Modelos dinámicos (capitulos 14 y 15 de SW)

1 Modelos dinámicos uniecuacionales

Consideramos aquí modelos de la siguiente forma

$$C(L)y_t = B(L)x_t + \varepsilon_t \ ; \qquad E(\ \varepsilon_t | y_{t-1}, y_{t-2}, ...,\ x_t, .x_{t-1}, ...) = 0 \qquad (1)$$

donde $C(L) = 1 - \gamma_1 L - \gamma_2 L^2 - ... - \gamma_p L^p$ and $B(L) = \beta_0 - \beta_1 L - \beta_2 L^2 - ... - \beta_r L^r$.

La ecuación (1) también puede ser expresada así

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} + \sum_{i=0}^{r} \beta_i x_{t-i} + \varepsilon_t.$$

Este modelo se llama ARDL(p,r), donde ARDL son las siglas de Autoregressive Distributed Lag model (esto es modelo autorregresivo con retardos distribuidos). En algunos libros de texto (SW por ejemplo) consideran ARDL cuando $\beta_0 = 0$.

Notad que poco cambia si incluimos una constante

$$C(L)(y_t - \mu) = B(L)x_t + \varepsilon_t$$

$$y_t = \alpha + \sum_{i=1}^{p} \gamma_i y_{t-i} + \sum_{i=0}^{r} \beta_i x_{t-i} + \varepsilon_t,$$

con $\alpha = C(1)\mu$. Así que ignoraremos la constante para simplificar el análisis.

2 Casos particulares

- Si **r=0** y $\beta_0 = 0$ el modelo ARDL(p,r) se reduce a un modelo AR(p)

$$C(L)y_t = \varepsilon_t$$

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t.$$

- Si **p=0**, tenemos un modelo de retardos distribuidos, es decir, un modelo DL(r)[1]

$$y_t = B(L)x_t + \varepsilon_t$$

$$y_t = \sum_{i=0}^{r} \beta_i x_{t-i} + \varepsilon_t.$$

- Un modelo que se utiliza mucho en la literatura de series temporales y en aplicaciones es el modelo **ARDL(1,1)**

$$y_t = \alpha + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t.$$

[1]Estudiaremos aquí un modelo con sólo una variable independiente,x_t. Pero el análisis se puede generalizar a un modelo con varias variables independientes, x_{1t},x_{2t},...,x_{mt}.

Muchas veces el modelo ARDL(1,1) se estudia en una IMPORANTE especificación equivalente que tiene la interpretación de modelo de corrección del error, este es

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + (\gamma_1 - 1)(y_{t-1} - \theta x_{t-1}) + \varepsilon_t$$

donde $\theta = -(\ \beta_0 + \beta_1)/(\gamma_1 - 1)$. Notad que el sistema esta en equilibrio cuando $y_{t-1} - \theta x_{t-1} = 0$.

- La representación ARDL(p,r) se puede tranformar en un modelo DL(∞) de la siguiente forma

$$C(L) y_t = B(L) x_t + \varepsilon_t$$

$$\begin{aligned} y_t &= \frac{B(L)}{C(L)} x_t + \frac{1}{C(L)} \varepsilon_t && (2) \\ &= \sum_{i=0}^{\infty} \delta_i x_{t-i} + \eta_t && (3) \\ &= D(L) x_t + \eta_t. \end{aligned}$$

Notad que el error η_t ya no es ruido blanco, ahora esta autocorrelacionado.

3 Estabilidad de un modelo dinámico

Un modelo es estable en los dos casos siguientes

1. Ante una variación puntual en x_t ... y_t retorna a su valor de equilibrio

2. Ante una variación permanente en x_t ... y_t evoluciona hacia su nuevo valor de equilibrio

Se demuestra que para que un modelo dinámico sea estable las raices del polinomio $C(L)$ deben estar fuera del círculo unidad, esto es, deben ser en valor absoluto mayores que la unidad. En concreto,

$$1 - \gamma_1\lambda - \gamma_2\lambda^2 - ... - \lambda_p L^p = 0 \qquad \Rightarrow \qquad |\lambda_1| > 1, |\lambda_2| > 1, ..., |\lambda_p| > 1.$$

Esta condición de estabilidad nos asegura que al pasar del modelo dinámico (1) al modelo (2) la suma de coeficientes del polinomio D(L) es finita, es decir, la suma $\sum_{i=1}^{\infty} \delta_i$ es convergente. Por tanto, el impacto sobre la variable endógena es finito. Pasado un tiempo se retorna al equilibrio o se tiende hacia un nuevo equilibrio.

4 Multiplicadores y Retardos

- **Multiplicadores de Impacto**: También denominado multiplicador contemporáneo, el multiplicador de impacto representa el cambio en y_t de una variación unitaria de x_t en el periodo actual. Este es,

$$m_0 = \frac{\partial y_t}{\partial x_t} = D(0) = \delta_0 = \frac{B(0)}{C(0)} = \beta_0.$$

- **Multiplicador de retardo j**: El multiplicador de retardo j, m_j cuantifica el efecto de una variación unitaria de la variable exógena en el periodo t-j, x_{t-j}, sobre la variable endógena en el periodo t, y_t.

$$m_0 = \frac{\partial y_t}{\partial x_{t-j}} = \delta_j \neq \beta_j.$$

- **Multiplicador acumulado en j periodos**: Como lo calcularias???

- **Multiplicador total**: El multiplicador total, m_T, es la suma de todos los multiplicadores.

$$m_T = \sum_{j=0}^{\infty} m_j = D(1) = \sum_{j=0}^{\infty} \delta_j = \frac{B(1)}{C(1)}.$$

- **Retardo medio**: El retardo medio se define como la media, ponderada por el retardo, de todos los coeficientes del polinomio D(L).

$$rmedio = \frac{\sum_{j=0}^{\infty} j\delta_j}{\sum_{j=0}^{\infty} \delta_j} = \frac{D'(1)}{D(1)} = \frac{B'(1)}{B(1)} - \frac{C'(1)}{C(1)}$$

donde $D'(1) = \frac{dD(L)}{dL}|_{L=1}$. El retardo medio es una medida de información sobre si el impacto de x_t sobre y_t es concetrado o diluido en el tiempo.

- **Retardo mediano**: El retardo mediano se define como el instante en el que se alcanza el 50% del impacto total que se produce en y_t debido a una variación en x_t.

$$rmediano = \min \left\{ q | \frac{\sum_{j=0}^{q} \delta_j}{\sum_{j=0}^{\infty} \delta_j} \geq 0.5 \right\}.$$

4.1 Ejemplo

Considera el modelo ARDL(1,0) siguiente

$$y_t = 0.8y_{t-1} + 3x_t + \varepsilon_t.$$

Los polinomios C(L) y D(L) son

$$
\begin{aligned}
C(L) &= 1 - 0.8L \\
B(L) &= 3.
\end{aligned}
$$

Las raíz característica asociada al polinomio C(L) es mayor que uno, y en consecuencia el modelo es estable:

$$1 - 0.8L \Rightarrow \lambda_1 = 1.25 > 1.$$

El polinomio D(L) se obtiene dividiendo B(L) por C(L)

$$
\begin{aligned}
D(L) &= \frac{B(L)}{C(L)} \\
&= \frac{3}{1 - 0.8L} \\
&= 3(1 + 0.8L + 0.8^2L^2 + ...) \\
&= 3 + 2.4L + 1.92L^2 + 1.536L^3 + ... \\
&= \delta_0 + \delta_1 L + \delta_2 L^2 + \delta_3 L^3 + ...
\end{aligned}
$$

El multiplicador total es

$$m_T = D(1) = \frac{B(1)}{C(1)} = \frac{3}{1 - 0.8} = 15 = \sum_{j=0}^{\infty} \delta_j.$$

El retardo medio

$$rmedio = \frac{B'(1)}{B(1)} - \frac{C'(1)}{C(1)} = \frac{0}{3} - \frac{-0.8}{1-0.8} = 4.$$

Y para calcular el retardo mediano definiremos el multiplicador acumulativo, m_0^q. Este es,

$$m_0^q = \sum_{j=0}^{q} \delta_j.$$

Dividiendo cada m_0^q entre el multiplicador total , m_T, se obtiene para cada q qué porcenataje del effecto total se ha acumulado hasta el momento q. Es decir,

$$\begin{aligned}
\frac{m_0^0}{m_T} &= \frac{3}{15} \\
\frac{m_0^1}{m_T} &= \frac{5.4}{15} \\
\frac{m_0^2}{m_T} &= \frac{7.32}{15} \\
\frac{m_0^3}{m_T} &= \frac{8.856}{15} \\
&\ldots
\end{aligned}$$

Y así el retardo medio

$$rmediano = \min \left\{ q | \frac{\sum_{j=0}^{q} \delta_j}{\sum_{j=0}^{\infty} \delta_j} \geq 0.5 \right\} = 3.$$

5 Causalidad en sentido de Granger

Pensad en como definir causalidad de X a Y via los modelos ARDL(p,r) con $\beta_0 = 0$. Luega pensad en como se puede contrastar dicha causalidad o no-causalidad de Granger.

6 Determinación de "p" y "r" en los modelos ARDL(p, r)

- Criterios de Información: AIC, BIC, ...(vease SW 14)

  Recordad: AIC(p)=ln$[\frac{SSr(p)}{T}]+(p+1)\frac{2}{T}$; BIC(p)=ln$[\frac{SSr(p)}{T}]+(p+1)\frac{\ln(T)}{T}$; ....

- De lo General a lo Particular (via contrastes F). Hay varias formas de llevar a cabo esta estrategia (GP). En todas ellas queremos que el error del modelo no este correlacionado.

7 Modelos Económicos

Existen varios modelos económicos que se pueden representar como modelos de regresión dinámicos. Ejemplos:

(i) modelos de expectativas adaptativas

(ii) modelos de ajuste parcial

(iii) modelos de optimización dinámica

entre otros. Se expondran aqui los casos (i) y (ii).

7.1 Expectativas Adaptativas

Consideramos el modelo

$$y_t = \alpha + \beta x^e_{t+1|t} + \varepsilon_t \tag{4}$$

$$x^e_{t+1|t} = \lambda x^e_{t|t-1} + (1-\lambda)x_t, \tag{5}$$

donde $x^e_{t+1|t}$ es la nueva expectativa formulada, $x^e_{t|t-1}$ es la anterior expectativa, x_t es el valor realizado y $\lambda \in [0,1]$.

La ecuación (5) puede ser reescrita del siguiente modo

$$x^e_{t+1|t} = \lambda L x^e_{t+1|t} + (1-\lambda)x_t$$

$$(1-\lambda L)x^e_{t|t-1} = (1-\lambda)x_t$$

$$\begin{aligned} x^e_{t|t-1} &= \frac{(1-\lambda)}{(1-\lambda L)}x_t \qquad (6) \\ &= (1-\lambda)(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + ...), \end{aligned}$$

dando más peso a los valores más recientes.

Notad que cuando $\lambda = 0$, la corrección es total e inmediata. Y cuando $\lambda = 1$ las expectativas son estáticas.

El modelo inicial puede contrastarse mediante el uso de un modelo econométrico dinámico. Para ello basta reescribir la ecuación (4) sustituyendo en ella la ecuación (6), esto es

$$y_t = \alpha + \beta\frac{(1-\lambda)}{(1-\lambda L)}x_t + \varepsilon_t$$

o lo que es lo mismo

$$(1-\lambda L)y_t = (1-\lambda L)\alpha + \beta(1-\lambda)x_t + (1-\lambda L)\varepsilon_t$$

$$y_t = \alpha^* + \gamma_1^* y_{t-1} + \beta_0^* x_t + \varepsilon_t^*$$

donde $\alpha^* = (1-\lambda)\alpha$, $\gamma_1^* = \lambda$, $\beta_0^* = \beta(1-\lambda)$, and $\varepsilon_t^* = (1-\lambda L)\varepsilon_t$. Por tanto, $y_t \sim ARDL(1,0)$ con $\varepsilon_t \sim MA(1)$.

7.2 Ajuste Parcial

El modelo de ajuste parcial supone que existe una variable objetivo o valor deseado, y_t^*, que depende de x_t, esto es

$$y_t^* = \beta x_t.$$

Además

$$\begin{aligned} y_t &= y_{t-1} + \lambda(y_t^* - y_{t-1}) + \varepsilon_t \\ &= (1-\lambda) y_{t-1} + \lambda y_t^* + \varepsilon_t \\ &= (1-\lambda) y_{t-1} + \lambda \beta x_t + \varepsilon_t \\ &= \gamma_1^* y_{t-1} + \beta_0^* x_t + \varepsilon_t \end{aligned}$$

donde ahora $\gamma_1^* = (1-\lambda)$ y $\beta_0^* = \lambda\beta$. Así resulta un modelo $y_t \sim ARDL(1,0)$ sin correlación serial en ε_t.

8 Estimación de modelos dinámicos

Problemas importantes:

1. Multicolinealidad
2. Excesivo número de parámetros

3. Correlación entre regressores y el error

Los problemas (1) y (2) sugieren que es aconsejable elegir valores pequeños de p y r para un modelo ARDL(p,r). Por ello, el modelo ARDL(1,1) es bastante popular:

$$y_t = \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t.$$

También hemos visto que las hipótesis de expectativas adaptativas resultan en modelos ARDL con pocos parámetros. Ahora a por el problema 3.

8.1 Correlación entre regresores y perturbación

Una hipótesis fundamental en el análisis de regresión es que no hay correlación entre los regresores y la perturbación. Sin embargo, si no se cumple esta condición, la estimación MCO (OLS) resulta en estimadores sesgados e inconsistentes.

Esto ocurre por ejemplo en un modelo ARDL(p,r) con autocorrelación en el error.

$$\begin{aligned} y_t &= \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\ y_{t-1} &= \mu + \gamma_1 y_{t-2} + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \varepsilon_{t-1.} \end{aligned}$$

El remedio más utilizado en este caso es la estimación del modelo a través del método de **variables instrumentales**. La idea de este método es sustituir los retardos de y_t por p variables instrumentales (o p combinaciones de h instrumentos, con $h > p$, si se dispone de más instrumentos que variables a instrumentalizar).

Las variables instrumentales son variables que :

(i) están correlacionadas con las variables a las que sustituyen

(ii) no están correlacionadas con la perturbación

(iii) no forman parte de la regresión

Para el caso en el que se dispone del mismo número de instrumentos que variables a instrumentalizar, si Z es la matriz que contiene en sus columnas tanto las observaciones de las variables explicativas sin problemas de endogeneidad del modelo que no han sido sustituidas (no "instrumentadas") como las observaciones de las variables instrumentales, entonces:

$$\hat{\beta}_{IV} = (Z'X)^{-1}(Z'Y)$$

$$Var(\hat{\beta}_{IV}) = \sigma^2 (Z'X)^{-1}(Z'Z)(Z'X)^{-1}$$

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k} = \frac{(Y - X\hat{\beta}_{IV})'(Y - X\hat{\beta}_{IV})}{n-k},$$

donde n es el número de observaciones y k el número de coeficientes estimados.

Ejemplo: El caso de una única variable explicativa:

$$\hat{\beta}_{ols} = \frac{\sum_t x_t y_t}{\sum_t x_t^2} = \frac{\sum_t x_t(x_t\beta + \varepsilon_t)}{\sum_t x_t^2} = \beta + \frac{\sum_t x_t\varepsilon_t}{\sum_t x_t^2}$$

donde el término $\frac{\sum_t x_t\varepsilon_t}{\sum_t x_t^2}$ no desaparece, ni tan siquiera, asintóticamente. Esto es $\hat{\beta}_{ols}$ es sesgado e inconsistente. En cambio el estimador de variables instrumentales en este caso es

$$\hat{\beta}_{IV} = \frac{\sum_t z_t y_t}{\sum_t z_t x_t} = \frac{\sum_t z_t(x_t\beta + \varepsilon_t)}{\sum_t z_t x_t} = \beta + \frac{\sum_t z_t\varepsilon_t}{\sum_t z_t x_t}$$

donde ahora el término $\frac{\sum_t z_t\varepsilon_t}{\sum_t z_t x_t}$ desaparece asintóticamente. Esto es, el estimador de variables instrumentales es consistente.

8.2 Método de Mínimos Cuadrados en Dos Etapas

El método de mínimos cuadrados en dos etapas, MC2E (o two-stages least squares, 2SLS) ayuda a realizar una estimación con variables instrumentales. En particular, una ventaja de este estimador, MC2E, es que combina de manera eficiente la información de multiples instrumentos para regresiones *sobreidentificadas* (donde hay más instrumentos que regresores iniciales).

- En la primera etapa: Se determinan cuáles son las variables exógenas y cuáles son endógenas, explicables con instrumentos. Y se realiza una regression de cada variable del modelo sobre las exógenas (instrumentos y exógenas incluidas en el modelo) (ols).

- En la segunda etapa: Se hace una regresión de la ecuación original donde se reemplazan todas las variables con los valores estimados en la primera etapa.

La expresión del estimador MC2E es

$$\hat{\beta}_{2SLS} = [X'Z(Z'Z)^{-1}Z'X]^{-1}[X'Z(Z'Z)^{-1}Z'Y]$$

con varianza

$$Var(\hat{\beta}_{2SLS}) = \hat{\sigma}^2[X'Z(Z'Z)^{-1}Z'X]^{-1}.$$