# 4 I MINIMI QUADRATI GENERALIZZATI

## Indice del capitolo

## 4.1 L'ipotesi di sfericità dei residui

Nel capitolo precedente abbiamo analizzato la violazione della prima fra le ipotesi stocastiche alla base del modello lineare, quella di omogeneità del campione. Nel corso della trattazione siamo stati condotti ad occuparci anche della violazione di un'altra ipotesi, quella di normalità dei residui, che è connessa alla prima nel caso in cui si manifestino osservazioni anomale. Ci occupiamo ora di un'altra ipotesi stocastica, quella relativa ai momenti secondi della distribuzione dei residui

$$E(\tilde{u}_t \cdot \tilde{u}_s) = \begin{cases} 0 & t \neq s \\ \sigma^2 & t = s \end{cases} \quad \forall t,s \tag{4.1.1}$$

detta di *sfericità* dei residui, alludendo al fatto che sotto l'ipotesi (4.1.1) le regioni dello spazio a $n$ dimensioni nei quali il residuo assume valori con una prefissata probabilità $\alpha$ sono delle ipersfere.[1]

> *Osservazione* 4.1 – Nei modelli specificati su serie storiche, in cui i residui $u_t$ sono una successione temporale di variabili aleatorie, la (4.1.1) viene anche detta ipotesi di *bianchezza* dei residui, poiché nella teoria dei processi stocastici una successione di variabili aleatorie incorrelate a media nulla e varianza costante viene definita *rumore bianco*.[2]

Come si è detto nel paragrafo 1.6, l'ipotesi di sfericità (o, a seconda del tipo di modelli, di bianchezza) è raramente verificata nella realtà, poiché frequentemente i residui sono correlati tra di loro e posseggono varianze diverse. Sia l'autocorrelazione che la non costanza delle varianze sono fenomeni connessi in qualche modo alla crescita economica, che determina la presenza di una forte componente inerziale nelle serie storiche (da cui l'autocorrelazione) e il fatto che le loro fluttuazioni attorno alla tendenza aumentino di ampiezza col passare del tempo (da cui la non costanza delle varianze).

Si noti che la (4.1.1) non prevede alcuna specifica ipotesi distribuzionale. Siamo quindi nel contesto delle ipotesi deboli (o di Gauss-Markov). La violazione delle (4.1.1) quindi determinerà sia la perdita delle proprietà BLU dello stimatore OLS, sia l'impossibilità di effettuare inferenze statistiche valide. Questi punti vengono affrontati in dettaglio nel paragrafo 4.2, dove si presenta uno stimatore generalizzato, detto appunto stimatore dei *minimi quadrati generalizzati*, o di

---

[1] In particolare, sfere a $n$ dimensioni. Più che di sfericità dei residui bisognerebbe in effetti parlare di sfericità della distribuzione (di forma non specificata) dei residui.

[2] Si veda il paragrafo 1.6.

Aitken, o GLS,[3] che mantiene le proprietà BLU e permette inferenze valide anche in presenza di violazioni della (4.1.1).

In pratica, l'applicazione di questo stimatore è limitata dalla necessità di conoscere la matrice di covarianze vera dei residui. Sono quindi stati proposti diversi stimatori cosiddetti "GLS fattibili",[4] detti anche *stimatori di Aitken a due stadi*, i quali utilizzano, a fronte di violazioni specifiche della (4.1.1), specifici stimatori della matrice di covarianze (o specifiche ipotesi sulla sua struttura) per pervenire alla costruzione delle stime generalizzate. I diversi stimatori fattibili sono quindi determinati dalla natura della violazione ipotizzata.

In questo capitolo considereremo in particolare quelli proposti per ovviare alla presenza di eteroschedasticità dei residui. Il più diffuso è lo stimatore dei *minimi quadrati ponderati*, o WLS,[5] presentato nel paragrafo 4.3. Gli stimatori GLS fattibili proposti a fronte di violazioni dell'ipotesi di non autocorrelazione dei residui verranno presentati nel modulo V, dopo aver introdotto i concetti di dinamica econometrica necessari per rappresentare l'autocorrelazione stessa.

[3] Dall'inglese *Generalized Least Squares*.

[4] In inglese *feasible GLS*.

[5] Dall'inglese *Weighted Least Squares*.

## 4.2 Lo stimatore dei minimi quadrati generalizzati

Come si è visto nel precedente paragrafo, spesso, per trattare situazioni più aderenti alla realtà, è necessario indebolire ulteriormente l'ipotesi stocastica (4.1.1), o, se si vuole, estendere il modello di base, ipotizzando che la matrice di dispersione di $\tilde{\mathbf{u}}$ sia del tipo più generale

$$Cov(\tilde{\mathbf{u}}) = E(\tilde{\mathbf{u}}\tilde{\mathbf{u}}') = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2n}^2 \\ \vdots & & \ddots & \vdots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \cdots & \sigma_{nn}^2 \end{bmatrix} = \boldsymbol{\Sigma} = \sigma^2\mathbf{V} \tag{4.2.1}$$

dove $\boldsymbol{\Sigma}$ e $\mathbf{V}$ sono matrici simmetriche definite positive. La scelta della notazione $\boldsymbol{\Sigma}$ o dell'altra $\sigma^2\mathbf{V}$ dipende dal contesto di analisi econometriche in cui ci si trova, essendo la seconda motivata dall'avere conformazione analoga all'ipotesi $\sigma^2\mathbf{I}$ di residui sferici.

> *Osservazione* 4.2 – La matrice di dispersione $\boldsymbol{\Sigma}$ di un vettore aleatorio $\tilde{\mathbf{y}}$ è sempre definita positiva in quanto è:
>
> $0 < Var(\boldsymbol{\alpha}'\tilde{\mathbf{y}}) = \boldsymbol{\alpha}'\,Cov(\tilde{\mathbf{y}})\,\boldsymbol{\alpha} = \boldsymbol{\alpha}'\,\boldsymbol{\Sigma}\,\boldsymbol{\alpha}$
>
> dove $\boldsymbol{\alpha}$ è un qualsiasi vettore di costanti non tutte nulle e $\boldsymbol{\alpha}'\tilde{\mathbf{y}}$, essendo una variabile aleatoria scalare (non degenere), ha varianza necessariamente nonnulla.

*Proprietà dello stimatore OLS in presenza di disturbi non sferici*

Prima di procedere mostriamo in che modo l'ipotesi (4.2.1) modifica le proprietà dello stimatore OLS. Questa analisi è essenziale allo scopo di farci capire quali sono i rischi che corriamo nell'applicare lo stimatore OLS in contesti nei quali una delle ipotesi alla base di esso, la (4.1.1), non è verificata.

Ricordiamo dal capitolo 1 che sotto le ipotesi standard lo stimatore OLS è BLU. Sotto la (4.2.1) questo stimatore mantiene la proprietà di non distorsione, ma non è più ottimo nel senso di Gauss-Markov. Avremo infatti

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{u}}] = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\tilde{\mathbf{u}}] = \boldsymbol{\beta}$$

Viceversa, la sua matrice di varianze e covarianze ora diventa

$$\begin{aligned} Cov(\hat{\boldsymbol{\beta}}) &= E[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'] = E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{u}}\tilde{\mathbf{u}}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \neq \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \tag{4.2.2}$$

cioè diversa dalla (1.6.18) derivata nel caso di validità delle ipotesi classiche. Ma quest'ultima matrice di dispersione è quella sulla quale si basa la dimostrazione del teorema di Gauss-Markov (si veda il paragrafo 1.8). Ne consegue che quando la matrice di dispersione dei residui è la (4.2.1) lo stimatore OLS non è più quello di varianza minima fra tutti gli stimatori lineari non distorti. L'intuizione sottostante a questo risultato è semplice: lo stimatore OLS non è più efficiente in quanto non sfrutta tutta l'informazione (teoricamente) disponibile. In particolare, esso ignora la conformazione della matrice di covarianze (4.2.1), e risulterà quindi più disperso di uno stimatore che invece ne tenga conto.

La (4.2.2) indica anche che lo stimatore campionario della matrice di covarianze dei residui $\overline{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$ non è più appropriato, perché non riflette la struttura della matrice di covarianze vera (4.2.2). Come abbiamo visto nel capitolo 2, tutta l'inferenza statistica sui parametri del modello utilizza una stima della loro matrice di dispersione e di conseguenza in presenza di residui non sferici le inferenze effettuate utilizzando la matrice di varianze e covarianze "standard" non sono valide. In altre parole, le statistiche $t$ e $F$ calcolate dai software econometrici potrebbero fornire risultati fuorvianti (cioè, a seconda delle circostanze, spingerci a non rifiutare un'ipotesi nulla falsa o a rifiutare una nulla vera).

Le conseguenze dell'ipotesi (4.2.1) sullo stimatore OLS sono quindi piuttosto gravi. Vediamo ora come costruire uno stimatore generalizzato che, tenendo conto della (4.2.1), ci consenta di superare questi problemi.

*Derivazione dello stimatore generalizzato*

Se valgono le ipotesi (4.2.1), le prime due delle (1.6.10) e la (1.4.9) è possibile determinare un nuovo stimatore che segue il criterio dei minimi quadrati, detto *stimatore dei minimi quadrati generalizzati*[6].

Per ottenerlo è sufficiente far ricorso al teorema XIX-1.10, che consente di fattorizzare la matrice $\mathbf{V}$ secondo la forma

$$\mathbf{V}=\mathbf{P}\mathbf{P}' \tag{4.2.3}$$

dove la matrice $\mathbf{P}$ è un'appropriata matrice quadrata, anch'essa di ordine $n$; dalla (4.2.3) si trae, con due inversioni,

[6] O *di Aitken*; in lingua inglese: *Generalized Least Squares*, GLS. Riportiamo per comodità del lettore le quattro ipotesi alla base dello stimatore GLS:

1) $\mathbf{X}$ matrice di costanti
2) $E(\mathbf{u}) = 0$
3) $Cov(\mathbf{u}) = \sigma^2\mathbf{V}$
4) $\det(\mathbf{X}'\mathbf{X}) \neq 0$

$$\mathbf{P}^{-1}\mathbf{V}(\mathbf{P}')^{-1}=\mathbf{I}_n$$

l'invertibilità di $\mathbf{P}$ essendo assicurata dallo stesso teorema XIX-1.10. Se premoltiplichiamo il modello (1.4.4) per $\mathbf{P}^{-1}$ otteniamo

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{P}^{-1}\mathbf{X}\beta + \mathbf{P}^{-1}\mathbf{u}$$

cioè

$$\mathbf{y}_* = \mathbf{X}_*\beta+\mathbf{u}_* \qquad (4.2.4)$$

avendo posto

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{y}_*, \qquad \mathbf{P}^{-1}\mathbf{X} = \mathbf{X}_*, \qquad \mathbf{P}^{-1}\mathbf{u} = \mathbf{u}_* \qquad (4.2.5)$$

Il vantaggio delle (4.2.5) è che in seguito alla loro applicazione i residui del modello (4.2.4), specificato in termini di variabili trasformate, rispettano le ipotesi standard:

$$E(\tilde{\mathbf{u}}_*) = E(\mathbf{P}^{-1}\tilde{\mathbf{u}}) = \mathbf{0}$$
$$Cov(\tilde{\mathbf{u}}_*) = E[\mathbf{P}^{-1}\tilde{\mathbf{u}}\tilde{\mathbf{u}}'(\mathbf{P}^{-1})'] = \sigma^2\mathbf{P}^{-1}\mathbf{V}(\mathbf{P}')^{-1} = \sigma^2\mathbf{I}_n$$

dove si è fatto uso della (4.2.3). Lo stimatore dei minimi quadrati ordinari dei parametri della (4.2.4) è

$$\tilde{\beta}_* = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\tilde{\mathbf{y}}_* = [\mathbf{X}'(\mathbf{P}^{-1})'\mathbf{P}^{-1}\mathbf{X}]^{-1}\mathbf{X}'(\mathbf{P}^{-1})'\mathbf{P}^{-1}\tilde{\mathbf{y}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\tilde{\mathbf{y}} \qquad (4.2.6)$$

per ottenere il quale abbiamo invertito i due membri della (4.2.3). Il vettore aleatorio (4.2.6) è lo stimatore dei minimi quadrati generalizzati; esso è BLU perché ottenuto tramite il criterio dei minimi quadrati ordinari operando su un modello, il (4.2.4), che rispetta le ipotesi standard. La matrice di dispersione di tale stimatore è

$$Cov(\tilde{\beta}_*) = \sigma^2(\mathbf{X}_*'\mathbf{X}_*)^{-1} = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \qquad (4.2.7$$

mentre la stima non distorta di $\sigma^2$ ha l'espressione

$$\begin{aligned}\overline{\sigma}_*^2 &= \frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*}{n-k} = (\mathbf{y}_* - \mathbf{X}_*\beta_*)'(\mathbf{y}_* - \mathbf{X}_*\beta_*)/(n-k) = \\ &= (\mathbf{y} - \mathbf{X}\beta_*)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta_*)/(n-k) = (\mathbf{y}'\mathbf{V}^{-1}\mathbf{y} - \beta_*\mathbf{X}'\mathbf{V}^{-1}\mathbf{y})/(n-k)\end{aligned} \qquad (4.2.8)$$

che si ottiene sfruttando la (4.2.7).

Nelle applicazioni l'utilità dello stimatore dei minimi quadrati generalizzati è notevolmente limitata dalla necessità di conoscere $\mathbf{V}$. Nel paragrafo successivo vedremo come viene risolto il problema nel caso particolare di eteroschedasticità dei residui.

## 4.3 Eteroschedasticità dei residui

Nel modello lineare generale

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t \tag{4.3.1}$$

sono state fatte sinora le ipotesi deboli standard

$$E(\tilde{u}_t) = 0 \qquad \forall t \quad , \quad E(\tilde{u}_t \cdot \tilde{u}_s) = \begin{cases} 0 & t \neq s \\ \sigma^2 & t = s \end{cases} \qquad \forall\, t, s \tag{4.3.2}$$

Nelle variabili economiche, tuttavia, accade spesso che la variabilità non sia costante nel tempo, ma crescente o più raramente decrescente, oppure ancora crescente e poi decrescente a tratti. Se una tale situazione vale per la (4.3.1) e se la componente sistematica (la combinazione lineare delle $x$) non rappresenta sufficientemente questa variabilità non costante, tale tipo di variabilità si trasferisce sui residui $u_t$ per cui la seconda della (4.3.2) si trasforma nella

$$E(\tilde{u}_t \cdot \tilde{u}_s) = \begin{cases} 0 & t \neq s \\ \sigma_t^2 & t = s \end{cases} \tag{4.3.3}$$

caratterizzandone la *eteroschedasticità*.

In tale caso l'analisi svolta nel paragrafo precedente mostra come non possano essere più utilizzati gli stimatori dei minimi quadrati ordinari, per i quali è necessario che valgano le (4.3.2). Nel resto del paragrafo mostriamo come costruire stimatori GLS nel caso di eteroschedasticità e poi richiamiamo alcuni test comunemente utilizzati per verificare l'ipotesi di omoschedasticità.

*La stima dei minimi quadrati ponderati (WLS)*

Viene naturale ipotizzare che l'eteroschedasticità dei residui sia causata da alcune variabili note che indichiamo con $z_{1t}, z_{2t}, \ldots, z_{st}$. Queste possono essere, tutte o in parte, anche variabili esplicative del modello (4.3.1). Sotto l'ulteriore ipotesi che $\sigma_t^2$ sia funzione crescente (l'adattamento al caso decrescente è banale) di queste variabili, possiamo porre

$$\sigma_t^2 = \exp(\alpha_1 z_{1t}) \cdot \exp(\alpha_2 z_{2t}) \cdot \ldots \cdot \exp(\alpha_s z_{st}) \tag{4.3.4}$$

dove la crescenza è rappresentata mediante l'esponenziale per comodità di sviluppo analitico. Sempre per comodità è conveniente specializzare ulteriormente la (4.3.4) senza che le ipotesi addizionali condizionino troppo le situazioni reali.

Poniamo, dunque, in primo luogo $s = 2$, $z_{1t} = 1$ $\forall t$, per cui la (4.3.4) diventa

$$\sigma_t^2 = \exp(\alpha_1) \cdot \exp(\alpha_2 z_{2t}) = \sigma^2 \cdot w_t^{\alpha_2} \tag{4.3.5}$$

avendo posto

$$\sigma^2 = \exp(\alpha_1)$$

$$z_{2t} = \ln w_t$$

In secondo luogo poniamo $\alpha_2 = 2$, per cui in conclusione si ha

$$\sigma_t^2 = \sigma^2 \cdot w_t^2 \tag{4.3.6}$$

Se $w_t = 1$, $\forall t$, si ritorna all'ipotesi standard di omoschedasticità.

Sotto l'ipotesi (4.3.6), per eliminare l'eteroschedasticità basta dividere il modello (4.3.1) per $w_t$

$$\frac{y_t}{w_t} = b_1 \frac{x_{1t}}{w_t} + b_2 \frac{x_{2t}}{w_t} + \ldots + b_k \frac{x_{kt}}{w_t} + \frac{u_t}{w_t} \tag{4.3.7}$$

che si può stimare con gli OLS; infatti

$$E\left(\frac{\tilde{u}_t}{w_t}\right) = \frac{1}{w_t} E(\tilde{u}_t) = 0 \qquad \forall t$$

$$E\left(\frac{\tilde{u}_t}{w_t} \cdot \frac{\tilde{u}_s}{w_s}\right) = \frac{1}{w_t \cdot w_s} E(\tilde{u}_t \cdot \tilde{u}_s) = \begin{cases} 0 & t \neq s \\ \dfrac{1}{w_t^2} \sigma_t^2 = \sigma^2 & t = s \end{cases}$$

avendo fatto uso della (4.3.6).

La stima effettuata in questo modo è detta dei *minimi quadrati ponderati* o *WLS*, poiché ogni elemento $t$-esimo del campione viene pesato con il fattore $1/w_t$.

*Test di omoschedasticità*

Illustriamo ora, senza la dimostrazione che può essere trovata negli articoli originali, due test che sono comunemente usati per verificare l'eteroschedasticità dei residui. Nel primo, dovuto a Goldfeld e Quandt [1965], l'ipotesi nulla è quella di omoschedasticità

$$H_0 : \sigma_t^2 = \sigma^2 \qquad \forall t \tag{4.3.8}$$

e il test viene sviluppato secondo i seguenti passi

1) si ritiene per ipotesi che una sola variabile $z_t$ sia responsabile dell'eteroschedasticità;

2) si ordinano le osservazioni di $y_t$ secondo la dimensione dei valori di $z_t$;[7]

[7] In modelli stimati su serie storiche la necessità di riordinare i dati secondo i valori della variabile ritenuta responsabile dell'eteroschedasticità può sconvolgere il naturale ordinamento cronologico delle osservazioni e quindi alterare la dinamica espressa dai dati.

3) si omettono $c$ osservazioni centrali;
4) si stimano una regressione sui primi $(n-c)/2$ dati ed un'altra sugli $(n-c)/2$ ultimi dati, con $(n-c)/2>k$;
5) si calcolano le devianze $S_1$ e $S_2$ dei residui nelle due regressioni, per le quali, sotto $H_0$, è

$$\frac{S_1^2}{\sigma^2} \sim \chi^2_{(n-c)/2-k} \qquad ; \qquad \frac{S_2^2}{\sigma^2} \sim \chi^2_{(n-c)/2-k}$$

e si costruisce la statistica

$$\frac{S_2^2}{S_1^2} \sim F_{(n-c)/2-k,(n-c)/2-k} \tag{4.3.9}$$

con la quale si esegue un normale test della $F$ di Fisher.

Il secondo test è dovuto a Breusch e Pagan [1979] e presuppone che sotto l'alternativa $H_1: \ \sigma_t^2 \neq \sigma^2$ valga una relazione del tipo di (4.3.4)

$$\sigma_t^2 = h(\alpha_1 z_{1t} + \alpha_2 z_{2t} + \ldots + \alpha_s z_{st})$$

dove $h$ è una funzione indeterminata poiché il test ne è indipendente. Se $z_{1t} = 1$, l'ipotesi nulla

$$H_0\colon \alpha_2 = \alpha_3 = \ldots = \alpha_s = 0$$

suggerisce omoschedasticità poiché in questo caso

$$\sigma_t^2 = h(\alpha_1) = \sigma^2 = \text{costante}$$

I passi da percorrere in questo secondo test sono i seguenti:

1) si stima il modello (4.3.1) con gli OLS e si calcolano i residui stimati $\hat{u}_t$;
2) si calcolano le quantità

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{u}_t^2 \qquad \frac{\hat{u}_t^2}{\hat{\sigma}^2} \qquad \forall t$$

3) si utilizza la $\hat{u}_t^2 / \hat{\sigma}^2$ come variabile proxy di $\sigma_t^2$ e quindi si stimano i parametri della regressione[8]

---

Per questo motivo in tal caso al test di Goldfeld e Quandt se ne preferiscono altri, quali quello di Breusch e Pagan [1979], esposto più avanti.

[8] La divisione per la costante $\hat{\sigma}^2$ serve unicamente a semplificare le elaborazioni metodologiche contenute nel lavoro originale di Breusch e Pagan.

$$\frac{\hat{u}_t^2}{\hat{\sigma}^2} = \alpha_1 + \alpha_2 z_{2t} + \ldots + \alpha_s z_{st} + v_t \qquad (4.3.10)$$

4) si calcola la devianza residua $\sum_{t=1}^{n} \hat{v}_t^2$

5) sotto $H_0$ la *SSE*, differenza tra devianza totale e devianza residua della (4.3.9), è tale che

$$\frac{\tilde{SSE}}{2} \sim \chi^2_{s-1} \qquad (4.3.11)$$

per cui si può effettuare un comune test del chi quadrato per la verifica dell'omoschedasticità.

Il significato intuitivo del test è questo: se sussiste l'eteroschedasticità, e se questa è effettivamente spiegata dalle $s$-1 variabili $z_{2t}, \ldots, z_{st}$ prescelte, allora queste stesse variabili forniranno una buona spiegazione dell'andamento della $\hat{u}_t^2 / \hat{\sigma}^2$ nella (4.3.10), per cui la devianza spiegata della (4.3.10) è abbastanza elevata e la statistica (4.3.11) è maggiore del valore soglia, cadendo quindi nella zona di rifiuto del test del $\chi^2$.

Questo fondamento intuitivo è alla base di una formulazione alternativa del test, proposta da Koenker [1981], che risulta di più rapida implementazione in quanto prescinde dal calcolo di $\hat{\sigma}^2$. Per effettuare il test basta infatti stimare con i minimi quadrati il modello

$$\hat{u}_t^2 = \alpha_1 + \alpha_2 z_{2t} + \ldots + \alpha_s z_{st} + v_t \qquad (4.3.12)$$

Si dimostra che la (4.3.11) può in tal caso essere espressa come:

$$nR^2 \sim \chi^2_{s-1} \qquad (4.3.13)$$

dove $R^2$ è il coefficiente di determinazione non centrato (1.5.8) della (4.3.12). L'ipotesi di omoschedasticità viene quindi respinta se le variabili $z_{jt}$ prescelte ($j$ = 2, ..., $s$) spiegano bene l'andamento del quadrato dei residui.

*Osservazione* 4.3 – La (4.3.12) è un esempio di regressione *ausiliaria*, intendendosi con questo termine una regressione priva di diretto significato economico, che viene stimata generalmente usando grandezze derivate dalla stima di un modello econometrico (ad esempio, i residui OLS) per permettere o semplicemente per facilitare il calcolo delle statistiche di determinati test. La teoria moderna della verifica delle ipotesi utilizza largamente le regressioni ausiliarie.

Rimane da spiegare come vengano scelte queste $z_{jt}$. In mancanza di ipotesi *a priori* specifiche sulla natura dell'eventuale eteroschedasticità, il test viene effettuato ponendo $s = 2$ e $z_{2t} = \hat{y}_t^2$. Si considera quindi come variabile esplicativa dell'omoschedasticità il quadrato dei valori stimati, che è, a sua volta, una combinazione lineare di tutte le esplicative che entrano nel modello. In questo caso la statistica (4.3.13) si distribuisce come un $\chi^2$ con un grado di libertà.

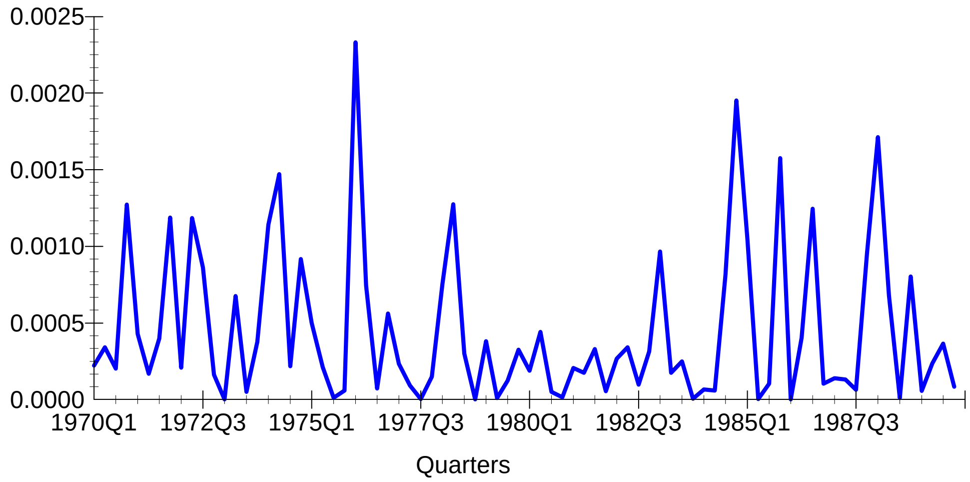


*Figura 4.1 – Il grafico del quadrato dei residui OLS dell'equazione (3.3.29). Il grafico mostra il quadrato dei residui rappresentati nella figura 3.20, considerato come proxy della varianza dei residui della (3.3.29).*

La figura 4.1 rappresenta il quadrato del residuo OLS dell'equazione (3.3.29). Come abbiamo già notato analizzando il grafico dei residui 3.20, anche nella figura 4.1 si distinguono tre episodi: uno intermedio, caratterizzato da bassa variabilità, e due (quello iniziale e quello finale) con variabilità più elevata. Per effettuare il test di Breusch e Pagan nella formulazione (4.3.13) stimiamo la (4.3.12) prendendo come unica esplicativa (oltre all'intercetta) il quadrato dei valori stimati. La stima fornisce questo risultato

$$\hat{u}_t^2 = \underset{(0.6)}{0.0007} \quad \underset{(-0.2)}{- 0.000005} \, (\ln \hat{y}_t)^2 \tag{4.3.14}$$

$$n = 80, \; R^2 = 0.0008$$

(dato che la (4.3.14) è una regressione ausiliaria, non riportiamo per essa la diagnostica completa, ma solo le due grandezze che servono a calcolare la statistica del test che qui interessa, cioè il (4.3.13)). Come si vede, il quadrato dei valori stimati non spiega significativamente l'andamento del quadrato dei residui e di conseguenza la statistica $nR^2$ è pari a 0.06, molto inferiore al valore soglia di un $\chi_1^2$, che al livello 0.05 è pari a 3.8. Di conseguenza l'ipotesi nulla di omoschedasticità dei residui della (3.3.29) non viene respinta dai dati. Questo risultato conferma in qualche misura quello del test di uguaglianza delle varianze condotto con la (3.5.13).

Si noti però che la figura 3.16 mostra come i valori stimati $\ln\hat{y}_t$ della (3.3.29) abbiano una forte tendenza positiva. Questa tendenza sarà ancor più accentuata nella serie $(\ln\hat{y}_t)^2$. D'altra parte, la figura 4.1 mostra che il quadrato dei residui della (3.3.29) non ha una tendenza definita, il che è in un certo senso prevedibile, dato che, come è noto dalla (1.4.18), i residui sommano a zero, per cui le osservazioni positive devono essere necessariamente compensate da altre negative. Ne consegue che quando la variabile dipendente è caratterizzata da una forte tendenza, la regressione ausiliaria (4.3.14) è malspecificata, perché pone in relazione due serie delle quali una sola è caratterizzata da tendenza. In questo caso le due serie nel lungo periodo sono incorrelate e quindi il test è distorto verso l'accettazione dell'ipotesi nulla. In pratica, quindi, il risultato ottenuto con la (4.3.14) è solo apparentemente rassicurante, perché potrebbe essere viziato dalla tendenza della serie delle importazioni.

*Lo stimatore WLS come stimatore dei minimi quadrati generalizzati*

Come abbiamo anticipato, è possibile dimostrare che lo stimatore dei minimi quadrati ponderati (4.3.7) deriva dall'applicazione del principio dei minimi quadrati generalizzati.

Infatti, in termini matriciali l'ipotesi generale (4.3.3) diventa

$$E(\tilde{\mathbf{u}}\tilde{\mathbf{u}}') = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = \mathbf{\Sigma}$$

e inserendovi l'ipotesi più particolare (4.3.6) si ottiene

$$\mathbf{\Sigma} = \sigma^2 \begin{bmatrix} w_1^2 & 0 & \cdots & 0 \\ 0 & w_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & w_n^2 \end{bmatrix} = \sigma^2\mathbf{V}$$

avendo fatto uso della (4.2.1). La $\mathbf{V}$ può essere fattorizzata secondo la (4.2.3), con

$$\mathbf{P} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}$$

come facilmente si verifica, e l'uso dei minimi quadrati generalizzati impone

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{P}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{P}^{-1}\mathbf{u}$$

cioè la (4.3.7).

I minimi quadrati ponderati sono pertanto un esempio di stimatore dei minimi quadrati generalizzati computabile o fattibile, dove queste qualificazioni alludono al fatto che nel caso della (4.3.7) la matrice **V** può essere calcolata, cosa non sempre possibile nel caso generale considerato dal paragrafo 4.2.

## 4.4 Bibliografia

Breusch, T.S., Pagan, A.R. [1979], “A Simple Test for Heteroskedasticity and Random Coefficient Variation”, *Econometrica*, **47**, pp. 1287-1294.

Goldfeld, S.M., Quandt, R.E. [1965], “Some Tests for Heteroscedasticity”, *Journal of the American Statistical Association*, **60**, pp. 539-547.

Johnston, J. [1984], *Econometric Methods*, III edizione, Singapore: McGraw-Hill.

Koenker, R. [1981], “A Note on Studentizing a Test for Heteroscedasticity”, *Journal of Econometrics*, **17**, pp. 107-112.