

5 LA PROIEZIONE ED I RESIDUI RICORSIVI

Indice del capitolo

| | | |
|-----|---|----|
| 5.1 | La proiezione dei minimi quadrati..... | 2 |
| | <i>L'errore di proiezione.....</i> | 3 |
| | <i>L'errore quadratico medio di proiezione.....</i> | 3 |
| | <i>Proiezione ex ante ed ex post.....</i> | 4 |
| | <i>Indicatori dell'accuratezza delle proiezioni.....</i> | 8 |
| 5.2 | Intervalli di confidenza per le proiezioni..... | 11 |
| 5.3 | Il caso di variabili esplicative non note nel periodo di proiezione..... | 15 |
| | <i>L'errore quadratico di proiezione nel caso di x_{n+h} non noto.....</i> | 15 |
| | <i>Il teorema di Cebiscev.....</i> | 16 |
| 5.4 | I residui ricorsivi..... | 18 |
| | <i>Due formule per il calcolo ricorsivo.....</i> | 19 |
| | <i>La devianza residua ricorsiva.....</i> | 20 |
| | <i>Minimi quadrati ricorsivi e diagnostica del modello.....</i> | 21 |
| 5.5 | Due test di cambiamento strutturale nel caso $n_2 \leq k$ | 27 |
| | <i>Il test del Chow basato sulle proiezioni.....</i> | 28 |
| | <i>Il test preliminare di uguaglianza delle varianze.....</i> | 30 |
| | <i>Applicazioni alla funzione delle importazioni.....</i> | 30 |
| 5.6 | Verifica della stabilità dei parametri..... | 33 |
| | <i>I test CUSUM e CUSUMSQ.....</i> | 35 |
| 5.7 | L'impiego delle variabili di comodo nella proiezione..... | 38 |
| | <i>L'errore quadratico medio di proiezione e l'intervallo di confidenza della proiezione.....</i> | 38 |
| | <i>Il test di significatività dell'errore di proiezione.....</i> | 40 |
| | <i>Un'interpretazione alternativa del test del Chow.....</i> | 41 |
| | <i>Un'applicazione: effettuiamo il test del Chow con EasyReg.....</i> | 42 |
| 5.8 | Bibliografia..... | 46 |

Il paragrafo 5.7 e le applicazioni sono state redatte da Alberto Bagnai.

5.1 La proiezione dei minimi quadrati

Consideriamo nuovamente il modello lineare multiplo (1.3.4)

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + u_t \quad (5.1.1)$$

e poniamoci il problema di *proiettare* y_t fuori del campione che percorre il tempo $t = 1, 2, \dots, n$. In altre parole, vogliamo determinare y_{n+h} , per $h = 1, 2, \dots, m$, dove l'intervallo temporale $n+1, n+2, \dots, n+m$ è detto *periodo di proiezione*. Se utilizziamo il modello (5.1.1), stimato nel periodo campionario, per proiettare y_t occorre supporre che la struttura economica, già ipotizzata sostanzialmente invariante nel campione, rimanga la stessa nei due periodi, rendendo così possibile l'utilizzazione della stima dei minimi quadrati nella proiezione di y_t . Se inoltre supponiamo di conoscere i valori futuri delle variabili esplicative per $h = 1, 2, \dots, m$, e di far valere anche per il futuro le ipotesi deboli per i residui

$$E(\tilde{u}_{n+h}) = 0 \quad , \quad E(\tilde{u}_{n+h} \cdot \tilde{u}_s) = \begin{cases} \sigma^2 & s = n+h \\ 0 & s \neq n+h; \quad s = 1, 2, \dots, n+m \end{cases} \quad (5.1.2)$$

per ogni h , per cui risulta “naturale” prendere come proiezioni dei residui il loro valor medio, che è nullo, la proiezione \hat{y}_{n+h} al tempo $n+h$ è

$$\hat{y}_{n+h} = \hat{\beta}_1 x_{1n+h} + \hat{\beta}_2 x_{2n+h} + \dots + \hat{\beta}_k x_{kn+h} \quad (5.1.3)$$

In effetti \hat{y}_{n+h} non è tanto la proiezione di y_t quanto della sua componente sistematica, poiché la proiezione di u_t è stata posta arbitrariamente uguale a zero. Questa procedura, tuttavia, può essere giustificata in senso probabilistico se si considera, come ad esempio fatto dal de Finetti [1970] in ambito soggettivista, la proiezione di una variabile aleatoria come suo valor medio; in questo caso si ha

$$\hat{y}_{n+h} = E(\tilde{y}_{n+h}) = \beta_1 x_{1n+h} + \beta_2 x_{2n+h} + \dots + \beta_k x_{kn+h}$$

e i parametri β_i , sconosciuti, devono essere sostituiti da stime.

La proiezione (5.1.3) considerata come una variabile aleatoria funzione degli stimatori $\hat{\beta}_i$ diventa un *proiettore* della parte sistematica del modello, che è BLU se si utilizza il criterio di stima degli OLS: infatti è *lineare* rispetto alle y_t poiché lineari sono gli stimatori OLS; è *non distorto*

$$E(\hat{y}_{n+h}) = E(\hat{\beta}_1 x_{1n+h} + \hat{\beta}_2 x_{2n+h} + \dots + \hat{\beta}_k x_{kn+h}) = \beta_1 x_{1n+h} + \beta_2 x_{2n+h} + \dots + \beta_k x_{kn+h}$$

ed ottimo

$$\text{Var}(\hat{y}_{n+h}) = \text{Var}(\hat{\beta}_1 x_{1n+h} + \hat{\beta}_2 x_{2n+h} + \dots + \hat{\beta}_k x_{kn+h}) \leq \text{Var}(\tilde{b}_1 x_{1n+h} + \tilde{b}_2 x_{2n+h} + \dots + \tilde{b}_k x_{kn+h})$$

dove $\tilde{b}_1, \dots, \tilde{b}_k$ sono stimatori qualsiasi nella classe dei lineari e non distorti, in quanto vale la (1.8.2) se si prendono le costanti c_i pari a $x_{i,n+h}$, $i = 1, 2, \dots, k$.

L'errore di proiezione

Rimarchiamo il fatto che, a meno di non accettare l'impostazione del de Finetti, il proiettore (5.1.3) non è uno stimatore non distorto di y_{n+h} definito dalla (5.1.1), mentre lo è della sua componente sistematica. Esso, tuttavia, può essere considerato non distorto in un altro senso, che illustriamo facendo ricorso all'*errore di proiezione* definito nella maniera seguente

$$e_{n+h} = y_{n+h} - \hat{y}_{n+h} = \sum_{i=1}^k (\beta_i - \hat{\beta}_i) x_{in+h} + u_{n+h} \quad (5.1.4)$$

Poiché il valor medio dell'errore (5.1.4) è nullo

$$E(\tilde{e}_{n+h}) = \sum_{i=1}^k E(\beta_i - \hat{\beta}_i) x_{in+h} + E(\tilde{u}_{n+h}) = 0 \quad (5.1.5)$$

il proiettore \hat{y}_{n+h} può essere considerato come uno stimatore non distorto di y_{n+h} nel senso che il valor medio dell'errore di proiezione è nullo. In questo caso si dice che \hat{y}_{n+h} è un proiettore *incondizionatamente non distorto*.

L'errore quadratico medio di proiezione

La varianza dell'errore di proiezione è facilmente trovata in termini matriciali

$$\begin{aligned} Var(\tilde{e}_{n+h}) &= E(\tilde{e}_{n+h}^2) = Var[\mathbf{x}'_{n+h}(\mathbf{b} - \hat{\mathbf{b}})] + Var(\tilde{u}_{n+h}) = \\ &= \mathbf{x}'_{n+h} Cov(\mathbf{b} - \hat{\mathbf{b}}) \mathbf{x}_{n+h} + \sigma^2 = \sigma^2 [1 + \mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h}] \end{aligned} \quad (5.1.6)$$

dove nel secondo passaggio sono stati utilizzati la (5.1.4) ed il fatto che \tilde{u}_{n+h} è incorrelato per la (5.1.2) con tutti i residui del periodo campionario e quindi con gli stimatori $\hat{\beta}_i$; nel terzo passaggio è stata operata la (1.8.1) e nel quarto, infine, la (1.6.18). La varianza (5.1.6) è detta *errore quadratico medio di proiezione*¹ ed è spesso considerata come un indicatore della precisione della proiezione. Tanto più piccolo è questo errore tanto più precisa è la proiezione, per cui il proiettore (5.1.3) gode di una grande rilevanza valendo il seguente

Teorema 5.1 - Nella classe dei proiettori lineari (rispetto alle y_i) ed incondizionatamente non distorti, se i $\hat{\beta}_i$, $i = 1, 2, \dots, k$, sono gli stimatori dei minimi quadrati ordinari il proiettore (5.1.3) è quello che possiede errore quadratico medio minimo.

Infatti, utilizzando l'algebra delle matrici, ogni altro proiettore lineare nelle y_i è del tipo

¹ In lingua inglese: *Mean Square Error (MSE) of prediction*.

$$\bar{y}_{n+h} = [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}']\mathbf{y} \quad (5.1.7)$$

dove \mathbf{d} è un vettore di costanti arbitrarie di dimensione n . Il suo errore di proiezione è

$$\begin{aligned} y_{n+h} - \bar{y}_{n+h} &= \mathbf{x}'_{n+h} \mathbf{b} + u_{n+h} - [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}'](\mathbf{X}\mathbf{b} + \mathbf{u}) = \\ &= \mathbf{x}'_{n+h} \mathbf{b} - [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}']\mathbf{X}\mathbf{b} - [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}']\mathbf{u} + u_{n+h} = \\ &= u_{n+h} - \mathbf{d}'\mathbf{X}\mathbf{b} - [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}']\mathbf{u} \end{aligned}$$

che deve essere nullo, per ogni \mathbf{b} , sotto l'operatore valor medio affinché \bar{y}_{n+h} sia incondizionatamente non distorto. Occorre, quindi, che sia

$$\mathbf{d}'\mathbf{X} = \mathbf{0} \quad (5.1.8)$$

L'errore quadratico medio del proiettore (5.1.7) è dunque

$$\begin{aligned} E[(y_{n+h} - \bar{y}_{n+h})^2] &= \text{Var}(\tilde{u}_{n+h}) + E\left\{[\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \mathbf{d}']\tilde{\mathbf{u}}\tilde{\mathbf{u}}'[\mathbf{d} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h}]\right\} = \\ &= \sigma^2 \mathbf{d}'\mathbf{d} + \sigma^2 + \sigma^2 [\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h} + 2\mathbf{d}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h}] = \\ &= \sigma^2 \mathbf{d}'\mathbf{d} + \text{Var}(\tilde{e}_{n+h}) \end{aligned} \quad (5.1.9)$$

dove sono utilizzati la (5.1.6), la (5.1.8) ed il fatto che lo scalare $\mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{d}$ è uguale al suo trasposto; inizialmente, d'altro canto, si è fatto uso della non correlazione (5.1.2) tra \tilde{u}_{n+h} e ciascun elemento del vettore $\tilde{\mathbf{u}}$ associato ai tempi da 1 ad n .

La (5.1.9) è minima se $\mathbf{d}=\mathbf{0}$, cioè se il proiettore è quello definito dalla (5.1.3). La tesi è così dimostrata.

Proiezione ex ante ed ex post

Nelle pagine precedenti abbiamo considerato un periodo di proiezione esteso fuori del campione da $n+1$ a $n+m$ su un intervallo temporale nel quale i valori della variabile dipendente non sono ancora noti. Una proiezione di questo tipo viene detta proiezione *ex ante*, ad indicare appunto che essa viene effettuata *prima* di conoscere i valori storici della variabili proiettata. Si definisce invece proiezione *ex post* la proiezione effettuata *all'interno del campione*, cioè riferita a un intervallo temporale nel quale i valori storici della variabile dipendente sono disponibili (il termine *ex post* allude appunto al fatto che la proiezione viene effettuata *dopo* che i valori della dipendente sono stati rilevati). In quest'ultimo caso il campione di n osservazioni viene scisso in due sottocampioni, in modo analogo a quanto viene fatto quando si imposta un test di cambiamento di struttura (si veda il par. 3.4). Il primo campione consta di n_1 osservazioni e su esso si effettua la stima del modello, mentre il secondo (campione di proiezione *ex post*) consta di n_2 osservazioni. La proiezione viene quindi calcolata sulle n_2 osservazioni da n_1+1 a $n_1+n_2 = n$.

La proiezione *ex post* è effettuata sul passato, anziché sul futuro, della variabile, e quindi non ha una rilevanza operativa, ma ha una grande importanza in termini diagnostici. Infatti solo *ex post* è possibile misurare gli errori di proiezione del modello (errori che *ex ante* sono variabili aleatorie) e quindi valutare l'accuratezza della proiezione stessa. In termini intuitivi, si può affermare che se un'equazione non prevede bene il passato della variabile dipendente, generalmente le sue previsioni saranno poco accurate anche riguardo al futuro. La precisione della proiezione *ex post* è quindi una condizione necessaria, ma non sufficiente, perché si possa prestare fiducia al modello quando lo si impiega nella proiezione *ex ante*.

Per esemplificare i concetti esposti finora riprendiamo la funzione delle importazioni (3.3.29) e utilizziamola per effettuare una proiezione *ex post*. Il campione disponibile, lo ricordiamo, consta di dati trimestrali dal 1970 al 1989 per un totale di $n = 80$ osservazioni. Nel paragrafo 3.5 questa equazione è stata sottoposta a test di cambiamento di struttura considerando due sottocampioni rispettivamente di $n_1 = 25$ e $n_2 = 55$ osservazioni (l'osservazione n_1 -esima corrisponde al primo trimestre del 1976), ottenendo la statistica $F_{8,63}$ riportata nella (3.5.12), che risultava significativa, evidenziando la presenza di problemi di cambiamento di struttura.

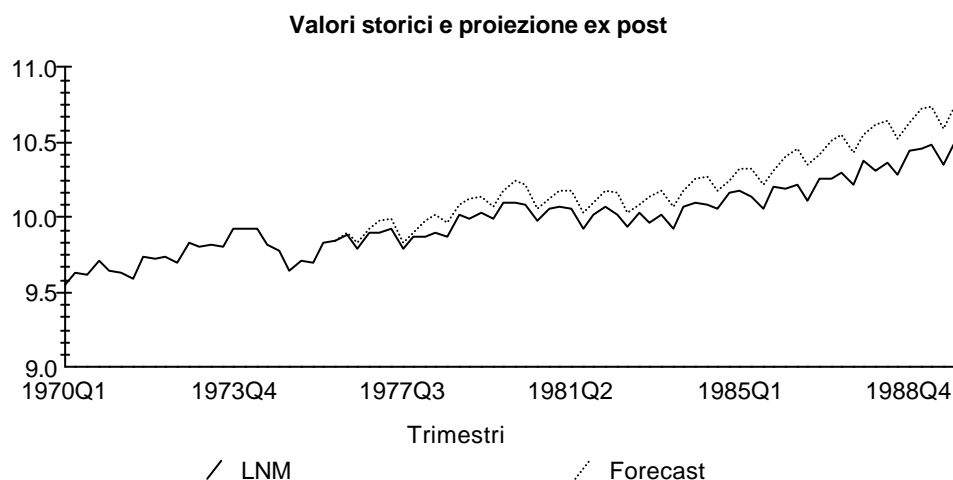


Figura 5.1 - I valori storici del logaritmo delle importazioni e quelli proiettati mediante la (5.1.8).

La probabile presenza di un cambiamento di struttura segnala che la proiezione *ex post* effettuata sul sottocampione da 1976:2 a 1989:4 potrà essere accurata. Per costruire la proiezione applichiamo la (5.1.3) all'equazione (3.5.11), tenendo presente che i coefficienti stimati sul primo sottoperiodo corrispondono a quelli

delle variabili esplicative non “spezzate”, cioè non moltiplicate per la variabile di comodo d_t , che serve appunto a rappresentare l’eventuale spostamento (*shift*) dei parametri fra il primo e il secondo sottoperiodo. In altri termini, la formula utilizzata per il calcolo delle proiezioni è

$$\ln \hat{y}_{n_1+h} = -10.94 + 0.92 \ln x_{1,n_1+h} + 0.99 \ln x_{2,n_1+h} - 0.28 \ln x_{3,n_1+h} + \\ + 0.22 \ln x_{4,n_1+h} + 0.10 d_{1,n_1+h} + 0.06 d_{2,n_1+h} + 0.03 d_{3,n_1+h} \quad (5.1.10)$$

dove, come abbiamo specificato, $n_1 = 25$ (corrispondente al primo trimestre del 1976) e $h = 1, \dots, n_2$, con $n_2 = 55$. Si noti che abbiamo ommesso la variabile di comodo puntuale $d_{73,t}$ poiché essa nel periodo di proiezione vale sempre zero.

Utilizzando la (5.1.10) e i valori delle esplicative nel secondo sottoperiodo si costruisce la proiezione *ex post* rappresentata nella figura 5.1. Come è lecito attendersi, data la presenza accertata di un cambiamento di struttura, la proiezione *ex post* è sistematicamente distorta. In particolare, il modello sovrastima sistematicamente l’andamento delle importazioni nel secondo sottoperiodo, cioè l’errore di proiezione è costantemente negativo. Da questo punto di vista quindi il grafico 5.1 non fa che confermare il risultato (3.5.12). La conseguenza operativa di questo risultato è che per ottenere un modello da utilizzare a scopo di previsione *ex ante* dovremo quanto meno stimare il modello eliminando le osservazioni antecedenti al cambiamento di struttura. I parametri stimati sull’intero campione sono infatti una mistura dei parametri corrispondenti al primo e al secondo regime, e quindi forniscono proiezioni distorte.²

Proviamo quindi a costruire un modello previsionale omettendo le osservazioni relative agli anni ’70: la prima osservazione corrisponde quindi al primo trimestre del 1980; omettiamo inoltre le ultime otto osservazioni, corrispondenti all’ultimo biennio del campione, per costruire con esse la proiezione *ex post*. Abbiamo quindi $n_1 = 32$ (corrispondente a 1987:4) e $n_2 = 8$. La stima della (3.3.29) sul primo sottocampione fornisce questi risultati

$$\ln \hat{y}_t = -13.82 + 1.15 \ln x_{1t} + 1.024 \ln x_{2t} + 0.007 \ln x_{3t} - 0.123 \ln x_{4t} + \\ (-6.8) \quad (5.4) \quad (13.6) \quad (0.17) \quad (-2.2) \\ + 0.036 d_{1t} - 0.023 d_{2t} - 0.013 d_{3t} \quad (5.1.11) \\ (3.0) \quad (-2.4) \quad (-1.0) \\ n = 32, R^2_c = 0.986, RSS = 0.0039, SEE = 0.012, JB = 0.67, BP = 7.3$$

² In questo ragionamento sono implicite le ipotesi che il cambiamento di struttura sia unico e sia stato correttamente individuato, e inoltre che il secondo regime si mantenga anche al di fuori del campione di stima. L’ultima ipotesi

dove gli indicatori diagnostici hanno il consueto significato e inoltre BP è la statistica del test di omoschedasticità di Breusch e Pagan [1979] condotto utilizzando una regressione ausiliaria analoga alla (4.3.14). Le stime non sono particolarmente soddisfacenti, né sotto il profilo dell'interpretazione economica (le variabili di prezzo hanno segno contrario a quello atteso e il deflatore delle importazioni non è significativo), né sotto il profilo statistico (la statistica di Breusch e Pagan si distribuisce come un χ_1^2 , con un valore soglia al 5% pari a 3.84, quindi l'ipotesi di omoschedasticità dei residui è respinta).

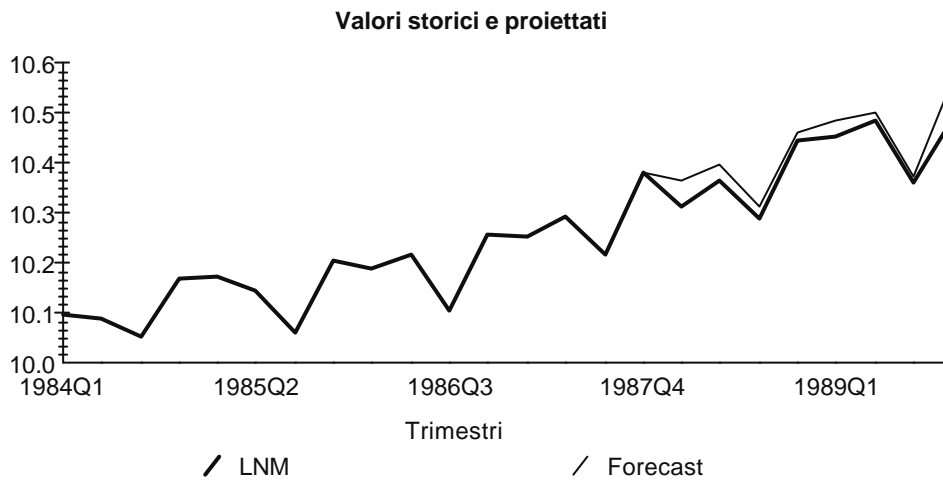


Figura 5.2 – I valori storici del logaritmo delle importazioni e quelli proiettati con la (5.1.9)-

L'emergere di problemi simili nella stima su un sottocampione evidenzia la necessità di una rispecificazione del modello. Anche in questo caso ci aspettiamo quindi che la proiezione *ex post* non sarà particolarmente accurata. I risultati relativi sono riportati in dettaglio nella tavola 5.1 e nella figura 5.2.

| | storici | proiettati | errore |
|--------|---------|------------|---------|
| 1988Q1 | 10.3132 | 10.3649 | -0.0517 |
| 1988Q2 | 10.3661 | 10.3993 | -0.0332 |
| 1988Q3 | 10.2915 | 10.3134 | -0.0219 |
| 1988Q4 | 10.4462 | 10.4624 | -0.0162 |
| 1989Q1 | 10.4544 | 10.4870 | -0.0326 |
| 1989Q2 | 10.4840 | 10.5007 | -0.0167 |
| 1989Q3 | 10.3607 | 10.3757 | -0.0149 |
| 1989Q4 | 10.4860 | 10.5611 | -0.0751 |

Tavola 5.1 – I valori storici del logaritmo delle importazione e quelli proiettati con la (5.1.9).

Anche in questo caso le proiezioni sono sistematicamente distorte verso l'alto, se pure in misura minore rispetto a quanto riscontrato nel precedente esperimento di proiezione. Si noti che siccome le variabili sono espresse in logaritmi, l'errore di proiezione ha dimensione percentuale. Ad esempio, nel primo trimestre del campione di proiezione *ex post* la proiezione sovrastima il valore storico in misura pari al 5.17%.

Dal punto di vista operativo questi risultati lasciano intuire che il fallimento del modello sotto il profilo dell'accuratezza delle previsioni non può essere semplicemente ricondotto all'esistenza di un singolo cambiamento di struttura nel 1976.

Indicatori dell'accuratezza delle proiezioni

Al fine di misurare l'accuratezza delle proiezioni vengono costruiti alcuni indicatori che fanno uso dell'insieme degli errori di previsione (5.1.4) considerati in tutto il periodo di proiezione. Naturalmente questi indicatori possono essere costruiti solo *ex post*, cioè solo una volta che sia disponibile il valore storico della variabile dipendente da confrontare con quello proiettato. I cinque comunemente più usati sono l'*errore medio*

$$ME = \frac{1}{n_2} \sum_{h=1}^{n_2} e_{n_1+h} \quad (5.1.12)$$

che è nella sostanza la media aritmetica degli n_1 errori; l'*errore medio assoluto*

$$MAE = \frac{1}{n_2} \sum_{h=1}^{n_2} |e_{n_1+h}| \quad (5.1.13)$$

che è la media aritmetica degli errori presi in valore assoluto (in modo da considerarli simmetricamente, sia i positivi che i negativi, mentre nell'*ME* gli uni si elidono con gli altri); l'*errore quadratico medio*

$$MSE = \frac{1}{n_2} \sum_{h=1}^{n_2} e_{n_1+h}^2 \quad (5.1.14)$$

che è la stima campionaria dell'errore quadratico medio di proiezione (5.1.6) e che spesso è considerato sotto radice quadrata aritmetica dando luogo alla *radice dell'errore quadratico medio*

$$RMSE = MSE^{1/2} \quad (5.1.15)$$

in modo da ottenere un indice della stessa dimensione dell'errore; ed infine il coefficiente di disuguaglianza di Theil³

$$U = \frac{\left(\sum_{h=1}^{n_2} e_{n_1+h}^2 \right)^{1/2}}{\left(\sum_{h=1}^{n_2} y_{n_1+h}^2 \right)^{1/2} + \left(\sum_{h=1}^{n_2} \hat{y}_{n_1+h}^2 \right)^{1/2}} \quad (5.1.16)$$

che vale zero quando tutti gli errori di proiezione sono nulli e tende ad uno man mano che l'accuratezza delle proiezioni peggiora.

Spesso è utile valutare alcuni degli indicatori precedenti in termini percentuali rispetto ad y_{n_1+h} al fine di disporre di una misura di errore indipendentemente dalla dimensione della variabile che si proietta: sostituendo nelle (5.1.13) e (5.1.14) al posto degli errori e_{n_1+h} i rapporti e_{n_1+h} / y_{n_1+h} si ottengono l'errore medio assoluto percentuale (*MAPE*) e l'errore quadratico medio percentuale (*MSPE*).

Queste misure di accuratezza sono utilizzate in pratica con gli errori e_{n_1+h} , $h = 1, 2, \dots, n_2$, calcolati tramite le proiezioni *ex post* \hat{y}_{n_1+h} ; le misure di accuratezza sono, dunque, numeriche.

Nel caso della (5.1.11) l'impianto dei calcoli per la costruzione delle misure di accuratezza è mostrato nella tavola 5.2, nella quale abbiamo riportato i valori storici e proiettati convertiti in unità naturali. In effetti, se le variabili sono logaritmizzate gli scarti fra valori storici e stimati hanno già la dimensione di una percentuale, quindi in questo caso non ha senso calcolare le misure descrittive percentuali. D'altra parte nelle applicazioni della proiezione è spesso opportuno, come strategia espositiva, riportare i risultati in unità naturali, in particolar modo quando si possa presumere che l'utente finale delle proiezioni abbia scarsa familiarità con le proprietà dei logaritmi (come generalmente avviene al di fuori dell'ambito della ricerca scientifica).

Nella tavola 5.2 si noti innanzitutto che gli errori medio e medio assoluto coincidono, come accade necessariamente quando l'errore è sistematicamente positivo o negativo. La coincidenza di questi due indicatori è quindi indizio di scarsa affidabilità delle previsioni del modello, poiché, indipendentemente dalle dimensioni dell'errore, la sua sistematicità denota appunto il fatto che il modello non esprime in modo adeguato tutta la componente sistematica del fenomeno

³ In lingua inglese gli indicatori sono: *mean prediction error (MPE)*, *mean absolute error (MAE)*, *mean square error (MSE)*, *root mean square error (RMSE)*, *Theil's inequality coefficient (U)*, rispettivamente.

osservato. Si noti poi, come anticipato commentando la tavola 5.1, che l'errore assoluto percentuale è approssimato dal valore assoluto dell'errore riferito ai logaritmi delle variabili. In termini puramente descrittivi la prestazione del modello non è pessima: l'errore assoluto percentuale è in media pari al 3.3%, valore che potrebbe anche essere considerato accettabile. Tuttavia, come abbiamo già rilevato, questo indicatore preso di per sé è ingannevole perché esprime unicamente la dimensione dell'errore, e non la sua sistematicità. Quest'ultima è indizio di gravi difetti nella specificazione del modello, difetti che, peraltro, risultano evidenti anche dall'interpretazione economica e della diagnostica dell'equazione (5.1.11).

| | Valori | | Errore | | | Errore percentuale | |
|--------------|----------|------------|----------|----------|------------|--------------------|------------|
| | storici | proiettati | semplice | assoluto | quadratico | assoluto | quadratico |
| 1988Q1 | 30127.69 | 31726.26 | -1598.57 | 1598.57 | 2555421.51 | 0.0531 | 0.0028 |
| 1988Q2 | 31764.35 | 32836.63 | -1072.28 | 1072.28 | 1149779.85 | 0.0338 | 0.0011 |
| 1988Q3 | 29480.96 | 30133.72 | -652.75 | 652.75 | 426088.64 | 0.0221 | 0.0005 |
| 1988Q4 | 34413.36 | 34975.39 | -562.04 | 562.04 | 315885.09 | 0.0163 | 0.0003 |
| 1989Q1 | 34696.70 | 35846.46 | -1149.75 | 1149.75 | 1321928.99 | 0.0331 | 0.0011 |
| 1989Q2 | 35739.08 | 36340.93 | -601.85 | 601.85 | 362228.36 | 0.0168 | 0.0003 |
| 1989Q3 | 31593.29 | 32070.76 | -477.47 | 477.47 | 227978.95 | 0.0151 | 0.0002 |
| 1989Q4 | 35810.63 | 38603.57 | -2792.94 | 2792.94 | 7800516.74 | 0.0780 | 0.0061 |
| medie | | | -1113.46 | 1113.46 | 1769978.52 | 0.0335 | 0.0016 |
| RMSE | | | | | 1330.41 | | |

Tavola 5.2 – Il calcolo delle misure descrittive di accuratezza della proiezione per la (5.1.11).

5.2 Intervalli di confidenza per le proiezioni

Supponiamo, ora, che valgano le ipotesi forti per i residui, sia nel periodo campionario che in quello di proiezione, per cui è anche

$$\tilde{u}_{n+h} \sim N(0, \sigma^2) \quad h = 1, 2, \dots, m$$

In questo caso l'errore di proiezione (5.1.4) è una variabile aleatoria costituita da una combinazione lineare degli stimatori $\hat{\mathbf{b}}$ distribuiti normalmente e di \tilde{u}_{n+h} anch'essa normale, per cui è

$$\tilde{e}_{n+h} \sim N(0, \sigma^2 \cdot a_{n+h}^2) \quad h = 1, 2, \dots, m$$

se poniamo

$$a_{n+h}^2 = 1 + \mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h} \quad (5.2.1)$$

in virtù delle (5.1.5) e (5.1.6).

Volendo trovare gli intervalli di confidenza per le y_{n+h} , che in realtà hanno un significato leggermente diverso dal consueto in quanto *ex ante* la \tilde{y}_{n+h} è una quantità aleatoria (funzione di \tilde{u}_{n+h}), consideriamo che

$$\frac{\tilde{e}_{n+h}}{\sigma \cdot a_{n+h}} = \frac{\tilde{y}_{n+h} - \hat{y}_{n+h}}{\sigma \cdot a_{n+h}} \sim N(0,1) \quad h = 1, 2, \dots, m$$

e che

$$\frac{\tilde{y}_{n+h} - \hat{y}_{n+h}}{a_{n+h} \left[\sum_{t=1}^n \hat{u}_t^2 / (n-k) \right]^{1/2}} = \frac{\tilde{y}_{n+h} - \hat{y}_{n+h}}{a_{n+h} \bar{\sigma}} \sim t_{n-k} \quad h = 1, 2, \dots, m \quad (5.2.2)$$

poiché \tilde{e}_{n+h} e $\sum_{t=1}^n \hat{u}_t$ sono variabili aleatorie stocasticamente indipendenti. Infatti

per la (5.1.4) \tilde{e}_{n+h} è formata dalle \hat{b}_i non correlate con le \hat{u}_i per la (2.3.4) e dalla \tilde{u}_{n+h} non correlata con le \hat{u}_i per la (5.1.2); essendo tutte queste variabili distribuite normalmente, vale l'asserita indipendenza stocastica.

La (5.2.2) può anche essere utilizzata *ex post* per costruire un test di significatività che permette di valutare l'accuratezza della proiezione in modo più rigoroso rispetto alle misure numeriche di accuratezza della proiezione (*MAE*, *MAPE*, ...) viste al termine del paragrafo precedente. Le informazioni sull'ordine di grandezza dell'errore di proiezione fornite da queste ultime non tengono conto della variabilità del fenomeno studiato, per cui, ad esempio, un errore medio assoluto percentuale del 5% può essere accettabile se la variabile dipendente è caratterizzata da estrema variabilità (ed è quindi intrinsecamente più difficile da

prevedere), mentre sarà inaccettabile se il fenomeno studiato ha andamento meno irregolare. La (5.2.2) normalizza l'errore di proiezione dividendolo per il proprio scarto quadratico medio e quindi elimina l'effetto della variabilità intrinseca del fenomeno, fornendo così una metrica nella quale è possibile effettuare in modo sensato valutazioni comparative fra modelli diversi.

Se la statistica definita dalla (5.2.2) è inferiore in valore assoluto al valore soglia pertinente della distribuzione t_{n-k} allora l'errore di proiezione non è statisticamente significativo. In questo caso, quindi, anche se l'errore risultasse "grande" in termini descrittivi, si dovrebbe concludere che le sue dimensioni sono legate alla variabilità del fenomeno, e non denotano né un evento aleatorio eccezionale (un'osservazione anomala), né un fallimento del modello (ad esempio, un cambiamento di struttura nei parametri di regressione). Viceversa, se la (5.2.2) è superiore (in valore assoluto) al relativo valore soglia, allora l'errore è statisticamente significativo, e quindi valgono considerazioni inverse, nel senso che anche qualora fosse l'errore "piccolo" in termini descrittivi esso denoterebbe pur sempre o il verificarsi di un evento eccezionale, o la presenza di problemi statistici nel modello (in particolare, di cambiamenti di struttura).

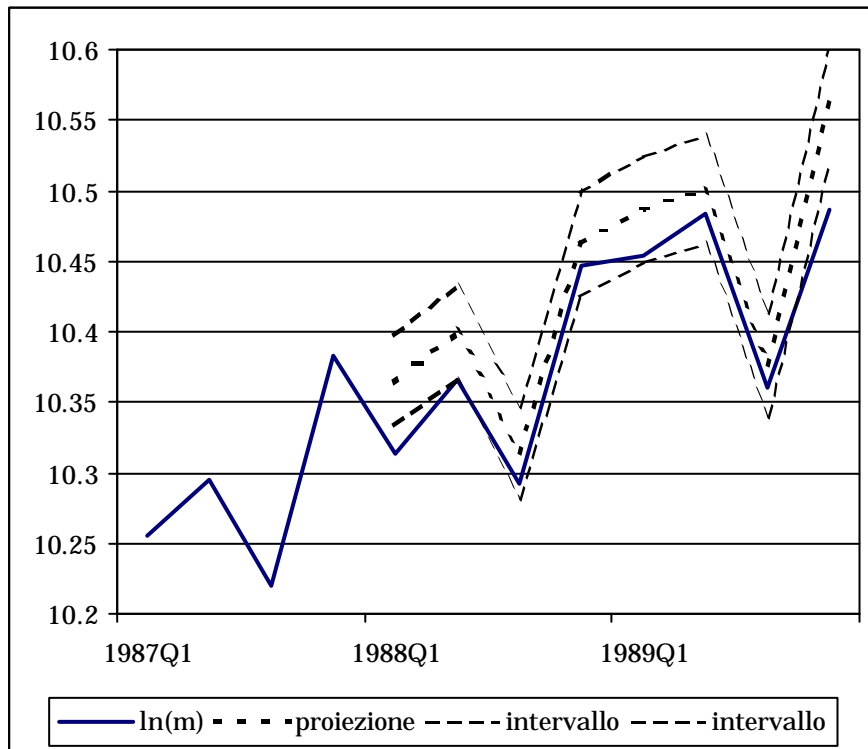


Figura 5.3 – L'intervallo di confidenza della proiezione ex post effettuata con la (5.1.11).

Sfruttando la consueta dualità fra verifica delle ipotesi e stima intervallare, l'intervallo di confidenza al livello α è trovato partendo dalla condizione

$$P\left(t'_{n-k} \leq \frac{\tilde{y}_{n+h} - \hat{y}_{n+h}}{a_{n+h} \bar{\sigma}} \leq t''_{n-k}\right) = 1 - \alpha$$

dove t'_{n-k} e t''_{n-k} sono i quantili di probabilità $\alpha/2$ e $1-\alpha/2$ rispettivamente, forniti dalle tavole della t di Student con $n-k$ gradi di libertà. L'intervallo risulta, dunque,

$$\hat{y}_{n+h} + a_{n+h} \cdot \bar{\sigma} \cdot t'_{n-k} \leq \tilde{y}_{n+h} \leq \hat{y}_{n+h} + a_{n+h} \cdot \bar{\sigma} \cdot t''_{n-k} \quad (5.2.3)$$

per $h = 1, 2, \dots, n_1$, e va interpretato come l'intervallo più corto che contiene il valore aleatorio \tilde{y}_{n+h} con probabilità $1-\alpha$.

La dualità fra verifica delle ipotesi e stima intervallare implica che *ex post* il test definito dalla (5.2.2) e l'intervallo di confidenza (5.2.3) forniscano esattamente le medesime informazioni, se pure in forma diversa. In altre parole, quando il valore storico ricade all'interno dell'intervallo di confidenza (5.2.3), la statistica (5.2.2) (cioè l'errore di proiezione studentizzato) non risulta statisticamente significativa, mentre, di converso, se la statistica (5.2.2) è maggiore, in valore assoluto, del rispettivo valore soglia, allora il valore storico cade al di fuori dell'intervallo di confidenza.

Riprendiamo ora la (5.1.11) corredando la tavola 5.1, che ne riporta la proiezione *ex post*, con gli scarti quadratici medi degli errori di proiezione, costruiti come radice quadrata della (5.1.6), e con le statistiche t (5.2.2). Questi risultati sono esposti nella tavola 5.3. Nel caso della proiezione *ex post* le statistiche (5.2.2) si distribuiscono come una t con n_1-k gradi di libertà; il relativo valore soglia, considerando che nella (5.1.11) si ha $n_1-k = 32-8 = 24$, è pari a 2.06. Ne consegue che il primo, il secondo e l'ultimo errore di proiezione sono significativamente diversi da zero, ovvero, in altri termini, che i relativi valori storici giacciono al di fuori dell'intervallo di confidenza al 5% della proiezione, come puntualmente si osserva nella figura 5.3, che riporta appunto la proiezione effettuata con la (5.1.11) insieme al relativo intervallo di confidenza.

| | storici | proiettati | errore | s.q.m. dell'errore | test t |
|--------|---------|------------|---------|--------------------|--------|
| 1988Q1 | 10.3132 | 10.3649 | -0.0517 | 0.0155 | -3.34 |
| 1988Q2 | 10.3661 | 10.3993 | -0.0332 | 0.0160 | -2.08 |
| 1988Q3 | 10.2915 | 10.3134 | -0.0219 | 0.0157 | -1.39 |
| 1988Q4 | 10.4462 | 10.4624 | -0.0162 | 0.0177 | -0.91 |
| 1989Q1 | 10.4544 | 10.4870 | -0.0326 | 0.0183 | -1.78 |
| 1989Q2 | 10.4840 | 10.5007 | -0.0167 | 0.0188 | -0.89 |
| 1989Q3 | 10.3607 | 10.3757 | -0.0149 | 0.0178 | -0.84 |
| 1989Q4 | 10.4860 | 10.5611 | -0.0751 | 0.0201 | -3.74 |

Tavola 5.3 – Il test di significatività dell'errore di proiezione sulle proiezioni *ex post* condotte con l'equazione (5.1.11). L'errore risulta significativamente diverso da zero in tre casi.

Un altro dato che segnala la scarsa affidabilità delle proiezioni effettuate con la (5.1.11) è dato dalla localizzazione degli errori significativi, che, contrariamente a quanto ci si potrebbe attendere, sono localizzati in maggioranza all'inizio, e non alla fine, del campione di proiezione *ex post*. Se astraiamo dall'ultima osservazione del campione di proiezione, gli errori più grandi si verificano nel primo e nel secondo periodo. In altri termini, sembra che il modello faccia più fatica a prevedere nell'immediato futuro che non nel futuro più remoto. Generalmente un risultato di questo tipo segnala la presenza di malaspecificazione del modello.

5.3 Il caso di variabili esplicative non note nel periodo di proiezione

Nei due paragrafi precedenti si è supposto che l'errore di proiezione e_{n+h} dipendesse

- dall'incertezza sui parametri riuniti nel vettore \mathbf{b} ,
- dal residuo aleatorio,

ipotizzando di conoscere il vettore \mathbf{x}_{n+h} delle osservazioni delle variabili esplicative al tempo $n+h$. In virtù di questa ultima ipotesi, le proiezioni sono definite *ex post*, in quanto appunto si suppone di conoscere i valori "futuri" delle esplicative. In molte circostanze, tuttavia, questi valori non sono noti e devono quindi essi stessi essere proiettati dando luogo alle proiezioni $\hat{\mathbf{x}}_{n+h}$, che supponiamo essere indipendenti da \tilde{u}_{n+h} , per cui la (5.1.3) diventa

$$\hat{y}_{n+h} = \hat{\mathbf{x}}'_{n+h} \hat{\mathbf{b}} \quad (5.3.1)$$

e l'errore di proiezione viene ad essere il seguente

$$\begin{aligned} e_{n+h} &= y_{n+h} - \hat{y}_{n+h} = \mathbf{x}'_{n+h} \mathbf{b} + u_{n+h} - \hat{\mathbf{x}}'_{n+h} \hat{\mathbf{b}} = \\ &= u_{n+h} - \hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b}) + (\mathbf{x}'_{n+h} - \hat{\mathbf{x}}'_{n+h}) \mathbf{b} = \\ &= u_{n+h} - \hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b}) + \mathbf{b}' (\mathbf{x}'_{n+h} - \hat{\mathbf{x}}'_{n+h}) \end{aligned} \quad (5.3.2)$$

scomposto nelle tre fonti di incertezza, la u_{n+h} , la $\hat{\mathbf{b}}$ e la $\hat{\mathbf{x}}_{n+h}$.

Se si fanno le due ulteriori ipotesi

$$E(\hat{\mathbf{x}}_{n+h}) = \mathbf{x}_{n+h} \quad E[\hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b})] = 0 \quad (5.3.3)$$

prendendo il valor medio della (5.3.2) si ottiene

$$E(\tilde{e}_{n+h}) = 0 \quad (5.3.4)$$

ed il proiettore (5.3.1) è ancora incondizionatamente non distorto.

L'errore quadratico di proiezione nel caso di \mathbf{x}_{n+h} non noto

La varianza dell'errore di proiezione (5.3.2) è data da

$$\begin{aligned} \text{Var}(\tilde{e}_{n+h}) &= E(\tilde{e}_{n+h}^2) = E[\tilde{u}_{n+h}^2 + \hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})' \hat{\mathbf{x}}_{n+h} + \\ &+ \mathbf{b}' (\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})(\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})' \mathbf{b} - 2\tilde{u}_{n+h} \hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b}) + \\ &+ 2\tilde{u}_{n+h} \mathbf{b}' (\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h}) - 2\hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b}) \mathbf{b}' (\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})] = \\ &= \sigma^2 + E[\hat{\mathbf{x}}'_{n+h} (\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})' \hat{\mathbf{x}}_{n+h}] + E[\mathbf{b}' (\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h}) \cdot (\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})' \mathbf{b}] \end{aligned} \quad (5.3.5)$$

poiché i valori medi dei prodotti incrociati si annullano in virtù dell'indipendenza di \tilde{u}_{n+h} , $\hat{\mathbf{x}}_{n+h}$ e $\hat{\mathbf{b}}$. Dal momento che è

$$\begin{aligned}
 E[\hat{\mathbf{x}}'_{n+h}(\hat{\mathbf{b}}-\mathbf{b})(\hat{\mathbf{b}}-\mathbf{b})'\hat{\mathbf{x}}_{n+h}] &= E\left\{tr\left[\hat{\mathbf{x}}'_{n+h}(\hat{\mathbf{b}}-\mathbf{b})(\hat{\mathbf{b}}-\mathbf{b})'\hat{\mathbf{x}}_{n+h}\right]\right\}= \\
 &= E\left\{tr\left[(\hat{\mathbf{b}}-\mathbf{b})(\hat{\mathbf{b}}-\mathbf{b})'\hat{\mathbf{x}}_{n+h}\hat{\mathbf{x}}'_{n+h}\right]\right\}= \\
 &= tr\left\{E\left[(\hat{\mathbf{b}}-\mathbf{b})(\hat{\mathbf{b}}-\mathbf{b})'\right]E\left[\hat{\mathbf{x}}_{n+h}\hat{\mathbf{x}}'_{n+h}\right]\right\}= \\
 &= tr\left\{Cov(\hat{\mathbf{b}})\cdot[Cov(\hat{\mathbf{x}}_{n+h})+\mathbf{x}_{n+h}\mathbf{x}'_{n+h}]\right\}
 \end{aligned}$$

poiché, in virtù della prima delle (5.3.3),

$$Cov(\hat{\mathbf{x}}_{n+h}) = E\left[(\hat{\mathbf{x}}_{n+h} - \mathbf{x}_{n+h})(\hat{\mathbf{x}}_{n+h} - \mathbf{x}_{n+h})'\right] = E(\hat{\mathbf{x}}_{n+h}\hat{\mathbf{x}}'_{n+h}) - \mathbf{x}_{n+h}\mathbf{x}'_{n+h}$$

ed inoltre è

$$E\left[\mathbf{b}'(\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})(\mathbf{x}_{n+h} - \hat{\mathbf{x}}_{n+h})'\mathbf{b}\right] = \mathbf{b}'Cov(\hat{\mathbf{x}}_{n+h})\mathbf{b}$$

la (5.3.5) diventa

$$\begin{aligned}
 Var(\tilde{e}_{n+h}) &= \sigma^2 + tr\left\{Cov(\hat{\mathbf{b}})\cdot[Cov(\hat{\mathbf{x}}_{n+h})+\mathbf{x}_{n+h}\mathbf{x}'_{n+h}]\right\} + \mathbf{b}'Cov(\hat{\mathbf{x}}_{n+h})\mathbf{b} = \\
 &= \sigma^2 + tr\left[Cov(\hat{\mathbf{b}})\cdot Cov(\hat{\mathbf{x}}_{n+h})\right] + tr\left[Cov(\hat{\mathbf{b}})\cdot\mathbf{x}_{n+h}\mathbf{x}'_{n+h}\right] + \mathbf{b}'Cov(\hat{\mathbf{x}}_{n+h})\mathbf{b} = \quad (5.3.6) \\
 &= \sigma^2 + tr\left[Cov(\hat{\mathbf{b}})\cdot Cov(\hat{\mathbf{x}}_{n+h})\right] + \mathbf{x}'_{n+h}Cov(\hat{\mathbf{b}})\mathbf{x}_{n+h} + \mathbf{b}'Cov(\hat{\mathbf{x}}_{n+h})\mathbf{b}
 \end{aligned}$$

uguale alla (5.1.6) qualora \mathbf{x}_{n+h} sia nota e la matrice di dispersione di $\hat{\mathbf{x}}_{n+h}$ sia quindi nulla.

In seguito vedremo come sia possibile calcolare questa matrice in certi casi di \mathbf{x}_{n+h} ignoto, e quindi determinare l'errore quadratico medio di proiezione (5.3.6).

In termini numerici, la proiezione \hat{y}_{n+h} viene determinata mediante la (5.3.1) una volta che siano disponibili le stime $\hat{\mathbf{b}}$ e le proiezioni $\hat{\mathbf{x}}_{n+h}$. Gli intervalli di confidenza sono viceversa difficilmente calcolabili perché, anche nell'ipotesi di $\hat{\mathbf{x}}_{n+h}$ ed \tilde{u}_{n+h} normali, l'errore di proiezione (5.3.2) è generalmente non normale poiché è funzione del prodotto dei due vettori aleatori $\hat{\mathbf{x}}'_{n+h}$ e $\hat{\mathbf{b}}$, ambedue con distribuzione normale multivariata.

Il teorema di Cebiscev

Una modo approssimato, suggerito in Feldstein [1971], per determinare intervalli di confidenza per la \tilde{y}_{n+h} si basa sulla diseuguaglianza di Cebiscev che viene dimostrata nel seguente

Teorema 5.2 - La probabilità che una variabile aleatoria \tilde{z} si discosti dal proprio valor medio μ più di δ volte lo scarto quadratico medio è minore o uguale ad $1/\delta^2$. In altri termini è

$$P(|\tilde{z} - \mu| \geq \delta\sigma) \leq 1/\delta^2 \quad (5.3.7)$$

ovvero

$$P(-\delta\sigma < \tilde{z} - \mu < \delta\sigma) > 1 - 1/\delta^2 \quad (5.3.8)$$

Poiché l'errore di proiezione $\tilde{e}_{n+h} = \tilde{y}_{n+h} - \hat{y}_{n+h}$ ha valor medio nullo e varianza σ_e^2 data dalla (5.3.6), applicando la (5.3.8) si ha

$$P(-\delta \cdot \sigma_e < \tilde{y}_{n+h} - \hat{y}_{n+h} < \delta \cdot \sigma_e) > 1 - 1/\delta^2$$

dalla quale si ricava l'intervallo di confidenza

$$\hat{y}_{n+h} - \delta \cdot \sigma_e < \tilde{y}_{n+h} < \hat{y}_{n+h} + \delta \cdot \sigma_e \quad (5.3.9)$$

che contiene il valore aleatorio \tilde{y}_{n+h} con probabilità maggiore di $1-1/\delta^2$. Se si prende soggettivamente $\alpha=1/\delta^2=0.05$, si ottiene $\delta=2 \cdot \sqrt{5}$, avendo considerato soltanto la radice quadrata aritmetica, e l'intervallo (5.3.9) diventa

$$\hat{y}_{n+h} - 2\sqrt{5} \cdot \sigma_e < \tilde{y}_{n+h} < \hat{y}_{n+h} + 2\sqrt{5} \cdot \sigma_e \quad (5.3.10)$$

che può essere determinato numericamente sostituendo a σ_e la radice quadrata aritmetica dell'errore quadratico medio (5.3.6) nella quale σ^2 è sostituita da una sua stima, ad esempio quella non distorta (1.7.2), \mathbf{x}_{n+h} è rimpiazzata da $\hat{\mathbf{x}}_{n+h}$, $Cov(\hat{\mathbf{b}})$ è ancora calcolata mediante la (1.6.17) e la stima (1.7.2), ed infine $Cov(\hat{\mathbf{x}}_{n+h})$ è determinata nella procedura di proiezione di \mathbf{x}_{n+h} .

L'accennata approssimazione dell'intervallo di confidenza (5.3.9) dipende appunto dalla sostituzione dei componenti dell'errore quadratico medio (5.3.6) con le loro stime. Se queste sono buone, l'approssimazione è molto accurata.

5.4 I residui ricorsivi

Il concetto, molto utile come vedremo in seguito, dei *minimi quadrati ricorsivi* si fonda sulla semplice procedura di stimare più volte il vettore dei parametri del modello lineare multiplo (5.1.1) utilizzando matrici \mathbf{X} con un numero di righe via via più grande. Proiettando la y_t del modello un tempo in avanti con l'uso delle stime ricorsive, si ottengono i *residui ricorsivi*, tramite i quali è possibile verificare alcune ipotesi relative all'invarianza nel tempo dei caratteri del modello lineare. La matrice \mathbf{X} di ordine $n \times k$ scritta in termini delle sue n righe diventa

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \dots \\ \mathbf{x}'_n \end{bmatrix}$$

e chiamando con \mathbf{X}_{t-1} la matrice di ordine $(t-1) \times k$ costituito dalle prime $t-1$ righe si può calcolare mediante essa la stima $\hat{\mathbf{b}}_{t-1}$ dei minimi quadrati ricorsivi

$$\hat{\mathbf{b}}_{t-1} = (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{X}'_{t-1} \mathbf{y}_{t-1} \quad (5.4.1)$$

dove \mathbf{y}_{t-1} è il vettore costituito dai primi $t-1$ elementi di \mathbf{y} , sotto le ipotesi che siano

$$t-1 \geq k, \quad r(\mathbf{X}_{t-1}) = k \quad (5.4.2)$$

Tramite $\hat{\mathbf{b}}_{t-1}$ si possono determinare la proiezione \hat{y}_t un tempo in avanti, in virtù della (5.1.3),

$$\hat{y}_t = \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} \quad t = k+1, k+2, \dots, n \quad (5.4.3)$$

e gli errori di proiezione

$$e_t = y_t - \hat{y}_t \quad (5.4.4)$$

che considerati come variabili aleatorie posseggono varianza fornita dalla (5.1.6)

$$\text{Var}(\tilde{e}_t) = \sigma^2 [1 + \mathbf{x}'_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{x}_t] = \sigma^2 \cdot a_t^2 \quad (5.4.5)$$

avendo sfruttato la posizione (5.2.1). Dividendo ciascun errore (5.4.4) per a_t si ottiene il residuo ricorsivo

$$v_t = (y_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1}) / a_t \quad t = k+1, k+2, \dots, n \quad (5.4.6)$$

che in termini aleatori possiede distribuzione $N(0, \sigma^2)$ se valgono le ipotesi forti sui residui \mathbf{u} del modello lineare multiplo dal momento che in questo caso $\tilde{e}_t \sim N(0, \sigma^2 a_t^2)$. Poiché deve essere $t \geq k+1$ il numero di residui ricorsivi che possono essere formati sulla base di un campione di ampiezza n è $n-k$. La loro indipendenza stocastica è dimostrata nel seguente

Teorema 5.3 - Se sussistono le ipotesi (5.4.2) per $t = k+1, k+2, \dots, n$, e se $\tilde{u} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ allora il vettore $\tilde{\mathbf{n}} = (\tilde{v}_{k+1}, \tilde{v}_{k+2}, \dots, \tilde{v}_n)$ dei residui ricorsivi ha distribuzione normale multivariata $N(\mathbf{0}, \sigma^2 \mathbf{I}_{n-k})$.

Si è appena mostrato che sotto le ipotesi del teorema $v_t \sim N(0, \sigma^2)$ per cui la tesi è dimostrata se si trova che $Cov(\tilde{v}_t, \tilde{v}_s) = E(\tilde{v}_t \tilde{v}_s) = 0$, per ogni t ed s , con $t \neq s$. Ma, in virtù del fatto che

$$\tilde{y}_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} = \tilde{u}_t + \mathbf{x}'_t (\mathbf{b} - \hat{\mathbf{b}}_{t-1}) = \tilde{u}_t - \mathbf{x}'_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{X}'_{t-1} \tilde{\mathbf{u}}_{t-1}$$

si ha che

$$\begin{aligned} E(\tilde{v}_t \cdot \tilde{v}_s) &= E[(\tilde{y}_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1})(\tilde{y}_s - \mathbf{x}'_s \hat{\mathbf{b}}_{s-1})] / a_t a_s = E[(\tilde{y}_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1})(\tilde{y}_s - \mathbf{x}'_s \hat{\mathbf{b}}_{s-1})'] / a_t a_s = \\ &= E\left\{ \left[\tilde{u}_t - \mathbf{x}'_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{X}'_{t-1} \tilde{\mathbf{u}}_{t-1} \right] \left[\tilde{u}'_s - \tilde{\mathbf{u}}'_{s-1} \mathbf{X}_{s-1} (\mathbf{X}'_{s-1} \mathbf{X}_{s-1})^{-1} \mathbf{X}_{s-1} \mathbf{x}_s \right] \right\} / a_t a_s = 0 \end{aligned}$$

poiché il trasposto di uno scalare è uguale allo scalare stesso e perché inoltre, per ogni t ed s e supponendo che $s > t$ senza perdere in generalità, si ha

$$\begin{aligned} E(\tilde{u}_t \tilde{u}_s) &= 0 & E(\tilde{\mathbf{u}}_{t-1} \tilde{\mathbf{u}}_s) &= \mathbf{0} \\ E(\tilde{u}_t \tilde{\mathbf{u}}'_{s-1}) &= [0 \quad 0 \quad \dots \quad 0 \quad \sigma^2 \quad 0 \quad \dots \quad 0]' \\ & \quad \quad \quad \uparrow \end{aligned}$$

elemento di indice t

$$E(\tilde{\mathbf{u}}_{t-1} \tilde{\mathbf{u}}'_{s-1}) = \sigma^2 [\mathbf{I}_{t-1} \quad \mathbf{0}]$$

dove la matrice $[\mathbf{I}_{t-1} \quad \mathbf{0}]$ è costituita da una parte sinistra uguale ad \mathbf{I}_{t-1} e da una parte destra uguale alla matrice nulla di ordine $(t-1) \times (s-t)$, come immediatamente si verifica.

Due formule per il calcolo ricorsivo

Il calcolo dei residui ricorsivi (5.4.6) presuppone l'utilizzazione della formula (5.4.1) per $n-k$ volte, e può essere computazionalmente pesante se n e k sono grandi in quanto in ogni passo è richiesta l'inversione di una nuova matrice di ordine k . Tuttavia, Brown, Durbin ed Evans [1975] hanno sviluppato due formule ricorsive per il calcolo della (5.4.1) che permettono di semplificare notevolmente il lavoro computazionale; esse sono

$$(\mathbf{X}'_t \mathbf{X}_t)^{-1} = (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} - (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{x}_t \mathbf{x}'_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} / a_t^2 \quad (5.4.7)$$

$$\hat{\mathbf{b}}_t = \hat{\mathbf{b}}_{t-1} + (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{x}_t (y_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1}) \quad (5.4.8)$$

dove lo scalare a_t^2 è dato dalla (5.2.1). Per mezzo di $(\mathbf{X}'_{t-1}\mathbf{X}_{t-1})$ e di \mathbf{x}_t si calcola a_t^2 , nonché $(\mathbf{X}'_t\mathbf{X}_t)^{-1}$ mediante la (5.4.7); quindi si determina $\hat{\mathbf{b}}_t$ partendo dalla stima precedente $\hat{\mathbf{b}}_{t-1}$ attraverso la (5.4.8).

La formula ricorsiva (5.4.7) è facilmente dimostrata se si nota che, poiché

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{X}_{t-1} \\ \mathbf{x}'_t \end{bmatrix} \quad (5.4.9)$$

è anche $\mathbf{X}'_t\mathbf{X}_t = \mathbf{X}'_{t-1}\mathbf{X}_{t-1} + \mathbf{x}_t\mathbf{x}'_t$; moltiplicando il membro a sinistra della (5.4.7) per $\mathbf{X}'_t\mathbf{X}_t$ si ottiene la matrice \mathbf{I}_k , alla quale anche si arriva moltiplicando il membro a destra della (5.4.7) per $\mathbf{X}'_{t-1}\mathbf{X}_{t-1} + \mathbf{x}_t\mathbf{x}'_t$.

Sfruttando ancora la scomposizione (5.4.9) si dimostra anche la (5.4.8); infatti, ricordando la definizione (5.4.1) di stima dei minimi quadrati ricorsivi, si ha

$$\begin{aligned} (\mathbf{X}'_t\mathbf{X}_t)\hat{\mathbf{b}}_t &= \mathbf{X}'_t\mathbf{y}_t = [\mathbf{X}'_{t-1}\mathbf{x}_t]\mathbf{y}_t = \mathbf{X}'_{t-1}\mathbf{y}_{t-1} + \mathbf{x}_t\mathbf{y}_t = (\mathbf{X}'_{t-1}\mathbf{X}_{t-1})\hat{\mathbf{b}}_{t-1} + \mathbf{x}_t\mathbf{y}_t = \\ &= (\mathbf{X}'_{t-1}\mathbf{X}_{t-1})\hat{\mathbf{b}}_{t-1} + \mathbf{x}_t\mathbf{x}'_t\hat{\mathbf{b}}_{t-1} + \mathbf{x}_t(\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1}) = (\mathbf{X}'_t\mathbf{X}_t)\hat{\mathbf{b}}_{t-1} + \mathbf{x}_t(\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1}) \end{aligned}$$

La devianza residua ricorsiva

Anche per la devianza residua sussiste una formula ricorsiva molto utile, data nel seguente

Teorema 5.4 - Se $\hat{\mathbf{u}}_t = \mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_t$ è vettore di residui stimati con i minimi quadrati ricorsivi, allora vale la seguente formula per la devianza residua

$$\hat{\mathbf{u}}'_t\hat{\mathbf{u}}_t = \hat{\mathbf{u}}'_{t-1}\hat{\mathbf{u}}_{t-1} + v_t^2 \quad t = k+1, k+2, \dots, n \quad (5.4.10)$$

Infatti, in virtù della (5.4.8), si ha

$$\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_t = \mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1} - \mathbf{X}_t(\hat{\mathbf{b}}_t - \hat{\mathbf{b}}_{t-1}) = \mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1} - \mathbf{X}_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{x}_t(\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1})$$

per cui

$$\begin{aligned} \hat{\mathbf{u}}'_t\hat{\mathbf{u}}_t &= (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1})'(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1}) + (\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1})^2\mathbf{x}'_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{x}_t + \\ &\quad - 2(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1})'\mathbf{X}_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{x}_t(\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1}) = \\ &= (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1})'(\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1}) - (\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1})^2\mathbf{x}'_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{x}_t \end{aligned} \quad (5.4.11)$$

dove nell'ultimo passaggio si è sfruttata l'uguaglianza seguente

$$\begin{aligned} (\mathbf{y}_t - \mathbf{X}_t\hat{\mathbf{b}}_{t-1})'\mathbf{X}_t &= \mathbf{y}'_t\mathbf{X}_t - \hat{\mathbf{b}}'_{t-1}\mathbf{X}'_t\mathbf{X}_t = [\mathbf{y}'_{t-1} \quad \mathbf{y}_t] \begin{bmatrix} \mathbf{X}_{t-1} \\ \mathbf{x}'_t \end{bmatrix} - \hat{\mathbf{b}}'_{t-1}(\mathbf{X}'_{t-1}\mathbf{X}_{t-1} + \mathbf{x}_t\mathbf{x}'_t) = \\ &= \mathbf{y}'_{t-1}\mathbf{X}_{t-1} + \mathbf{y}_t\mathbf{x}'_t - (\mathbf{y}'_{t-1}\mathbf{X}_{t-1} + \hat{\mathbf{b}}'_{t-1}\mathbf{x}_t\mathbf{x}'_t) = (\mathbf{y}_t - \mathbf{x}'_t\hat{\mathbf{b}}_{t-1})\mathbf{x}'_t \end{aligned}$$

in virtù della (5.4.1) ed essendo $\hat{\mathbf{b}}'_{t-1} \mathbf{x}_t$ uno scalare uguale al suo trasposto.

$$\begin{aligned} \text{Quindi, ricordando ancora la (5.4.9) e la scomposizione } \mathbf{y}'_t &= [\mathbf{y}'_{t-1} \quad y_t], \text{ è} \\ (\mathbf{y}_t - \mathbf{X}_t \hat{\mathbf{b}}_{t-1})' (\mathbf{y}_t - \mathbf{X}_t \hat{\mathbf{b}}_{t-1}) &= \mathbf{y}'_t \mathbf{y}_t - 2\mathbf{y}'_t \mathbf{X}_t \hat{\mathbf{b}}_{t-1} + \hat{\mathbf{b}}'_{t-1} \mathbf{X}'_t \mathbf{X}_t \hat{\mathbf{b}}_{t-1} = \\ &= \mathbf{y}'_{t-1} \mathbf{y}_{t-1} + y_t^2 - 2(\mathbf{y}'_{t-1} \mathbf{X}_{t-1} + y_t \mathbf{x}'_t) \hat{\mathbf{b}}_{t-1} + \hat{\mathbf{b}}'_{t-1} \mathbf{X}'_{t-1} \mathbf{X}_{t-1} \hat{\mathbf{b}}_{t-1} + \hat{\mathbf{b}}'_{t-1} \mathbf{x}_t \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} = \\ &= \hat{\mathbf{u}}'_{t-1} \hat{\mathbf{u}}_{t-1} + y_t^2 - 2y_t \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} + \hat{\mathbf{b}}'_{t-1} \mathbf{x}_t \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} \end{aligned}$$

per cui, sostituendo e sfruttando le (5.4.6) e (5.4.7), la (5.4.11) diventa

$$\begin{aligned} \hat{\mathbf{u}}'_t \hat{\mathbf{u}}_t &= \hat{\mathbf{u}}'_{t-1} \hat{\mathbf{u}}_{t-1} + y_t^2 - 2y_t \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} + \hat{\mathbf{b}}'_{t-1} \mathbf{x}_t \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} + \\ &\quad - v_t^2 a_t^2 \mathbf{x}'_t [(\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} - (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} \mathbf{x}_t \mathbf{x}'_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1} / a_t^2] \cdot \mathbf{x}_t = \\ &= \hat{\mathbf{u}}'_t \hat{\mathbf{u}}_t + y_t^2 - (2y_t - \hat{\mathbf{b}}'_{t-1} \mathbf{x}_t) \mathbf{x}'_t \hat{\mathbf{b}}_{t-1} - v_t^2 [a_t^2 - \mathbf{x}'_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{x}_t] \cdot \mathbf{x}'_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{x}_t = \\ &= \hat{\mathbf{u}}'_t \hat{\mathbf{u}}_t + y_t^2 - (y_t + v_t a_t)(y_t - v_t a_t) - v_t^2 \mathbf{x}'_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{x}_t = \\ &= \hat{\mathbf{u}}'_t \hat{\mathbf{u}}_t + v_t^2 a_t^2 - v_t^2 \mathbf{x}'_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{x}_t = \hat{\mathbf{u}}'_{t-1} \hat{\mathbf{u}}_{t-1} + v_t^2 \end{aligned}$$

e la tesi è dimostrata.

Minimi quadrati ricorsivi e diagnostica del modello

Se le ipotesi stocastiche forti sono verificate, avremo in particolare che i parametri del modello sono costanti e deve anche verificarsi che con l'aumentare della numerosità campionaria, aumentando l'informazione statistica disponibile per la stima dei parametri, le stime stesse divengano più precise, il che significa da un lato che le t di Student dei coefficienti dovrebbero diventare più significative, e dall'altro, per la dualità fra stima puntuale e stima intervallare, che l'intervallo di confidenza costruito attorno ai singoli coefficienti stimati dovrebbe restringersi; analogamente, con l'aumentare dell'informazione campionaria lo scarto quadratico medio dei residui del modello dovrebbe decrescere in modo pressoché monotono (in corrispondenza con il decrescere dell'incertezza sulle stime determinata dall'aumentare dell'informazione campionaria).

I metodi ricorsivi offrono uno strumento immediato per verificare se il modello si conforma a queste previsioni: basta rappresentare il grafico dei $\hat{\beta}_{i,t}$ insieme ai relativi intervalli di confidenza, e quello di $\hat{\sigma}_t$, costruito come radice quadrata di $\hat{\mathbf{u}}'_t \hat{\mathbf{u}}_t / (t-k)$ e osservarne l'andamento. L'applicazione di questi strumenti diagnostici alla (3.3.29) porta ai grafici rappresentati nelle figure dalla (5.5) alla (5.13).⁴

⁴ In sede di stima ricorsiva abbiamo eliminato la variabile di comodo puntuale $d_{73,t}$, che è perfettamente collineare con l'intercetta nelle prime dodici osservazioni del campione (da 70:1 a 72:4 la $d_{73,t}$ è un vettore di zeri, perfettamente collineare al vettore di uno che rappresenta l'intercetta).

Il primo grafico riporta l'indicatore diagnostico più sintetico, dato da $\hat{\sigma}_t$. Si vede bene come all'aumentare della numerosità campionaria lo scarto quadratico medio della regressione non tenda a diminuire, e anzi aumenti in modo marcato in almeno tre episodi, corrispondenti rispettivamente all'inizio del campione, ai trimestri da 1975:2 a 1976:1 (in corrispondenza dei quali abbiamo già localizzato un cambiamento di struttura, che abbiamo datato in 1976:1 sulla base dell'osservazione del grafico dei residui *OLS*), e ancora ai trimestri da 1984:2 a 1985:1. Quest'ultimo risultato mostra quindi che i cambiamenti di struttura nel campione a nostra disposizione sono stati almeno due: oltre a quello individuato dalla statistica (3.5.12) ne abbiamo un altro verosimilmente a metà degli anni '80. Quest'ultimo potrebbe spiegare come mai anche riducendo il campione di stima ai soli anni '80 (e quindi passando dalla (5.1.10) alla (5.1.11)) la proiezione *ex post* rimanga comunque distorta verso l'alto.

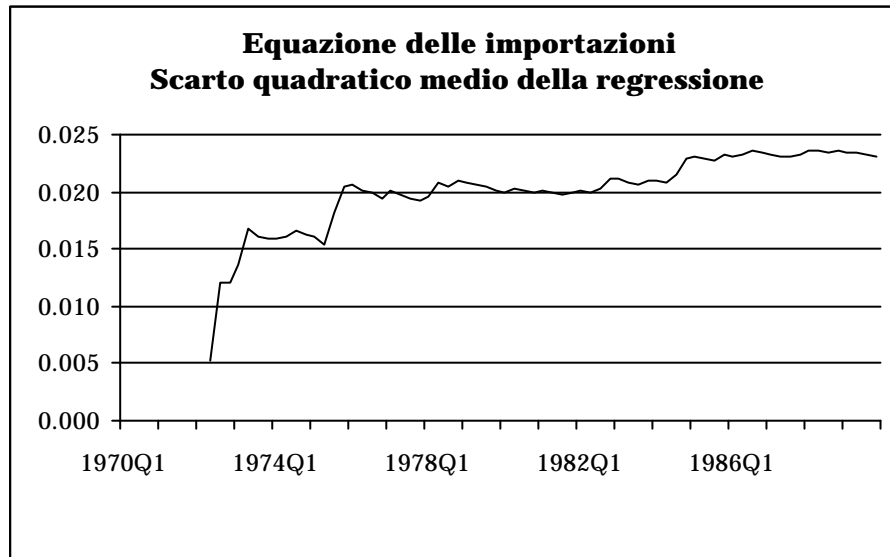


Figura 5.5 – Lo scarto quadratico medio della regressione calcolato partendo dalle stime ricorsive dell'equazione delle importazioni (3.3.29).

I grafici delle stime ricorsive intervallari dei singoli coefficienti possono fornirci indicazioni circa l'origine di questi cambiamenti di struttura, indicandoci in particolare se essi siano associati o meno ai coefficienti di specifiche variabili.

Notiamo intanto che tutti i grafici dal 5.6 al 5.13 manifestano una frattura attorno al 1976, e che generalmente in corrispondenza di essa le stime diventano molto meno disperse (cioè il loro intervallo di confidenza si restringe). Attorno al 1985 si verifica poi un significativo slittamento verso il basso dell'elasticità delle importazioni ai consumi (figura 5.7), in sincrono con un opposto e più deciso

slittamento verso l'alto dell'elasticità a investimenti e esportazioni (figura 5.8). La stima ricorsiva mostra che i coefficienti delle variabili di prezzo sono relativamente più stabili verso la metà degli anni '80, e inoltre che l'elasticità ai prezzi interni (figura 5.10) è non significativamente diversa da zero sulla maggior parte del campione.

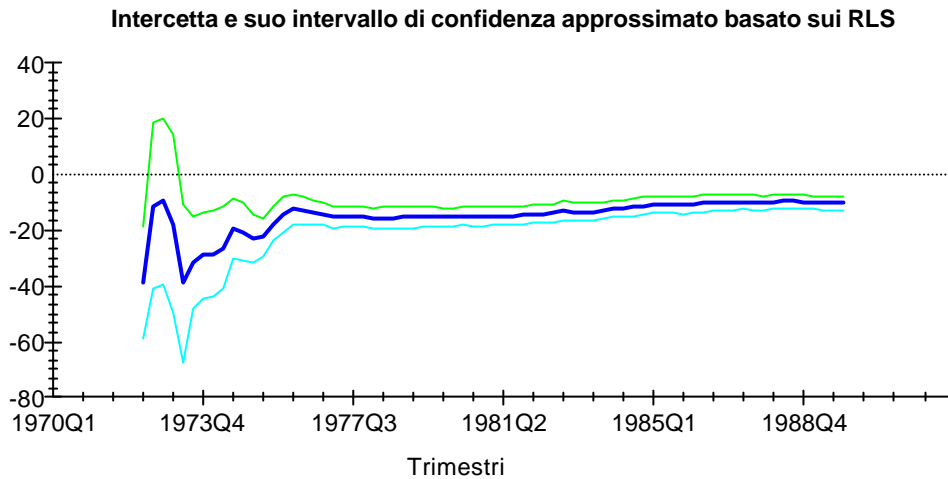


Figura 5.6 - La stima ricorsiva del coefficiente dell'intercetta.

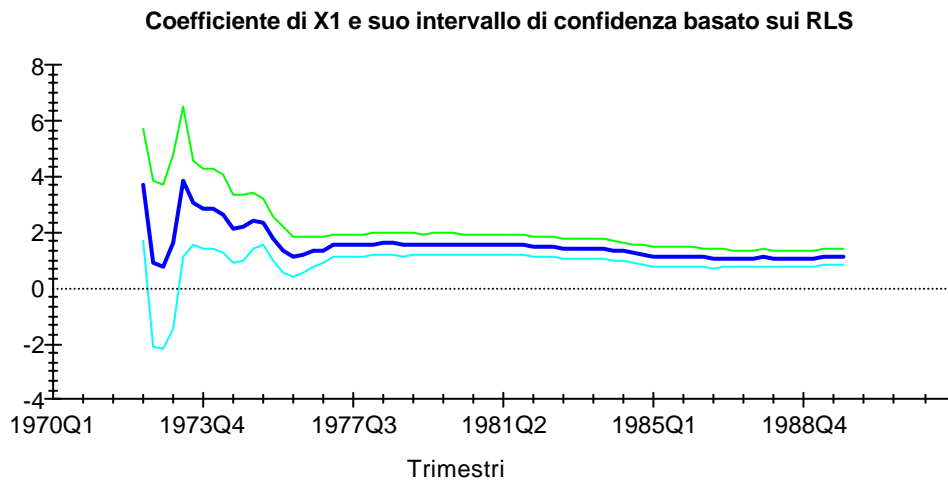


Figura 5.7 - La stima ricorsiva del coefficiente dell'elasticità delle importazioni ai consumi.

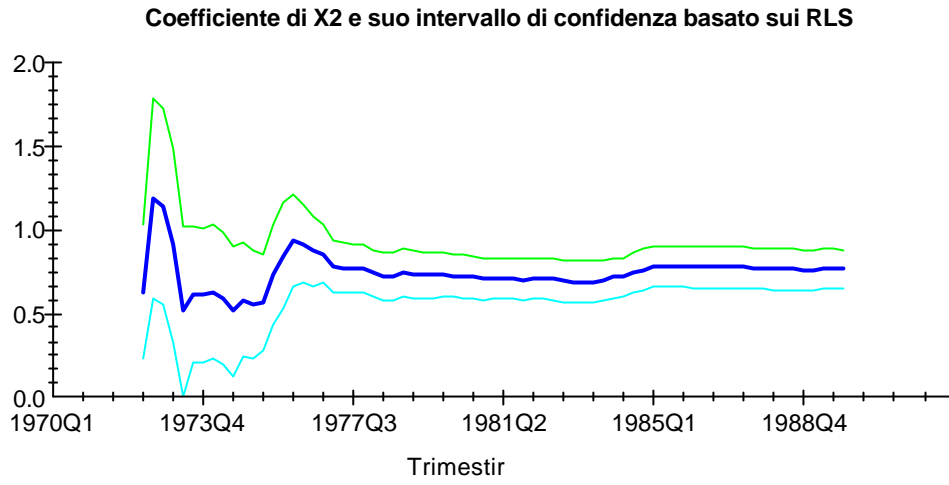


Figura 5.8 - La stima ricorsiva del coefficiente dell'elasticità delle importazioni a investimenti e esportazioni.

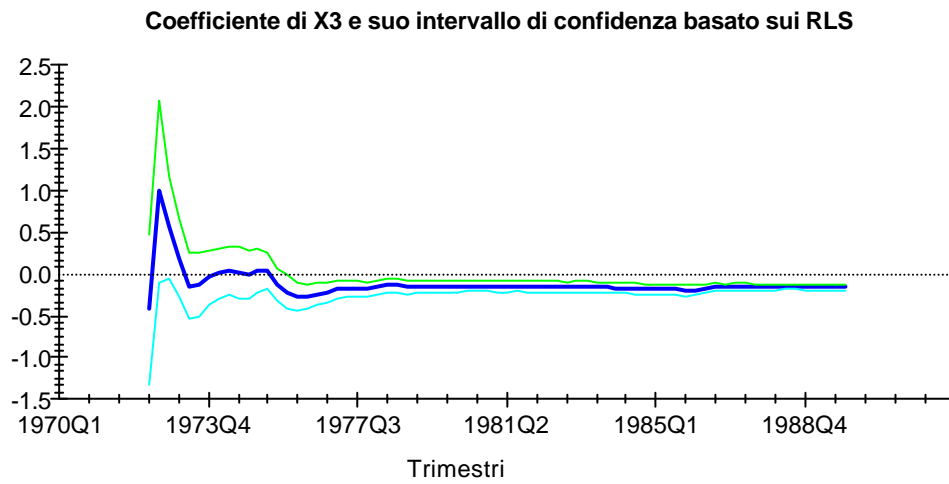


Figura 5.9 - La stima ricorsiva dell'elasticità delle importazioni ai prezzi import.

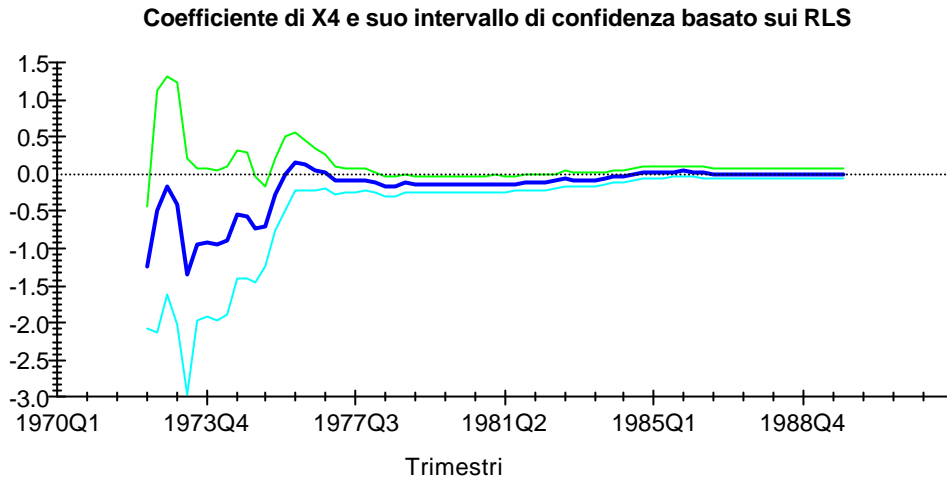


Figura 5.10 - La stima ricorsiva dell'elasticità delle importazioni ai prezzi interni.

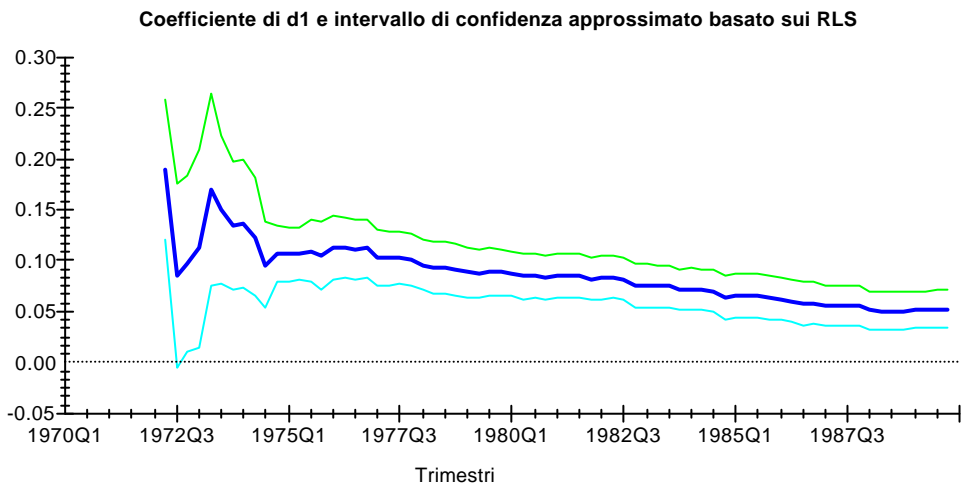


Figura 5.11 - La stima ricorsiva del coefficiente della prima variabile di comodo stagionale.

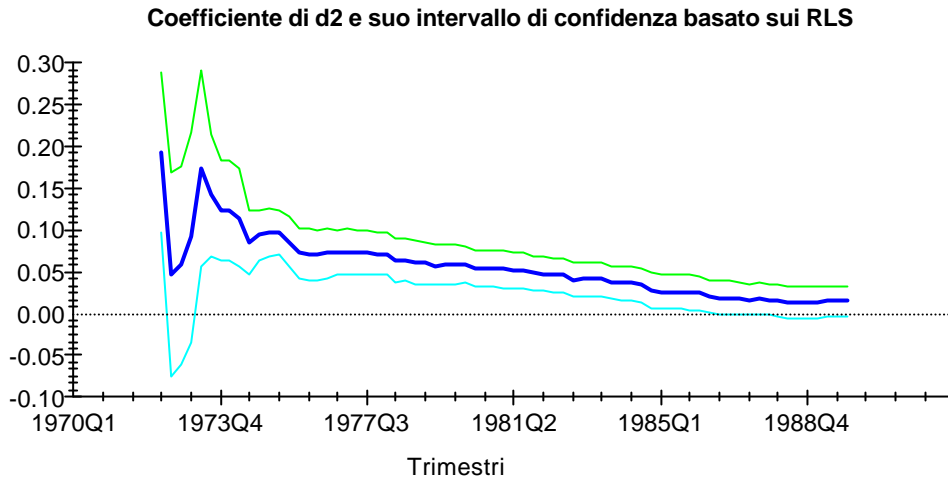


Figura 5.12 - La stima ricorsiva del coefficiente della seconda variabile di comodo stagionale.

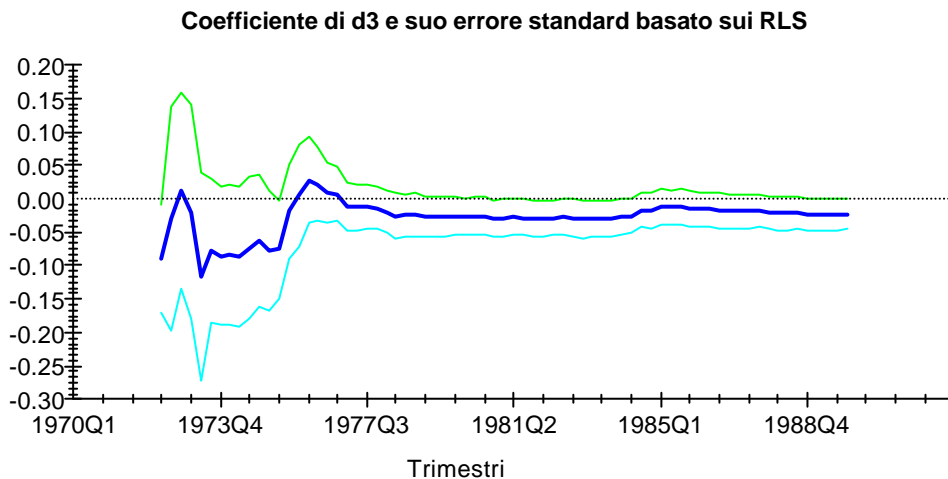


Figura 5.13 - La stima ricorsiva del coefficiente della terza variabile di comodo stagionale.

Viceversa, le variabili di comodo stagionali presentano coefficienti instabili su tutto il campione: in particolare, i coefficienti delle prime due hanno una marcata tendenza negativa (figure 5.11 e 5.12), mentre il coefficiente della terza risente del cambiamento di struttura a metà degli anni '80, in corrispondenza del quale diventa non significativo.

Questa evidenza conferma che nella (3.3.29) è la rappresentazione della stagionalità è malspecificata, e inoltre che a metà degli anni '80 si verifica un ulteriore cambiamento di struttura riguardante i coefficienti delle due variabili di scala.

5.5 Due test di cambiamento strutturale nel caso $n_2 \leq k$

Nel test (di malaspecificazione) basato sull'ipotesi (3.5.1) si è supposto che $n_1 > k$ ed $n_2 > k$ poiché la procedura costruita equivale a stimare separatamente con il criterio dei minimi quadrati ordinari una equazione nella prima parte del campione ed un'altra nella seconda.

Generalmente la condizione $n_1 > k$ è abbondantemente soddisfatta nelle applicazioni (se si utilizzano dati temporali è spesso possibile *allungare* le serie storiche all'indietro qualora n_1 sia troppo piccola), mentre sovente non lo è l'altra $n_2 > k$; è allora necessario ricorrere a test diversi dai precedenti, uno dei quali si fonda sul fatto che se è valida l'ipotesi nulla (3.5.1) non vi è cambiamento strutturale tra il primo ed il secondo sottoperiodo e si può calcolare la devianza residua $\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0$ considerando tutte le $n_1 + n_2$ osservazioni. Se l'ipotesi nulla non vale si può determinare la devianza residua relativamente al solo primo sottocampione, che indichiamo con $\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1$ ed alla quale sono relativi $n_1 - k$ gradi di libertà.

Se l'ipotesi nulla è vera, la differenza $\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0 - \hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1$ è piccola, e questa conseguenza induce a costruire la statistica

$$\frac{(\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0 - \hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1) / n_2}{\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1 / (n_1 - k)} = F \quad (5.5.1)$$

sulla quale basare il test di cambiamento strutturale. Questa statistica è simile ma non identica alla (3.5.5) per cui non possono essere adoperate le argomentazioni del capitolo 3 per affermare che ha distribuzione della F di Fisher. Per questo, è necessario far uso del teorema 5.4 relativo alla devianza residua ricorsiva, secondo l'impostazione di Harvey [1976].

Iterando la formula (5.4.10) ed osservando che $\hat{\mathbf{u}}_k' \hat{\mathbf{u}}_k = 0$ poiché un modello lineare con k parametri stimato su k osservazioni si adatta ad esse perfettamente così da avere residui tutti nulli, si ha

$$\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0 = \sum_{t=k+1}^{n_1+n_2} v_t^2 \quad (5.5.2)$$

e

$$\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1 = \sum_{t=k+1}^{n_1} v_t^2 \quad (5.5.3)$$

per cui la variabile aleatoria

$$(\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0 - \hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1) / \sigma^2 = \sum_{t=n_1+1}^{n_1+n_2} \tilde{v}_t^2 / \sigma^2$$

è formata dalla somma dei quadrati di n_2 variabili normali standardizzate ed indipendenti per il teorema 5.3, e quindi possiede una distribuzione del chi quadrato con n_2 gradi di libertà. D'altro canto la variabile aleatoria $\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1 / \sigma^2$ per gli stessi motivi è distribuita come un χ^2 con $n_1 - k$ gradi di libertà e quindi il rapporto (5.5.1) possiede distribuzione $F_{n_2, n_1 - k}$. Numeratore e denominatore della (5.5.1) sono indipendenti perché formati da residui ricorsivi diversi.

La procedura di esecuzione di questo test di cambiamento strutturale è composto da quattro passi:

- i) si stima modello lineare (5.1.1) sul campione costituito da tutte le $n_1 + n_2$ osservazioni e si calcola la devianza $\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0$;
- ii) si stima il modello con le prime n_1 osservazioni e si calcola devianza $\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1$;
- iii) si calcola la statistica F data dalla (5.5.1) e la si confronta con il valore critico F' determinato per mezzo della distribuzione della F di Fisher con n_2 e $n_1 - k$ gradi di libertà al livello di significatività α prescelto;
- iv) se $F \leq F'$ si è indotti ad accettare l'ipotesi nulla di assenza di cambiamento di struttura, altrimenti la si rifiuta, nel qual caso occorrerà procedere a una rispecificazione del modello.

Il test del Chow basato sulle proiezioni

Un secondo test utilizzabile per verificare l'ipotesi di cambiamento strutturale nel caso in cui sia $n_2 \leq k$ è stato sviluppato dal Chow [1960] sulla base delle proiezioni effettuate nel secondo sottoperiodo campionario mediante il modello stimato nel primo. Il modello lineare nel primo sotto periodo ($t=1, 2, \dots, n_1$) sia

$$\mathbf{y}_1 = \mathbf{X}_1 \mathbf{b} + \mathbf{u}_1 \quad (5.5.4)$$

con stima dei parametri

$$\hat{\mathbf{b}} = (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{y}_1 = \mathbf{b} + (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{u}_1 \quad (5.5.5)$$

mentre nel secondo sottoperiodo ($t= n_1 + 1, n_1 + 2, \dots, n_1 + n_2$) sia

$$\mathbf{y}_2 = \mathbf{X}_2 \mathbf{b} + \mathbf{u}_2$$

dove \mathbf{b} viene stimato mediante la (5.5.5), per cui si possono effettuare le proiezioni $\hat{y}_{n_1+1}, \hat{y}_{n_1+2}, \dots, \hat{y}_{n_1+n_2}$, raccolte nel vettore

$$\hat{\mathbf{y}}_2 = \mathbf{X}_2 \hat{\mathbf{b}}$$

Tramite gli errori di proiezione relativi, contenuti nel vettore

$$\mathbf{e} = \mathbf{y}_2 - \hat{\mathbf{y}}_2 = \mathbf{X}_2 \mathbf{b} + \mathbf{u}_2 - \mathbf{X}_2 \hat{\mathbf{b}} = \mathbf{X}_2 \mathbf{b} + \mathbf{u}_2 - \mathbf{X}_2 \mathbf{b} - \mathbf{X}_2 (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{u}_1 = \mathbf{u}_2 - \mathbf{X}_2 (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{u}_1$$

si può costruire il test basandolo sul fatto che se essi sono sufficientemente piccoli si può ritenere che le stime $\hat{\mathbf{b}}$ siano valide anche nel secondo sottoperiodo e che quindi valga l'ipotesi nulla di inesistenza di cambiamento strutturale.

Il valor medio del vettore degli errori di proiezione è

$$E(\tilde{\mathbf{e}}) = \mathbf{0}$$

e la sua matrice di dispersione

$$\begin{aligned} E(\tilde{\mathbf{e}}\tilde{\mathbf{e}}') &= E\left\{\left[\tilde{\mathbf{u}}_2 - \mathbf{X}_2(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\tilde{\mathbf{u}}_1\right]\left[\tilde{\mathbf{u}}_2' - \tilde{\mathbf{u}}_1'\mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_2\right]\right\} = \\ &= \sigma^2\left[\mathbf{I}_{n_2} + \mathbf{X}_2(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_2\right] = \sigma^2\mathbf{V} \end{aligned}$$

dove si è sfruttato il fatto che $E(\tilde{\mathbf{u}}_2\tilde{\mathbf{u}}_2') = \sigma^2\mathbf{I}_{n_2}$, $E(\tilde{\mathbf{u}}_1\tilde{\mathbf{u}}_1') = \sigma^2\mathbf{I}_{n_1}$, $E(\tilde{\mathbf{u}}_2\tilde{\mathbf{u}}_1') = \mathbf{0}$, $E(\tilde{\mathbf{u}}_1\tilde{\mathbf{u}}_2') = \mathbf{0}$, e si è definita la matrice \mathbf{V} . Ma il vettore aleatorio $\tilde{\mathbf{e}}$ è funzione lineare di $\tilde{\mathbf{u}}_2$ e $\hat{\mathbf{b}}$ che sono vettori normali multivariati ed indipendenti stocasticamente tra di loro poiché $\tilde{\mathbf{u}}_2$ è indipendente da $\tilde{\mathbf{u}}_1$, e quindi

$$\tilde{\mathbf{e}} \sim N(\mathbf{0}, \sigma^2\mathbf{V}) \quad (5.5.6)$$

Ora osserviamo che poiché $(\mathbf{X}'_1\mathbf{X}_1)^{-1}$ è una matrice definita positiva (e lo è poiché esiste la sua inversa $(\mathbf{X}'_1\mathbf{X}_1)$), per il teorema XIX-1.9 è definita positiva anche la matrice $\mathbf{X}_2(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_2$ e quindi anche la \mathbf{V} ; poiché questa è anche simmetrica, come facilmente si verifica, l'applicazione del teorema XIX-1.14 mostra che

$$\tilde{\mathbf{e}}'\mathbf{V}^{-1}\tilde{\mathbf{e}}/\sigma^2 \sim \chi_{n_2}^2 \quad (5.5.7)$$

per cui

$$\frac{\tilde{\mathbf{e}}'\mathbf{V}^{-1}\tilde{\mathbf{e}}/n_2}{\hat{\mathbf{u}}_1'\hat{\mathbf{u}}_1/(n_1-k)} = \tilde{F} \quad (5.5.8)$$

è il rapporto tra due variabili aleatorie distribuite come un chi quadrato con n_2 e (n_1-k) gradi di libertà rispettivamente. Se il rapporto (5.5.8) è piccolo anche gli errori di proiezione sono piccoli e si è indotti ad accettare l'ipotesi nulla di inesistenza di cambiamento strutturale. Però, per utilizzare la (5.5.8) come una statistica per il test, occorre trovarne la distribuzione, cosa che non è facile poiché non è possibile dimostrare in maniera diretta che numeratore e denominatore sono stocasticamente indipendenti. Il Chow [1960], tuttavia, ha dimostrato che rapporto (5.5.8) è uguale al (5.5.1), che abbiamo visto essere distribuito come una F di Fisher con n_2 ed n_1-k gradi di libertà. Se non vi è cambiamento strutturale gli errori di proiezione sono piccoli e le variabili aleatorie (5.5.7) ed (5.5.8) assumono valori ugualmente piccoli. Viceversa, se con il test della F si è portati a ritenere il numeratore del rapporto (5.5.8) non significativamente diverso da zero, si è anche indotti ad accettare l'ipotesi nulla di inesistenza di cambiamento strutturale.

Il test preliminare di uguaglianza delle varianze

Anche in questo caso, come in quello della verifica delle ipotesi (3.5.1), si presuppone che le varianze σ_1^2 e σ_2^2 dei residui siano uguali nei due sottoperiodi. Qualora non si fosse certi di questa uguaglianza si dovrebbe verificare preliminarmente l'ipotesi ulteriore (3.5.6), cosa che tuttavia non si può fare con la statistica (3.5.7) poiché in questo caso è $n_2 \leq k$ e non è possibile costruire uno stimatore della varianza dei residui nel secondo sottoperiodo. Al posto di questo stimatore, tuttavia, si può utilizzare la devianza di proiezione $\tilde{e}'\tilde{e}$, costruendo la statistica

$$\frac{\tilde{e}'\tilde{e}}{\hat{u}'_1\hat{u}_1/(n_1 - k)} \quad (5.5.9)$$

della quale però non si conosce la distribuzione poiché numeratore e denominatore non sono indipendenti. Tuttavia, in virtù del fatto che si dimostra che per $n_1 \rightarrow \infty$ il denominatore tende a σ^2 sotto l'ipotesi definita dalla (3.5.6), il rapporto (5.5.9) si distribuisce approssimativamente come un χ^2 con n_2 gradi di libertà e quindi tale statistica può essere usata, per n_1 sufficientemente grande, al fine di una verifica preliminare dell'uguaglianza delle varianze nei due sottoperiodi.

L'ipotesi alternativa è ancora del tipo (3.5.8) per cui il test del χ^2 è bilaterale, analogo a quello esposto alla fine del paragrafo 2.3. Per n_1 grande si calcola il rapporto (5.5.9) che, sulla falsariga del rapporto (2.3.22), viene confrontato con l'intervallo di accettazione dell'ipotesi nulla H_0 consistente in $\sigma_1^2 = \sigma_2^2$. Tale intervallo di accettazione è $[\chi_{n_2}^2, \chi_{n_2}^{\prime 2}]$ dove i quantili sono di probabilità $\alpha/2$ e $1-\alpha/2$ rispettivamente.

Applicazioni alla funzione delle importazioni

Nel precedente paragrafo 5.4 abbiamo constatato, osservando i grafici delle stime ricorsive e in particolare la figura 5.5, che riporta lo scarto quadratico medio della regressione, che molto probabilmente l'equazione (3.3.29) è soggetta a più di un cambiamento di struttura, l'ultimo dei quali appare localizzato all'inizio del 1985 (data dell'ultima "impennata" nel grafico dello scarto quadratico medio della regressione).

Se imputiamo il fallimento delle proiezioni alla presenza di cambiamento di struttura, viene quindi spontaneo ristimare l'equazione sull'ultimo sottoperiodo "stabile", ovvero scartando tutti i dati fino al 1984 compreso. Il campione si riduce quindi alle $n = 20$ osservazioni dal 1985 al 1989. Per verificare l'ipotesi di assenza di cambiamento di struttura possiamo suddividere ulteriormente questo campione nelle prime $n_1 = 16$ osservazioni dal 1985 al 1988, e utilizzare per la verifica il sottocampione di proiezione *ex post* comprendente le $n_2 = 4$ osservazioni relative al

1989. In questo caso abbiamo $n_2 < k$ (dato che i parametri da stimare sono 8)⁵ e quindi possiamo applicare i test di cambiamento di struttura presentati in questo paragrafo.

La stima della (3.3.29) sul sottocampione dal 1985 al 1988 fornisce questi risultati

$$\begin{aligned} \ln \hat{y}_t = & -8.45 & + 0.72 \ln x_{1t} & + 0.96 \ln x_{2t} & - 0.01 \ln x_{3t} & + 0.03 \ln x_{4t} + \\ & (-0.6) & (0.5) & (3.0) & (-0.1) & (0.0) \\ & - 0.003 d_{1t} & - 0.04 d_{2t} & - 0.02 d_{3t} & & \\ & (-0.7) & (-1.3) & (-0.6) & & \end{aligned} \quad (5.5.10)$$

$n = 32, R^2_c = 0.976, RSS = 0.0019, SEE = 0.015, JB = 0.18, BP = 0.0$

Notiamo che queste stime palesano alcune tendenze emerse dall'analisi dei grafici dal 5.6 al 5.13. Ad esempio, l'elasticità ai consumi è ora più piccola di quella a investimenti e esportazioni rispetto a quanto si verifica nelle stime (5.1.11) condotte su un campione più ampio, e inoltre non è statisticamente significativa (un risultato che può essere anticipato osservando le figure 5.7 e 5.8), mentre tutte le variabili di comodo hanno coefficienti più piccoli di quelli riscontrati nelle stime su campioni più ampi (anche questo risultato è in sintonia con i grafici delle stime ricorsive 5.11-5.13). Nonostante i test diagnostici non segnalino violazioni delle ipotesi stocastiche di base, e nonostante le variabili di prezzo abbiano il segno atteso, tuttavia nella (5.5.10) sussistono problemi di specificazione, evidenziati dal fatto che sette variabili su otto non sono statisticamente significative.

La statistica $F_{4,8}$ costruita con le (5.5.1)-(5.5.8) fornisce un valore pari a 2.30, che va confrontato con un valore soglia pari a 3.84. L'ipotesi di cambiamento di struttura viene quindi respinta. Tuttavia in questo caso il risultato non è motivo di particolare soddisfazione, visto che la struttura che non cambia passando dal primo al secondo sottoperiodo non è una "buona" struttura, ma anzi, come abbiamo appena rilevato, una pessima struttura, nella quale la maggior parte dei coefficienti risultano statisticamente non significativi. Questa osservazione può sembrare superflua, ma occorre sempre ricordare che il mancato rifiuto dell'ipotesi nulla in un test di assenza di cambiamento strutturale può indicare sia che l'equazione è ugualmente buona nei due sottoperiodi, sia che è ugualmente cattiva. Questo sembra essere appunto il caso della (5.5.10).

La tavola 5.4, analoga alla 5.3, riporta gli scarti quadratici medi dell'errore di proiezione e le relative misure di dispersione e test di significatività. In questo caso

⁵ Naturalmente se il campione di stima parte dal 1985 si omette dal modello la *dummy* puntuale relativa al primo trimestre del 1973, che invece figura nelle stime (3.3.29).

il valore soglia del test (5.2.2) è quello della t_8 , che al 5% è pari a 2.3. Nessun errore di proiezione risulta statisticamente significativo, ovvero nessun valore storico giace al di fuori dell'intervallo di confidenza della proiezione, gli errori di proiezione non hanno tutti lo stesso segno, e inoltre gli errori studentizzati sono relativamente più piccoli all'inizio rispetto alla fine del campione di proiezione *ex post* (cioè il modello prevede meglio il futuro prossimo di quello remoto, in accordo con l'intuizione e contrariamente a quanto si riscontrava ad esempio per l'equazione (5.1.11)). Anche in termini di dimensioni gli errori ora sono molto più ridotto, andando da un minimo dello 0.7% a un massimo del 5%. In questo senso, pur con le riserve espresse finora, possiamo affermare che, sulla base dei risultati della proiezione *ex post*, la (5.5.10) dovrebbe essere considerata, ai fini della proiezione *ex ante*, più affidabile delle altre equazioni presentate nelle pagine precedenti.

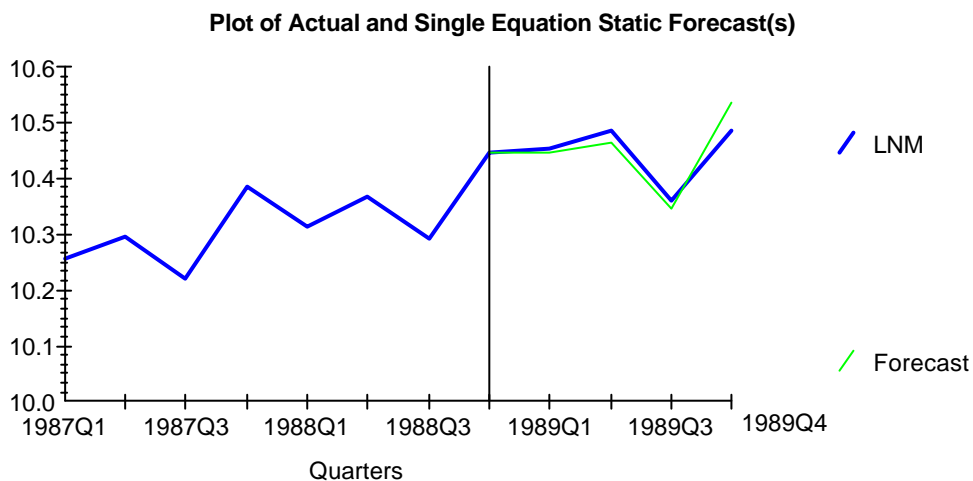


Figura 5.14 – I valori storici del logaritmo delle importazioni e quelli proiettati con la (5.5.10).

| | storici | proiettati | errore | s.q.m. dell'errore | test t |
|--------|---------|------------|---------|--------------------|---------|
| 1989Q1 | 10.4544 | 10.4473 | 0.0071 | 0.0238 | 0.2974 |
| 1989Q2 | 10.4840 | 10.4656 | 0.0184 | 0.0272 | 0.6749 |
| 1989Q3 | 10.3607 | 10.3450 | 0.0157 | 0.0293 | 0.5372 |
| 1989Q4 | 10.4860 | 10.5368 | -0.0508 | 0.0261 | -1.9482 |

Tavola 5.4 – Valori storici e proiettati, errori di proiezione e loro test di significatività relativi alla proiezione del logaritmo delle importazioni con la (5.5.10).

5.6 Verifica della stabilità dei parametri

Nei test di cambiamento strutturale esposti nel capitolo precedente e nel paragrafo 5.5 è stato ipotizzato di conoscere il tempo (tra n_1 ed n_1+1) in cui questo cambiamento si verifica. Supponiamo ora di non conoscerlo e di voler controllare l'ipotesi nulla che in uno o più istanti del campione possano esserci dei cambiamenti di struttura, in modo tale che il vettore di parametri \mathbf{b} non possa essere considerato costante in tutto il campione. La verifica di questa costanza equivale, dunque, alla verifica della *stabilità dei parametri* nel periodo campionario e si avvale dei residui ricorsivi esposti in precedenza.

Se \mathbf{b} non è stabile si può supporre che i residui ricorsivi (5.4.6) abbiano valor medio non nullo μ per cui

$$\tilde{v}_t \sim N(\mu, \sigma^2) \quad (5.6.1)$$

e si può considerare come stimatore di μ il valor medio campionario

$$\bar{v} = \frac{1}{n-k} \sum_{t=k+1}^n \tilde{v}_t \quad (5.6.2)$$

che ha distribuzione $N[\mu, \sigma^2/(n-k)]$ come facilmente si dimostra partendo dal fatto che esso è dato dalla somma di $n-k$ variabili distribuite come nella (5.6.1). Se per σ^2 si considera lo stimatore non distorto

$$\bar{\sigma}^2 = \frac{1}{(n-k)-1} \sum_{t=k+1}^n (\tilde{v}_t - \bar{v})^2 \quad (5.6.3)$$

si ha che

$$\frac{(\bar{v} - \mu) / [\sigma^2 / (n-k)]^{1/2}}{[\bar{\sigma}^2 / \sigma^2]^{1/2}} \sim t_{n-k-1}$$

dalla quale si ottiene la statistica

$$(n-k)^{1/2} \frac{\bar{v} - \mu}{\bar{\sigma}} \sim t_{n-k-1} \quad (5.6.4)$$

che può essere utilizzata per verificare l'ipotesi nulla di stabilità dei parametri corrispondente alla

$$H_0: \mu=0 \quad (5.6.5)$$

Si tende, dunque, ad accettare questa ipotesi se il valore $(n-k)^{1/2} \bar{v} / \bar{\sigma}$ è compreso nell'intervallo $[t', t'']$ dove t' e t'' sono gli estremi determinati con l'ausilio delle tavole statistiche della t di Student con $n-k-1$ gradi di libertà ad un prestabilito livello di confidenza α .

Osservazione 5.1 - In effetti, poiché i residui ricorsivi partono da $k+1$, la stabilità dei parametri è verificata non nell'intero periodo campionario bensì nel sottoperiodo che va da $k+1$ ad n .

Harvey e Collier [1977] interpretano questo test in termini più ampi come strumento diagnostico per la verifica della linearità della funzione di regressione (5.1.1). Infatti se anche il vettore di parametri \mathbf{b} è stabile ma l'ipotesi di linearità di y_t rispetto ad una variabile esplicativa x_{it} non è corretta, accade che la statistica $(n-k)^{1/2} \bar{v} / \bar{\sigma}$ tende ad essere grande e quindi ad uscire fuori dell'intervallo $[t', t'']$.

Osservazione 5.2 - In effetti, poiché

$$a_t v_t = (y_t - \mathbf{x}'_t \hat{\mathbf{b}}_{t-1}) = \mathbf{x}'_t (\mathbf{b}_t - \hat{\mathbf{b}}_{t-1}) + u_t$$

l'ipotesi nulla (5.6.5) è violata sia quando è $\mathbf{b}_t \neq \hat{\mathbf{b}}_{t-1}$ (parametri instabili) sia quando una o più variabili esplicative nel vettore \mathbf{x}_t non sono specificate appropriatamente in forma lineare anche se la differenza $\mathbf{b}_t - \hat{\mathbf{b}}_{t-1}$ è piccola. In questo secondo caso, tuttavia, può accadere che le non linearità di più variabili si compensino ed il test non sia in grado di identificarle. Per questo motivo è opportuno interpretarlo come test di non linearità soltanto quando si è certi che questa riguardi una sola variabile esplicativa.

Osservazione 5.3 - Il fatto che alla medesima statistica possano essere attribuiti significati diversi evidenzia come l'interpretazione dei test diagnostici di malaspecificazione sia caratterizzata da una fondamentale asimmetria, nel senso che *il rifiuto della nulla non implica l'accettazione dell'alternativa*. Ad esempio, la statistica (5.6.4) risulterà significativa (portando al rifiuto della nulla) sia in caso di non costanza dei parametri che in caso di linearità del modello, e il test di per sé non fornisce indizi circa la fonte più probabile di malaspecificazione. Da ciò scaturiscono due indicazioni operative di validità generale: la prima è che è in generale errato interpretare meccanicisticamente il rifiuto dell'ipotesi nulla di un test come accettazione di una particolare alternativa (poiché il medesimo test potrebbe essere sensibile anche rispetto a violazioni di ipotesi di base non previste nella specifica alternativa considerata); la seconda è che la diagnostica del modello deve basarsi su un insieme ampio di strumenti diagnostici, perché solo una considerazione complessiva delle sue proprietà statistiche può consentire di individuare in modo sufficientemente affidabile e costruttivo gli eventuali spazi per un miglioramento del modello stesso.

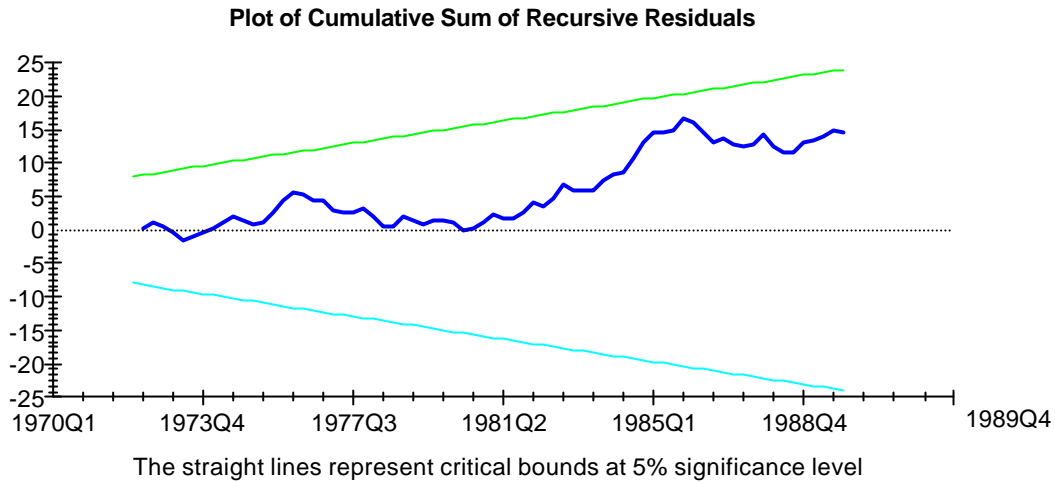


Figura 5.15 – Il test CUSUM per l'equazione (3.3.29). La statistica giace all'interno delle due rette che delimitano il valore soglia e l'ipotesi di stabilità dei parametri non viene respinta.

I test CUSUM e CUSUMSQ

Brown, Durbin ed Evans [1975] hanno suggerito di utilizzare altri due test per verificare la stabilità dei parametri. Il primo è basato sulla statistica

$$\tilde{w}_t = \sum_{i=k+1}^t \tilde{v}_i / \bar{\sigma} \quad t=k+1, k+2, \dots, n \quad (5.6.6)$$

dove $\bar{\sigma}^2$ è lo stimatore non distorto (1.7.2) calcolato sull'intero periodo campionario. All'aumentare di t la \tilde{w}_t stessa è detta *somma cumulata* da cui il nome CUSUM (CUmulative SUM in lingua inglese) dato al test. I tre autori hanno dimostrato che se vale l'ipotesi nulla di stabilità dei parametri la statistica (5.6.6) ha approssimativamente distribuzione $N(0, t-k)$ e che la regione di accettazione consiste nell'intervallo $[-z_{\alpha,t}, z_{\alpha,t}]$ dove

$$z_{\alpha,t} = a(n-k)^{1/2} + 2a(t-k)(n-k)^{1/2} \quad (5.6.7)$$

è una retta in t e il fattore a dipende dal livello di significatività α del test (vedi la tavola 5.5).

La verifica è eseguita riportando in un grafico le serie di valori w_t insieme alla due rette $[-z_{\alpha,t}, z_{\alpha,t}]$ definite dalla (5.6.7) e verificando che la serie w_t risulti tutta compresa nella regione definita dalle due rette oppure venga a trovarsi al di fuori di essa; nel primo caso si accetta l'ipotesi nulla di stabilità dei parametri mentre nel secondo la si rifiuta. Nel caso dell'equazione (3.3.29) la statistica CUSUM è riportata, insieme ai valori soglia al 5%, nella figura 5.15. La spezzata w_t giace all'interno della regione delimitata dalle due rette, e quindi, secondo il test CUSUM, i coefficienti della (3.3.29) sono costanti. In realtà abbiamo già raccolto

numerose e convincenti evidenze statistiche del contrario (a partire dal test F (3.5.12), fino ad arrivare ai grafici delle stime ricorsive riportate nel paragrafo 4 di questo capitolo). Il risultato della figura 5.15 vale quindi a evidenziare un limite del test $CUSUM$ attestato in letteratura, vale a dire la sua scarsa potenza, ovvero il fatto che ha una probabilità relativamente ridotta di rifiutare l'ipotesi nulla di costanza dei parametri quando è falsa. Tuttavia il grafico presenta due "impenne" che, se pure non statisticamente significative, vanno considerate come un segnale di allarme rispetto alla probabile presenza di cambiamenti di struttura (e in effetti sono localizzate nei pressi dei due episodi di cambiamento di struttura individuati in precedenza in corrispondenza del 1976 e del 1985). La consultazione del grafico del test quindi mantiene una sua efficacia diagnostica.

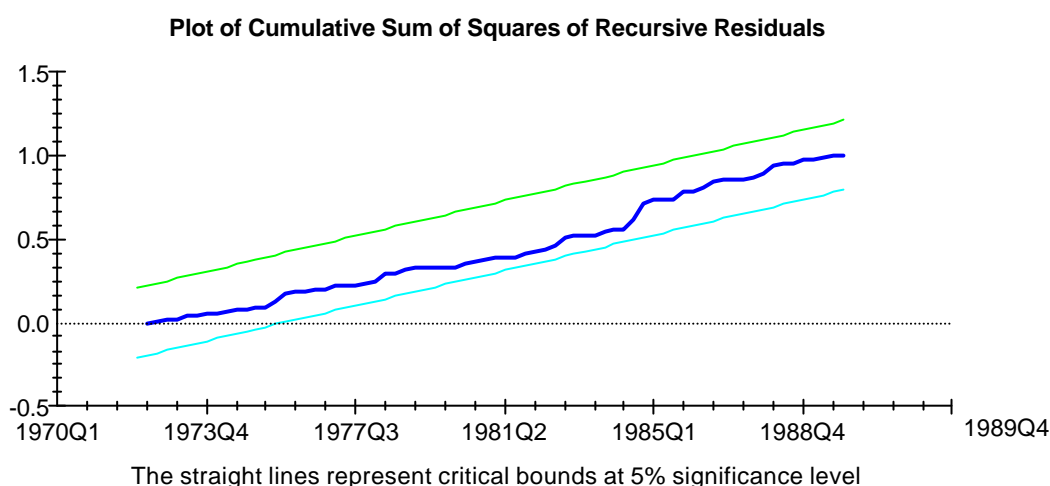


Figura 5.16 - Regione di accettazione nel test $CUSUMSQ$. Se la serie dei valori della statistica è compresa tra i due segmenti AB e $A'B'$ viene accettata l'ipotesi nulla di stabilità dei parametri, altrimenti la si rifiuta.

| Tavola 5.5 – Il fattore a del test $CUSUM$ di Brown, Durbin e Evans [1975] | | | | |
|--|--|---|-------|-------|
| | | <i>livello di significatività α</i> | | |
| | | 0.01 | 0.05 | 0.10 |
| <i>fattore a</i> | | 1.143 | 0.948 | 0.850 |

Considerazioni analoghe valgono per il secondo test costruito dai tre autori, che si basa sulla statistica

$$\tilde{s}_t = \frac{\sum_{i=k+1}^t \tilde{v}_i^2}{\sum_{i=k+1}^n \tilde{v}_i^2} \quad t=k+1, k+2, \dots, n \quad (5.6.8)$$

composta dai residui ricorsivi al quadrato cumulati, da cui il nome CUSUMSQ (CUmulative SUM of SQuares in lingua inglese), e normalizzati tramite la somma effettuata su tutto il campione. Tale statistica ha valor medio

$$E(\tilde{s}_t) = \frac{t-k}{n-k}$$

per cui sotto la validità dell'ipotesi nulla s_t non si discosta troppo dalla retta dei valori medi che vale zero per $t=k$ ed uno per $t=n$. Una regione di accettazione dell'ipotesi nulla può essere costruita con asse mediano dato da tale retta e con delimitazione fornita da questa stessa retta aumentata e diminuita di una fascia larga un valore c fornito da apposite tavole statistiche, dipendente dall'ampiezza del campione e dal livello di significatività come illustrato nella figura 5.16.

Anche il test CUSUMSQ non segnala violazioni dell'ipotesi nulla di stabilità dei parametri della (3.3.29).

5.7 L'impiego delle variabili di comodo nella proiezione

Alcuni calcoli esposti nei precedenti paragrafi di questo capitolo possono essere notevolmente semplificati mediante l'uso delle variabili di comodo, che oltre alla loro comodità sotto il profilo operativo forniscono anche, in certi casi, utili intuizioni circa il significato delle statistiche calcolate.

L'errore quadratico medio di proiezione e l'intervallo di confidenza della proiezione

Per calcolare l'errore quadratico medio di proiezione (5.1.6) occorre effettuare dei calcoli matriciali che coinvolgono la matrice di disegno del modello \mathbf{X} e il vettore delle variabili esplicative al tempo $n+h$. Questi calcoli vengono generalmente effettuati dai software econometrici. Può capitare, tuttavia, che il software che si sta utilizzando, pur effettuando le stime *OLS*, non calcoli l'errore quadratico medio di proiezione. In questo caso è utile ricorrere a un metodo di calcolo proposto originariamente da Salkever [1976] e ripreso da Pagan e Nicholls [1984].

Sia $\hat{\mathbf{y}}^* = [\hat{y}_{n+1}, \dots, \hat{y}_{n+m}]'$ il vettore delle m proiezioni e sia \mathbf{X}^* la matrice $m \times k$ che raccoglie gli m vettori delle k variabili esplicative relativi agli m tempi sui quali si effettua la proiezione

$$\mathbf{X}^* = \begin{bmatrix} \mathbf{x}'_{n+1} \\ \vdots \\ \mathbf{x}'_{n+m} \end{bmatrix} \quad (5.7.1)$$

Definiamo inoltre le m variabili di comodo $d_{h,t}^*$

$$d_{h,t}^* = \begin{cases} -1 & t = n+h \\ 0 & t \neq n+h \end{cases} \quad h = 1, \dots, m \quad (5.7.2)$$

raccolte in una matrice \mathbf{D} di ordine $(n+m) \times m$ così definita

$$\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m) = \begin{pmatrix} \mathbf{0} \\ -\mathbf{I}_m \end{pmatrix} \quad (5.7.3)$$

dove $\mathbf{0}$ è una matrice di zeri di ordine $n \times m$.

Consideriamo ora l'equazione costruita in questo modo: si aggiungono alle n osservazioni campionarie sulla variabile dipendente m zeri (uno per ogni data del campione di previsione), alle n osservazioni sulla i -esima esplicativa x_{it} le corrispondenti m osservazioni sul campione di previsione $x_{i,n+h}$ ($h = 1, \dots, m$), e si aumenta il modello con le m variabili di comodo (5.7.2). In termini matriciali l'equazione risultante è

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{X}^* & -\mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ \mathbf{d} \end{pmatrix} + \begin{pmatrix} \mathbf{u} \\ \mathbf{e}^* \end{pmatrix} \quad (5.7.4)$$

Nella (5.7.4) i \mathbf{d} sono gli m coefficienti delle variabili di comodo $d_{h,t}^*$. Si verifica immediatamente che le stime OLS dei \mathbf{d} coincidono con le m proiezioni OLS ai tempi $n+1, \dots, n+m$, e che le varianze di queste stime coincidono con gli errori quadratici medi delle proiezioni.

Per dimostrarlo, si osservi che per definizione le stime OLS dei parametri della (5.7.4) minimizzano la devianza dei residui, che in questo caso è

$$\sum_{t=1}^n u_t^2 + \sum_{t=n+1}^{n+m} (\varepsilon_t^*)^2 = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) + (-\mathbf{X}^*\mathbf{b} + \mathbf{d})'(-\mathbf{X}^*\mathbf{b} + \mathbf{d}) \quad (5.7.5)$$

Il primo addendo del quale consta la devianza (5.7.5) viene minimizzato, come è noto, dalla stima OLS $\hat{\mathbf{b}}$, mentre il secondo addendo raggiunge un minimo assoluto pari a zero quando si ponga $\mathbf{d} = \mathbf{X}^*\mathbf{b}$. Queste due condizioni implicano congiuntamente che sia $\hat{\mathbf{d}} = \mathbf{X}^*\hat{\mathbf{b}}$, per cui il coefficiente $\hat{\delta}_h$ della h -esima variabile di comodo $d_{h,t}^*$ è uguale a

$$\hat{\delta}_h = \mathbf{x}'_{n+h} \hat{\mathbf{b}} = \hat{y}_{n+h} \quad h = 1, \dots, m \quad (5.7.6)$$

cioè alla proiezione della variabile dipendente al tempo $n+h$.

Se chiamiamo \mathbf{Z} la matrice di disegno della (5.7.4)

$$\mathbf{Z} = \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{X}^* & -\mathbf{I}_m \end{pmatrix}$$

la matrice di varianze e covarianze delle stime OLS $[\hat{\mathbf{b}}', \hat{\mathbf{d}}']'$ è pari a $\sigma^2(\mathbf{Z}'\mathbf{Z})^{-1}$, per cui, in particolare, la matrice di varianze e covarianze dei soli $\hat{\mathbf{d}}$ sarà data dal blocco diagonale inferiore di ordine $m \times m$ della $\sigma^2(\mathbf{Z}'\mathbf{Z})^{-1}$. Per ricavarlo applichiamo la formula dell'inversa partizionata XIX-(1.4.8) alla matrice $\mathbf{Z}'\mathbf{Z}$

$$\mathbf{Z}'\mathbf{Z} = \begin{pmatrix} \mathbf{X}' & \mathbf{X}^{*'} \\ \mathbf{0}' & -\mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{X}^* & -\mathbf{I}_m \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{X} + \mathbf{X}^{*'}\mathbf{X}^* & -\mathbf{X}^{*'} \\ -\mathbf{X}^* & \mathbf{I}_m \end{pmatrix} \quad (5.7.7)$$

dove, nella notazione del capitolo XIX-1, la matrice $\mathbf{Z}'\mathbf{Z}$ corrisponde alla matrice \mathbf{A} e i blocchi sono $\mathbf{A}_{11} = \mathbf{X}'\mathbf{X} + \mathbf{X}^{*'}\mathbf{X}^*$, $\mathbf{A}_{12} = -\mathbf{X}^{*}'$, $\mathbf{A}_{21} = \mathbf{X}^*$ e $\mathbf{A}_{22} = \mathbf{I}_m$, per cui l'applicazione delle XIX-(1.4.8) e XIX-(1.4.9) mostra immediatamente che

$$\text{Var}(\hat{\mathbf{d}}) = \sigma^2[\mathbf{I}_m + \mathbf{X}^*(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}^{*'}] \quad (5.7.8)$$

e in particolare la varianza dell' h -esimo coefficiente $\hat{\delta}_h$ è data dall' h -esimo elemento diagonale della (5.7.8), pari a

$$\text{Var}(\hat{\delta}_h) = \sigma^2 [1 + \mathbf{x}'_{n+h} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_{n+h}] \quad (5.7.9)$$

il quale altro non è che l'errore quadratico medio di proiezione al tempo $n+h$ dato dalla (5.1.6).

Il test di significatività dell'errore di proiezione

Il ragionamento svolto al punto precedente può essere applicato anche *ex post*, con la differenza che in questo caso sono disponibili i valori storici della variabile dipendente sull'orizzonte di proiezione. La (5.7.4) diventa quindi, riprendendo la notazione del paragrafo 3.5,

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{X}_2 & -\mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{d} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{e}^* \end{pmatrix} \quad (5.7.10)$$

per cui le variabili di comodo $d_{h,t}^*$ ora sono in numero di n_2 (una per ogni elemento del sottocampione di proiezione *ex post*) e sono così costruite

$$d_{h,t}^* = \begin{cases} -1 & t = n_1 + h \\ 0 & t \neq n_1 + h \end{cases} \quad h = 1, \dots, n_2 \quad (5.7.11)$$

La devianza dei residui della (5.7.10) è

$$\begin{aligned} \sum_{t=1}^{n_1} u_t^2 + \sum_{t=n_1+1}^n (\varepsilon_t^*)^2 &= \\ &= (\mathbf{y}_1 - \mathbf{X}_1 \mathbf{b}_1)' (\mathbf{y}_1 - \mathbf{X}_1 \mathbf{b}_1) + (\mathbf{y}_2 - \mathbf{X}_2 \mathbf{b}_1 + \mathbf{d})' (\mathbf{y}_2 - \mathbf{X}_2 \mathbf{b}_1 + \mathbf{d}) \end{aligned} \quad (5.7.12)$$

e applicando il medesimo ragionamento sviluppato in precedenza si vede che il primo addendo è minimizzato dallo stimatore *OLS* applicato sul sottocampione di stima che consta di n_1 osservazioni, cioè da $\hat{\mathbf{b}}_1 = (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{y}_1$, mentre il secondo raggiunge il proprio minimo assoluto, pari a zero, ponendo $\mathbf{d} = \mathbf{X}_2 \mathbf{b}_1 - \mathbf{y}_2$. Queste due condizioni implicano congiuntamente che sia $\hat{\mathbf{d}} = \mathbf{X}_2 \hat{\mathbf{b}}_1 - \mathbf{y}_2$, per cui in questo caso l' h -esimo coefficiente stimato $\hat{\delta}_h$ è dato da

$$\hat{\delta}_h = \mathbf{x}'_{n_1+h} \hat{\mathbf{b}}_1 - y_{n_1+h} = \hat{y}_{n_1+h} - y_{n_1+h} = -e_{n_1+h} \quad h = 1, \dots, n_2 \quad (5.7.13)$$

cioè dal negativo dell' h -esimo errore di proiezione *ex post*.

Il modello (5.7.10) ha, tenuto conto dei cambiamenti di notazione, la stessa matrice di disegno del (5.7.4). Dato che la matrice di dispersione delle stime dipende solo dalla matrice di disegno, e non dal vettore della variabile dipendente, ne consegue che le stime $\hat{\mathbf{d}}$ della (5.7.10) avranno matrice di dispersione analoga alla (5.7.8), ovvero

$$\text{Var}(\hat{\mathbf{d}}) = \sigma^2 [\mathbf{I}_{n_2} + \mathbf{X}_2 (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_2'] \quad (5.7.14)$$

Di conseguenza, le t di Student dei coefficienti $\hat{\delta}_h$ nella (5.7.10) coincidono con le statistiche t definite dalla (5.2.2) cambiate di segno. Dato che il test t è bilaterale con distribuzione simmetrica, il cambiamento di segno non ne altera minimamente le conclusioni. Se ne conclude quindi che le t dei $\hat{\delta}_h$ possono essere impiegate per verificare l'ipotesi che l' h -esimo errore di proiezione *ex post* ($h = 1, \dots, n_2$) sia statisticamente significativo. Operativamente, t di Student maggiori di due in valore assoluto indicano un fallimento della proiezione *ex post*, che potrebbe essere determinato, in particolare, dalla presenza di un cambiamento di struttura nei coefficienti (ma si vedano anche le considerazioni svolte nelle osservazioni 5.2 e 5.3 di questo capitolo).

Un'interpretazione alternativa del test del Chow

Se le t di Student dei singoli coefficienti $\hat{\delta}_h$ della (5.7.10) forniscono test di accuratezza delle singole proiezioni *ex post* ai tempi n_1+h , viene naturale definire un test congiunto di accuratezza della proiezione *ex post* su tutto il campione da n_1+1 a $n_1+n_2 = n$ considerando il test F per la verifica dell'ipotesi che i δ_h siano congiuntamente nulli. Questo test viene talora definito nei testi in lingua inglese *predictive failure test*, cioè test di fallimento della proiezione (anche se forse sarebbe preferibile anche in questo caso denominare il test facendo riferimento all'ipotesi nulla, che è appunto quella di accuratezza della proiezione, cioè di nullità congiunta degli errori di proiezione *ex post*).

Dufour [1980] dimostra che questo test F coincide con il test del Chow presentato nel par. 5.5. Questa interpretazione di per sé non è particolarmente originale, in quanto, come abbiamo visto, lo stesso Chow interpretava la statistica (5.5.1)-(5.5.8) come test di accuratezza delle proiezioni. Tuttavia la dimostrazione di Dufour [1980], basata sull'impiego delle variabili di comodo, è più semplice e intuitiva di quelle precedentemente fornite da Chow [1960] e Harvey [1976].

In effetti, la statistica F per l'ipotesi di nullità dei coefficienti \mathbf{d} della (5.7.10) può essere costruita utilizzando la consueta formula (2.5.4)

$$\frac{\mathbf{u}_0' \mathbf{u}_0 - \hat{\mathbf{u}}' \hat{\mathbf{u}}}{\hat{\mathbf{u}}' \hat{\mathbf{u}}} \frac{(n-k)}{q} \sim F_{q, n-k} \quad (5.7.15)$$

dove $\mathbf{u}_0' \mathbf{u}_0$ è la devianza vincolata, cioè quella calcolata subordinatamente all'ipotesi nulla $H_0: \mathbf{d} = \mathbf{0}$, $\hat{\mathbf{u}}' \hat{\mathbf{u}}$ è la devianza libera, cioè stimata in assenza di vincoli, data dalla (5.7.12), q è il numero di vincoli, che nel nostro caso è pari a n_2 (perché tanti sono i coefficienti \mathbf{d}), e $n-k$ è il numero di gradi di libertà del modello, pari nel nostro caso a $n - n_2 - k$, dato che il modello (5.7.10) ha $k+n_2$ parametri.

Tuttavia, nel caso in esame la devianza vincolata $\mathbf{u}_0' \mathbf{u}_0$ coincide con la $\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0$ del test del Chow, dato che per $\mathbf{d} = \mathbf{0}$ la (5.7.10) diventa analoga alla (3.5.4)

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{X}_2 & -\mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{e}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \mathbf{b}_1 + \mathbf{u}$$

e quindi la devianza dei suoi residui coincide con quella stimata utilizzando tutte le n_1+n_2 osservazioni subordinatamente all'ipotesi di assenza di cambiamento strutturale, per cui si ha $\mathbf{u}_0' \mathbf{u}_0 = \hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0$. D'altra parte abbiamo già visto che la devianza libera $\hat{\mathbf{u}}' \hat{\mathbf{u}}$ per la (5.7.12) consta di due addendi, il secondo dei quali può essere annullato ponendo $\mathbf{d} = \mathbf{X}_2 \mathbf{b}_1 - \mathbf{y}_2$, mentre il primo coincide con la devianza dei residui relativa al primo sottocampione: $\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1 = (\mathbf{y}_1 - \mathbf{X}_1 \hat{\mathbf{b}}_1)' (\mathbf{y}_1 - \mathbf{X}_1 \hat{\mathbf{b}}_1)$, per cui si ha $\hat{\mathbf{u}}' \hat{\mathbf{u}} = \hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1$. In definitiva quindi, considerando che $n - n_2 - k = n_1 - k$, la statistica F per la verifica dell'ipotesi $H_0: \mathbf{d} = \mathbf{0}$ nella (5.7.10) è pari a

$$\frac{(\hat{\mathbf{u}}_0' \hat{\mathbf{u}}_0 - \hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1) / n_2}{\hat{\mathbf{u}}_1' \hat{\mathbf{u}}_1 / (n_1 - k)} \quad (5.7.16)$$

cioè coincide con la (5.5.1). Se si segue questo procedimento la (5.5.1) viene ricondotta a un caso particolare del test F esposto nel capitolo 3 e quindi per dimostrare l'indipendenza fra numeratore e denominatore non è necessario ricorrere, come fa Harvey [1976], alla nozione di devianza ricorsiva, ma basta applicare, come nel caso generale, il noto teorema sull'indipendenza fra forme quadratiche XIX-1.16.

Un'applicazione: effettuiamo il test del Chow con EasyReg

La disponibilità attuale di software econometrico potente e di facile utilizzo rende generalmente superfluo ricorrere ai metodi basati sulle variabili di comodo per semplificare il calcolo delle statistiche di bontà della proiezione. Questi metodi però conservano una validità didattica, perché consentono di dimostrare in modo semplice e intuitivo risultati come quello relativo alla distribuzione della statistica (5.7.16), e tornano comunque comodi in certi contesti specifici. Ad esempio, mentre software come Microfit e PcGive dedicano ampio spazio alla diagnostica di stabilità dei parametri, con l'implementazione di test del Chow e metodi ricorsivi, il freeware EasyReg non implementa il test del Chow. Vediamo allora come impostare il calcolo della relativa statistica senza calcolare "manualmente" il rapporto (5.5.1) e utilizzando invece il metodo delle variabili di comodo. Per esemplificare questo metodo di calcolo replichiamo i risultati relativi all'equazione (5.5.10) riportati nel paragrafo 5 di questo capitolo.

Dato che la proiezione *ex post* viene effettuata su un sottocampione di $n_2 = 4$ osservazioni, dobbiamo definire quattro variabili di comodo $d_{h,t}^*$ che valgono uno rispettivamente nel primo, secondo, terzo e quarto trimestre del 1989, e zero altrove. Ci serviamo a questo scopo del comando `Menù>Input>Transform variable` selezionando l'opzione `Dummy=1 for a subset of observations`. Le quattro variabili di

comodo definite sono riportate nella tavola 5.6, dove abbiamo omissis le osservazioni centrali, tutte ugualmente pari a zero. Si noti che nella 5.6 le variabili sono effettivamente state definite come $-d_{h,t}^*$, cioè come le *dummy* (5.7.2) cambiate di segno. In questo modo il loro coefficiente coincide con l'errore di proiezione non cambiato di segno.

| | $d_{1,t}^*$ | $d_{2,t}^*$ | $d_{3,t}^*$ | $d_{4,t}^*$ |
|--------|-------------|-------------|-------------|-------------|
| 1970.1 | 0 | 0 | 0 | 0 |
| 1970.2 | 0 | 0 | 0 | 0 |
| 1970.3 | 0 | 0 | 0 | 0 |
| 1970.4 | 0 | 0 | 0 | 0 |
| 1971.1 | 0 | 0 | 0 | 0 |
| ... | ... | ... | ... | ... |
| 1988.1 | 0 | 0 | 0 | 0 |
| 1988.2 | 0 | 0 | 0 | 0 |
| 1988.3 | 0 | 0 | 0 | 0 |
| 1988.4 | 0 | 0 | 0 | 0 |
| 1989.1 | 1 | 0 | 0 | 0 |
| 1989.2 | 0 | 1 | 0 | 0 |
| 1989.3 | 0 | 0 | 1 | 0 |
| 1989.4 | 0 | 0 | 0 | 1 |

Tavola 5.6 – Le variabili di comodo per l'effettuazione del test del Chow sulla (5.5.10).

La stima della (5.5.10) aumentata con le quattro variabili di comodo definite nella tavola 5.6 fornisce i risultati riportati nella tavola 5.7, che mostra direttamente l'output del programma di calcolo. Il tabulato, di facile interpretazione, riporta nella prima colonna i nomi delle variabili, nella seconda i coefficienti, nella terza le statistiche t , gli scarti quadratici medi dei coefficienti e i p -value delle t , e nella quarta le medesime informazioni della terza, ma calcolate con un diverso stimatore della varianza di $\hat{\mathbf{b}}$. Nel tabulato abbiamo evidenziato i coefficienti delle variabili $d1^*$, $d2^*$, $d3^*$ e $d4^*$, corrispondenti alle $d_{h,t}^*$. Si constata immediatamente che questi coefficienti corrispondono in effetti agli errori di proiezione *ex post* della (5.5.10), riportati nella tavola 5.4, mentre le relative t coincidono con le statistiche (5.2.2) di verifica della nullità dell'errore di proiezione, riportate anch'esse nella medesima tavola.

Per effettuare il test del Chow basta procedere fino al menù What to do next? di EasyReg e selezionare l'opzione Test parameter restrictions. I risultati del test di significatività congiunto sulle quattro variabili di comodo sono riportati nella tavola 5.8. La statistica F è pari a 2.31 e coincide quindi con quella del test del Chow calcolata nel paragrafo 5. Naturalmente le due statistiche, avendo il medesimo numero di gradi di libertà, portano alle stesse conclusioni, vale a dire che l'ipotesi di nullità congiunta dei coefficienti delle variabili di comodo, cioè

l'ipotesi che i quattro errori di proiezione *ex post* siano congiuntamente nulli, non viene respinta al livello di significatività del 5%. In altre parole, l'equazione appare strutturalmente stabile sul sottocampione di proiezione *ex post*.

| OLS estimation results for Y = LN[M] | | | |
|--------------------------------------|--------------|-----------------------------------|--|
| Variables | OLS estimate | t-value s.e. [p-value] | H.C. t-value(*) H.C. s.e. (*) [H.C. p-value] (*) |
| x(1) = LN[C] | .727278 | .584 1.2452E+00 [0.57528] | .813 8.9408E-01 [0.41597] |
| x(2) = LN[I] | .963265 | 3.095 3.1125E-01 [0.01478] | 4.730 2.0363E-01 [0.00000] |
| x(3) = LN[PM] | -.016768 | -.176 9.5501E-02 [0.86499] | -.295 5.6877E-02 [0.76814] |
| x(4) = LN[PY] | .030159 | .041 7.2832E-01 [0.96798] | .064 4.7320E-01 [0.94918] |
| x(5) = d1* | .007089 | .297 2.3833E-02 [0.77371] | .785 9.0253E-03 [0.43221] |
| x(6) = d2* | .018368 | .675 2.7215E-02 [0.51875] | 1.786 1.0284E-02 [0.07408] |
| x(7) = d3* | .015746 | .537 2.9313E-02 [0.60577] | 1.969 7.9977E-03 [0.04898] |
| x(8) = d4* | -.050845 | -1.948 2.6098E-02 [0.08723] | -5.442 9.3435E-03 [0.00000] |
| x(9) = Seasonal dummy 1 | -.003768 | -.071 5.3022E-02 [0.94509] | -.086 4.3704E-02 [0.93129] |
| x(10) = Seasonal dummy 2 | -.046929 | -1.353 3.4673E-02 [0.21289] | -1.790 2.6212E-02 [0.07340] |
| x(11) = Seasonal dummy 3 | -.028484 | -.625 4.5584E-02 [0.54946] | -.956 2.9809E-02 [0.33930] |
| x(12) = 1 | -8.456465 | -.664 1.2744E+01 [0.52564] | -.950 8.9016E+00 [0.34212] |

Tavola 5.7 - L'output del programma di calcolo EasyReg che presenta la stima dell'equazione (5.5.10) aumentata con le variabili di comodo per il calcolo della statistica del Chow.

| | | | |
|--------------------------|-----------|--------|-----------|
| x(1) = LN[C] | .727278 | .584 | .813 |
| x(2) = LN[I] | .963265 | 3.095 | 4.730 |
| x(3) = LN[PM] | -.016768 | -.176 | -.295 |
| x(4) = LN[PY] | .030159 | .041 | .064 |
| x(5) = d1* | .007089 | .297 | .785(*) |
| x(6) = d2* | .018368 | .675 | 1.786(*) |
| x(7) = d3* | .015746 | .537 | 1.969(*) |
| x(8) = d4* | -.050845 | -1.948 | -5.442(*) |
| x(9) = Seasonal dummy 1 | -.003768 | -.071 | -.086 |
| x(10) = Seasonal dummy 2 | -.046929 | -1.353 | -1.790 |
| x(11) = Seasonal dummy 3 | -.028484 | -.625 | -.956 |
| x(12) = 1 | -8.456465 | -.664 | -.950 |

Test of the null hypothesis that the parameters indicated by (*) are jointly zero:
 F-test statistic: 2.31 Null distribution: F(4,8)
 p-value = 0.14607
 Significance levels: 10% 5%
 Critical values: 2.81 3.84
 Conclusions: accept accept

Tavola 5.8 – I risultati del test F per l'ipotesi di significatività congiunta delle quattro variabili di comodo di tipo (5.7.2) nella (5.5.10). La statistica F coincide con quella del test del Chow calcolata secondo la (5.5.1), riportata nel paragrafo 5.

5.8 Bibliografia

- Breusch, T.S., Pagan, A.R. [1979], “A Simple Test for Heteroskedasticity and Random Coefficient Variation”, *Econometrica*, **47**, pp. 1287-1294.
- Brown, R.L., Durbin, J., Evans, J.M. [1975], “Techniques for testing the constancy of regression relationships over time”, *Journal of the Royal Statistical Society, Series B*, pp. 149-192.
- Chow, G.C. [1960], “Tests of equality between two sets of coefficients in two linear regressions”, *Econometrica*, **28**, pp. 591-605.
- de Finetti, B. [1970], *Teoria della probabilità*, Torino: Einaudi.
- Feldstein, M. [1971],
- Harvey, A.C. [1976], “An alternative proof and generalization of a test for structural change”, *The American Statistician*, **30**, pp. 122-123.
- Harvey e Collier [1977], “Testing for functional misspecification in regression analysis”, *Journal of Econometrics*, **6**, pp. 103-119.
- Pagan, A.R., e D.F. Nicholls [1984], “Estimating predictions, prediction errors and their standard deviations using constructed variables”, *Journal of Econometrics*, **24**, pp. 293-310.