boldsymbol{\alpha}, \Sigma)$ a los parámetros de interés en un determinado modelo

y: los datos

$\boldsymbol{\pi}(\boldsymbol{\theta})$: la distribución prior

$\mathbf{l}(\mathbf{y}/\boldsymbol{\theta})$**:** la likelihood o función de verosimilitud

$\boldsymbol{\pi}(\boldsymbol{\theta}/\mathbf{y})$: es la distribución posterior de θ dada los datos y:

$$\pi(\theta/y) = \frac{\pi(\theta) l(y/\theta)}{\int \pi(\theta) l(y/\theta)\, d\theta}$$

- Como el denominador es una constante estandarizada que no tiene aleatoriedad, la posterior es proporcional al producto de la likelihood y la prior:

$$\pi(\theta/y) \propto \pi(\theta) l(y/\theta)$$

- **El objetivo fundamental** es encontrar los momentos de la posterior para los parámetros de interés.

VAR Bayesiano – conceptos previos

- Un tema crucial en la econometría bayesiana es la formulación de la distribución prior de los parámetros, basada en información que refleja las creencias del investigador.
- La información "prior" generalmente se incorpora para reforzar las inferencias acerca de los valores verdaderos de los parámetros.
- *Problema*: una "prior" es intrínsecamente subjetiva y potencialmente puede dar lugar a manipulaciones.
- A medida que T crece, aumenta la importancia relativa de los datos respecto a la prior y el estimador coincide con el estimador de OLS sin restricciones.
- Investigaciones recientes señalan que los BVARs son una herramienta muy útil para tratar grandes bases de datos.

VAR Bayesiano – conceptos previos

- $Y = XA + E$, ó $y = (I_m \otimes X)\alpha + e$, $e \sim (0, \Sigma_e \otimes I_T)$

  donde Y, E son matrices Txm, X es una matriz Txk

  $$X_t = \left[y'_{t-1}, \dots, y'_{t-k}\right]$$

  y, e son vectores mTx1, I_m es la matriz identidad de dimensión m, $\alpha = vec(A)$es un vector de mkx1.

- La likelihood para un VAR(p) se puede descomponer en el producto de:
  - **-** una función de densidad normal para α, condicional en la estimación α_{OLS} y en Σ_e
  - **-** una densidad Wishart para ${\Sigma_e}^{-1}$, condicional en α_{OLS}

- Es decir, con las restricciones apropiadas, podemos derivar analíticamente la distribución "posterior" condicional de los coeficientes del VAR y de la matriz de covarianzas de los shocks de la forma reducida.

VAR Bayesiano - *priors*

- **Litterman/Minessota** prior: α normal y Σ_e fija (es conocida) es decir: $\alpha = \overline{\alpha} + v_\alpha, \quad v_\alpha \sim N(0, \bar{\bar{\Sigma}}_\alpha)$.

  La densidad posterior de α es Normal con media $\overline{\alpha}$ y varianza

$$\check{\Sigma}_\alpha = \left[\bar{\bar{\Sigma}}_\alpha{}^{-1} + \left(\Sigma_e{}^{-1} \otimes X'X\right)\right]^{-1}$$

  $\bar{\Sigma}_\alpha$ se elige en forma arbitraria, $\Sigma_e = \frac{1}{T-1}\sum_{t-1}^{T} e'_{t,OLS} e_{t,OLS}$

  $\overline{\alpha}, \Sigma_\alpha$ son funciones de un pequeño número de **hiperparámetros**:

  ϕ_0: estrechez de la varianza del primer rezago, controla la importancia relativa de la información de la muestra y de la prior. Si es grande, la información a priori es difusa y la distribución a posteriori refleja la información muestral. Si es pequeña, domina la información de la prior.

  ϕ_1: estrechez relativa de otras variables

  ϕ_2: estrechez relativa de las variables exógenas

  ϕ_3: estrechez de la varianza de los rezagos excepto el primero

  $\bar{\Sigma}_\alpha$: se elige diagonal

VAR Bayesiano - *priors*

- En Litterman/Minessota prior, a priori, las m series de tiempo son representadas como random walks (con media igual a cero); a posteriori, cada una puede seguir un proceso más complicado.

- La elección de los hiperparámetros $\varphi=(\varphi_0,\varphi_1,\varphi_2,\varphi_3)$, es importante: si es demasiado laxa, puede haber sobreparametrización, si es demasiado restrictiva, se le impide hablar a los datos.

- φ se puede estimar y luego sustituir en las expresiones para $\bar{\alpha}$ y $\bar{\Sigma}_\alpha$ para obtener la distribución posterior de α, condicional en las estimaciones de φ.

- φ se puede tratar como aleatorio, asumir una distribución prior y computar estimaciones de la posterior de α. (MCMC)

VAR Bayesiano - *priors*

- **Normal- Wishart prior**: α normal y Wishart para Σ_e
  Dos hiperparámetros: μ_1 = coeficiente de la media prior
  λ_1 = covarianza prior

- **Sims-Zha prior**: α normal y no informativa para Σ_e
  Cinco hiperparámetros:

$\mu_5 \geq 0$ controla la presencia de raíces unitarias
$\mu_5 \rightarrow \infty$, hay tantas RU como variables
$\mu_6 \geq 0$ controla las observaciones dummy iniciales
$\mu_6 \rightarrow \infty$, el modelo tiende a ser un modelo con todas las variables estacionarias con medias igual al promedio muestral de las condiciones iniciales ó hay componentes con RU sin deriva
λ_0= controla la rigidez total
λ_1= controla la tasa a la que la varianza de la prior disminuye al aumentar el tamaño de la muestra
λ_2= controla la rigidez de las creencias de la varianza residual

VAR Bayesiano - *priors*

- La *prior* contiene toda la información previa que se tiene sobre los parámetros y que no está en los datos.

- La econometría clásica, por el contrario, no incorpora dicha información y critica dicha utilización desde el punto de vista científico.

- Contra-argumentos bayesianos:

1. Es posible usar *priors* no informativas ("flat")
2. Es preferible utilizar más información de forma que la modelización sea más adecuada y las proyecciones más certeras.

- Para la econometría clásica los datos son variables aleatorias; para la econometría bayesiana los parámetros son variables aleatorias.

VAR Bayesiano - *priors*

- Generalmente las *priors* tienen que ver con motivaciones estadísticas o están basadas “*rules-of-thumb*” que resultan útiles para el pronóstico de series macroeconómicas.

- En ambos casos, la teoría económica no juega ningún papel, excepto quizá en establecer el rango de valores para las distribuciones de las *priors*.

- Un intento para darle un sustento de teoría económica a las *priors* consiste en utilizar *priors* derivadas de modelos DSGE. Así, la información de la *prior* mide la confianza del investigador en que la estructura DSGE ha generado los datos observados.

VAR Bayesiano

- La selección de variables a ser incluidas en un BVAR se realiza considerando que las especificaciones con diferentes variables pueden ser tratadas como diferentes modelos.
- Se elige la especificación con menor error de predicción un paso hacia adelante.
- Asimismo, debido a la importancia de la selección de la *prior*, se recomienda estimar el mismo modelo con diferentes *priors* y con distintos valores de una misma prior, para sensibilizar los resultados.
- La elección final dependerá, nuevamente, de aquella *prior* que dé el modelo con menor error de predicción. Generalmente se utiliza el coeficiente U de Theil: $\sqrt{\frac{(1/n)\sum_i (y_i - \hat{y}_i)^2}{(1/n)\sum_i y_i^2}}$

VAR Bayesiano

Antecedentes de proyección bayesiana

1. Banco de la Reserva Federal de Minneapolis. Litterman (1980).
2. Condiciones económicas en Iowa. Doan et al (1984), Whiteman (1996).
3. BVAR para Venezuela. Barráez et al (2008). Se evalúa el poder predictivo de los BVAR para variables macro, objetivo de política económica: inflación, pib y cantidad de dinero.
4. Banco de la República (Colombia), Ocampo y Rodríguez (2012). Implementan un VAR-X estructural y lo estiman por maximimaverosimilitud y métodos bayesianos.