

# Working Paper Series

---

01/14

## MODELLI STRUTTURALI E FILTRI DI KALMAN PER SERIE STORICHE UNIVARIATE - TEORIA ED APPLICAZIONI CON GRETL

PAOLO CHIRICO





# **Modelli strutturali e Filtri di Kalman per serie storiche univariate**

**Teoria ed applicazioni con Gretl**

Paolo Chirico<sup>1</sup>

**Abstract.** Il Filtro di Kalman è una tecnica statistica per fare previsioni e stimare parametri in opportuni modelli per serie storiche. Questi modelli sono i modelli strutturali nello Spazio degli Stati, così detti perchè con essi il dato storico è strutturato linearmente in componenti non osservabili, la cui variazione di stato (nel tempo) è regolata da equazioni lineari.

Formalmente il Filtro di Kalman è un predittore lineare che fornisce previsioni ottimali del processo stocastico allo studio; è un previsore particolare, perchè si costruisce come un palazzo: un piano (stato temporale) alla volta. Sembra una tecnica complessa perchè utilizza formule apparentemente complesse, ma, se non ci si spaventa di fronte a qualche “formulaccia”, ci si accorge che è una tecnica abbastanza duttile ed utile in molti contesti.

Proprio per non spaventare e demotivare lo studente, questa dispensa è stata pensata nel seguente modo: un primo capitolo in cui sono illustrati i punti salienti della metodologia (cercando nel possibile di limitare le “formulacce”!) e quattro successivi capitoli dedicati ognuno ad un caso studio. L’intento è quello di illustrare la tecnica in maniera pratica, attraverso delle applicazioni che lo studente è invitato a replicare. Anche per questo, è stata dedicata un’appendice all’uso del Filtro di Kalman in Gretl, il software con il quale sono state realizzate le applicazioni nei casi studio.

**Parole chiave:** Serie Storiche, Forma nello Spazio degli Stati, Filtro di Kalman,

**Classificazione JEL:** C22, C50

---

<sup>1</sup>paolo.chirico@unito.it  
Dipartimento di Economia e Statistica  
Università degli studi di Torino

# Indice

<b>1</b>	<b>Modelli nello Spazio degli Stati e Filtri di Kalman</b>	<b>4</b>
1.1	Introduzione . . . . .	4
1.2	Il Filtro di Kalman nel modello a livello locale . . . . .	6
1.2.1	Applicazione del filtro . . . . .	7
1.2.2	Analogie con l'Exponential Smoothing . . . . .	8
1.3	Il modello univariato invariante nel tempo . . . . .	8
1.3.1	Il filtro di Kalman . . . . .	9
1.3.2	Inizializzazione del filtro . . . . .	11
1.4	Stima dei parametri . . . . .	12
1.5	Smoothing . . . . .	13
1.6	Previsione con Filtri di Kalman . . . . .	14
1.7	Filtro di Kalman e modello non gaussiano . . . . .	15
<b>2</b>	<b>Caso Studio 1: Analisi del PIL Italia 1981-2010</b>	<b>17</b>
<b>3</b>	<b>Caso Studio 2: Analisi dell'indice FTSE-MIB</b>	<b>22</b>
<b>4</b>	<b>Caso Studio 3: Analisi della volatilità dei rendimenti finanziari</b>	<b>28</b>
<b>5</b>	<b>Caso Studio 4: Analisi della stagionalità nei prezzi all'ingrosso dell'elettricità in Italia</b>	<b>32</b>
<b>A</b>	<b>Introduzione al Filtro di Kalman in Gretl</b>	<b>38</b>
A.1	Comandi per il filtro . . . . .	39
A.1.1	Esempio 1: il modello a livello locale . . . . .	40
A.1.2	Esempio 2: il modello ARMA . . . . .	43
A.2	Comandi per la previsione . . . . .	45
A.2.1	Previsione di un modello ARIMA(1,1,1) . . . . .	46
<b>B</b>	<b>Derivazione analitica del Filtro di Kalman</b>	<b>48</b>



# Capitolo 1

## Modelli nello Spazio degli Stati e Filtri di Kalman

### 1.1 Introduzione

Cosa hanno in comune i seguenti tre modelli per serie storiche?

**Modello 1:**

$$y_t = \alpha + \beta t + \varepsilon_t \quad (1.1)$$

dove  $\varepsilon_t$  è un white noise con varianza  $\sigma_\varepsilon^2$ .

**Modello 2:**

$$y_t = \delta + \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \quad (1.2)$$

dove  $\varepsilon_t$  è un white noise con varianza  $\sigma_\varepsilon^2$ .

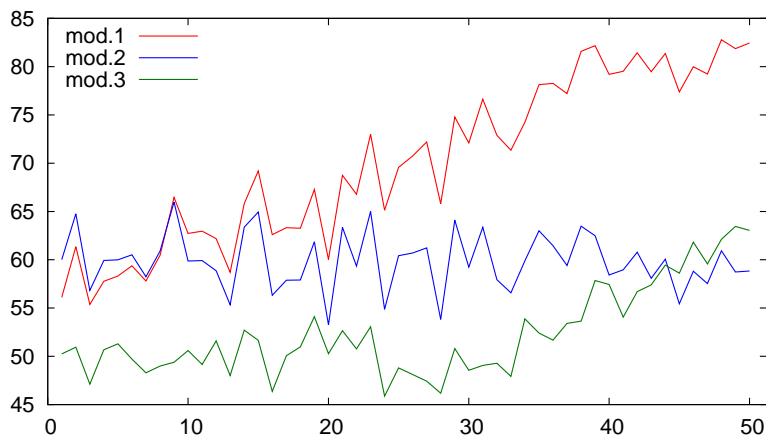
**Modello 3:**

$$\begin{aligned} y_t &= m_t + \varepsilon_t \\ m_{t+1} &= m_t + a_t \end{aligned} \quad (1.3)$$

dove  $\varepsilon_t$  e  $a_t$  sono whitte-noise con varianze  $\sigma_\varepsilon^2$  e  $\sigma_a^2$ .

Il primo è un modello estremamente semplice: ipotizza che i dati storici  $y_t$  si configurino come oscillazioni del tutto casuali intorno ad un trend rettilineo (Figura 1.1). Le previsioni con un tale modello sono semplici ed intuitive: si ottengono proiettando in avanti il trend rettilineo.

Figura 1.1: Esempi di dinamiche modelli 1, 2 e 3



Il secondo è un comune modello ARMA(1,1); se  $|\phi| < 1$ , i dati storici  $y_t$  sono configurati come oscillazioni aleatorie (ma non incorrelate) intorno ad un livello costante. Anche in questo caso le previsioni sono abbastanza semplici da ottenere.

Il terzo modello è noto in letteratura come *modello a livello locale* perché configura i dati storici  $y_t$  come oscillazioni del tutto casuali intorno ad un trend “random walk”. A differenza degli altri due modelli, le previsioni con tale modello non sono intuitive!

Questi tre modelli sembrano abbastanza differenti tra di loro da un punto di vista formale, eppure sia il primo che il secondo modello possono essere rappresentati nella forma del terzo. Infatti, il modello 1.1 può essere riscritto come:

$$\begin{aligned} y_t &= f_t + \varepsilon_t \\ f_{t+1} &= f_t + \beta \end{aligned} \quad (1.4)$$

dove  $f_t$  identifica lo stato al tempo  $t$  del trend rettilineo.

Parimenti il modello 1.2 diventa:

$$\begin{aligned} y_t &= \frac{\delta}{1-\phi} + z_t + \theta z_{t-1} \\ z_t &= \phi z_{t-1} + \varepsilon_t \end{aligned} \quad (1.5)$$

dove  $z_t$  è una variabile latente (non direttamente osservabile).

La “forma comune” alle rappresentazioni 1.3, 1.4 e 1.5 è detta *Forma nello Spazio degli Stati (SSF)* perché composta da due equazioni: una prima che definisce la struttura (*equazione di misura*) del dato storico  $y_t$  ed una seconda che

definisce come le componenti strutturali latenti evolvono da un stato (tempo) al successivo (*equazione di transizione*).

La ricerca di una forma standard di rappresentazione non è ovviamente dettata da ragioni di pura estetica, ma da motivazioni di natura pratica: la *SSF* permette il ricorso ad una metodologia nota come *Filtro di Kalman*, che permette di prevedere il processo mediante la stima delle sue componenti strutturali e dei parametri che le caratterizzano.

## 1.2 Il Filtro di Kalman nel modello a livello locale

Per comprendere meglio la tecnica del “filtro di Kalman” partiamo da una *SSF* semplice e conveniente: il *modello a livello locale gaussiano*:

$$\begin{aligned}y_t &= m_t + \varepsilon_t \\m_{t+1} &= m_t + a_t\end{aligned}$$

dove  $\varepsilon_t \sim NID(0, R)$  e  $a_t \sim NID(0, Q)$ . In pratica si assume che:

1.  $y_t$  differisca dal suo livello atteso per un *errore*  $\varepsilon_t$  white-noise e gaussiano;
2. il livello atteso  $m_t$  evolve nel tempo per *innovazioni*  $a_t$  a priori imprevedibili, indipendenti tra loro e gaussiane.

Per quanto semplice, tale modello si presta a interpretare diversi fenomeni reali, soprattutto in ambito finanziario (vedi Capitolo 3).

In questo modello un ruolo importante hanno i seguenti momenti:

$$m_{t+1|t} = E[m_{t+1}|I_t] \quad (1.6)$$

$$P_{t+1|t} = Var[m_{t+1}|I_t] \quad (1.7)$$

dove  $I_t = y_1, y_2, \dots, y_t$  è l'informazione nota al tempo  $t$ . Infatti la previsione ottima un tempo avanti di  $y_t$  risulta essere:

$$\begin{aligned}y_{t+1|t} &= E[y_{t+1}|I_t] \\&= E[m_{t+1} + w_{t+1}|I_t] \\&= m_{t+1|t}\end{aligned}$$

mentre l'errore quadratico medio di tale previsione risulta:

$$\begin{aligned}MSE_{t+1} &= E[(y_{t+1} - y_{t+1|t})^2|I_t] \\&= Var[m_{t+1} + w_{t+1}|I_t] \\&= P_{t+1|t} + R\end{aligned}$$



Poichè ogni variabile del modello è gaussiana, si può dimostrare (vedi Appendice B) quanto segue:

$$m_{t+1|t} = m_{t|t-1} + k_t(y_t - m_{t|t-1}) \quad (1.8)$$

$$P_{t+1|t} = P_{t|t-1} - k_t^2(P_{t|t-1} + R) + Q \quad (1.9)$$

dove:

$$k_t = \frac{P_{t|t-1}}{P_{t|t-1} + R} \quad (1.10)$$

Le relazioni 1.8, 1.9 e 1.10 (prese in ordine inverso) costituiscono un sistema ricorsivo di aggiornamento di  $m_{t+1|t}$  e  $P_{t+1|t}$  e sono dette Filtro di Kalman<sup>1</sup>.

Infine, combinando la 1.8 con la previsione di  $y_t$  un tempo avanti si arriva al seguente predittore:

$$y_{t+1|t} = y_{t|t-1} + k_t(y_t - y_{t|t-1}) \quad (1.11)$$

con  $k_t$  che viene aggiornato con le formule 1.10 e 1.9.

## 1.2.1 Applicazione del filtro

Per applicare il filtro di Kalman espresso dalle relazioni 1.8, 1.9 e 1.10, risulta evidente che occorre conoscere le varianze  $R$  e  $Q$ , ma anche i valori iniziali di  $m_{1|0}$  e  $P_{1|0}$ .

Supponiamo, in prima battuta, di conoscere le varianze e di non conoscere i valori iniziali. Tale lacuna, nel caso del modello a livello locale, non costituisce un problema, perchè è facile verificare (anche empiricamente) che, qualunque sia l'assegnazione iniziale di  $m_{1|0}$  e  $P_{1|0}$ :

- $P_{t+1|t}$  e  $k_t$  convergono a valori costanti che dipendono da  $R$  e  $Q$ , ma non dai loro valori iniziali;
- $m_{t+1|t}$  da un certo tempo in avanti assume valori che non dipendono dalle assegnazioni iniziali.

Come vedremo nel paragrafo 1.3, tale caratteristica non sempre risulta soddisfatta nei modelli in SSF, ma in genere lo è nella maggiorparte dei modelli di uso comune.

Nella realtà, però, le varianze  $R$  e  $Q$  non sono note e vanno stimate con il metodo della massima verosimiglianza (vedi paragrafo 1.4). Poichè i valori iniziali di  $m_{1|0}$  e  $P_{1|0}$  incidono sulla funzione di verosimiglianza, anche le stime di massima verosimiglianza  $\hat{R}$  e  $\hat{Q}$  dipenderanno da  $m_{1|0}$  e  $P_{1|0}$ ; ma anche i valori di

<sup>1</sup>In effetti se si sviluppa all'indietro la 1.8 si ottiene una combinazione di tutte le osservazioni note  $y_1, y_2, \dots, y_t$  da cui si "filtra" il livello atteso futuro  $m_{t+1}$ .

convergenza di  $m_{t+1|t}$  e  $P_{t+1|t}$ , in quanto dipendenti dalle varianze (stimate), diventeranno dipendenti da quei valori iniziali. Tuttavia esistono degli accorgimenti che permettono di attenuare gli effetti sconvenienti di questa criticità; il paragrafo 1.4, illustrerà meglio questo argomento.

## 1.2.2 Analogie con l'Exponential Smoothing

In ultimo vale la pena osservare che la tecnica del Filtro di Kalman, nel caso di un modello a livello locale, conduce ad un previsore formalmente simile all'Exponential Smoothing. Infatti, poichè  $k_t$  converge ad un valore costante  $\kappa$ , dopo un certo tempo il predittore 1.12 puo essere scritto come:

$$y_{t+1|t} = \kappa y_t + (1 - \kappa)y_{t|t-1} \quad (1.12)$$

che è la formula dell'Exponential Smoothing. Pertanto, se il coefficiente di smussamento nell'Exponential Smoothing viene preso pari al valore di convergenza di  $k_t$ , tale predittore è equivalente al predittore ottimo del Filtro di Kalman.

## 1.3 Il modello univariato invariante nel tempo

Il modello a livello locale illustrato nel paragrafo precedente, così come i modelli 1.4 e 1.5 possono essere ricondotti alla seguente *SSF*<sup>2</sup>:

$$y_t = A_t + \mathbf{h}'\mathbf{s}_t + w_t \quad (1.13)$$

$$\mathbf{s}_{t+1} = \mathbf{F}\mathbf{s}_t + \mathbf{v}_t \quad (1.14)$$

dove:

- $A_t$  è una funzione del tempo (al limite degenera, cioè costante);
- $\mathbf{s}_t$  è il vettore degli stati delle componenti strutturali;
- $\mathbf{h}$  è il vettore dei coefficienti delle componenti in  $y_t$ ;
- $\mathbf{F}$  è la matrice dei coefficienti di transizione delle componenti;
- $w_t$  è una deviazione white-noise (anche detta *componente irregolare* o *rumore*) che influisce sulla misurazione/osservazione di  $y_t$ , ma non ha effetto sulla evoluzione delle componenti strutturali; nell'impostazione base si assume distribuita normalmente con varianza  $R$ ;

<sup>2</sup>La notazione qui adottata è simile a quella proposta da Hamilton e riportata anche nel capitolo 28 del Manuale di Utilizzo di Gretl; la differenza principale sta nella denominazione delle componenti strutturali, che, per ragioni pratiche, sono qui identificate con  $\mathbf{s}_t$ , anziché  $\xi_t$ .

- $\mathbf{v}_t$  sono deviazioni white-noise (*innovazioni*) che incidono sull'evoluzione delle componenti strutturali<sup>3</sup>; anche questi sono assunti distribuiti normalmente con matrice di varianze e covarianze  $\mathbf{Q}$ .

Tabella 1.1: *Forme nello Spazio degli Stati*

modello 1.3	modello 1.4	modello 1.5
$A_t = 0$	$A_t = 0$	$A_t = \delta / (1 - \phi)$
$\mathbf{s}_t = [m_t]$	$\mathbf{s}_t = \begin{bmatrix} f_t \\ \beta \end{bmatrix}$	$\mathbf{s}_t = \begin{bmatrix} z_t \\ z_{t-1} \end{bmatrix}$
$\mathbf{h}' = [1]$	$\mathbf{h}' = [1 \ 0]$	$\mathbf{h}' = [1 \ \theta]$
$\mathbf{F} = [1]$	$\mathbf{F} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	$\mathbf{F} = \begin{bmatrix} \phi & 0 \\ 1 & 0 \end{bmatrix}$
$w_t = \varepsilon_t$	$w_t = \varepsilon_t$	$w_t = 0$
$\mathbf{v}_t = [a_t]$	$\mathbf{v}_t = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\mathbf{v}_{t+1} = \begin{bmatrix} \varepsilon_{t+1} \\ 0 \end{bmatrix}$

Propriamente, il modello 1.13-1.14 è un caso particolare di *SSF* ed identifica un modello *univariato, invariante nel tempo*: univariato perchè relativo ad una singola serie osservata  $y_t$  ed è invariante nel tempo perché le varianze degli errori e i coefficienti del modello non variano nel tempo. Benchè trattasi di una *SSF* ristretta rispetto a quella più generale, in tale forma possono essere espressi una varietà abbastanza ampia di modelli per serie storiche.

La relazione 1.13 definisce *equazione di misura* del modello, mentre la 1.14 rappresenta l'*equazione di transizione* degli stati.

### 1.3.1 Il filtro di Kalman

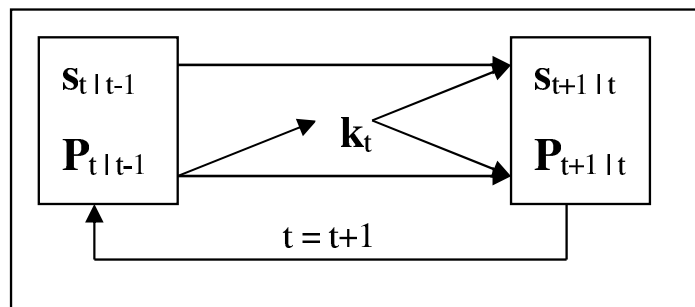
Con riferimento al modello 1.13-1.14, definiamo:

$$\mathbf{s}_{t+1|t} = E[\mathbf{s}_{t+1}|I_t] \quad (1.15)$$

$$\mathbf{P}_{t+1|t} = E[(\mathbf{s}_{t+1} - \mathbf{s}_{t+1|t})^2|I_t] = \text{Var}[\mathbf{s}_{t+1}|I_t] \quad (1.16)$$

<sup>3</sup>Poichè tali deviazioni/innovazioni hanno effetto sulla transizione di  $\mathbf{s}_t$ , possono, secondo la teoria sottostante, essere indicizzati sia al tempo  $t$  che al tempo  $t + 1$ ; in questo secondo caso l'equazione di transizione viene scritta come:  $\mathbf{s}_{t+1} = \mathbf{F}\mathbf{s}_t + \mathbf{v}_{t+1}$ . Tale variante non modifica la formula del filtro di Kalman (vedi par. 1.3.1).

Figura 1.2: Schema ricorsivo del Filtro di Kalman



dove  $I_t = \{y_1, y_2, \dots, y_t\}$  è l'insieme delle informazioni note al tempo  $t$ . È facile notare che  $\mathbf{s}_{t+1|t}$  è la previsione ottima di  $\mathbf{s}_{t+1}$  fatta al tempo  $t$  e  $\mathbf{P}_{t+1|t}$  è la corrispondente matrice MSE (cioè di errori quadratici medi di previsione)

Sia  $\mathbf{s}_{t+1|t}$  che  $\mathbf{P}_{t+1|t}$  hanno un ruolo fondamentale nella previsione ottima di  $y_{t+1}$  fatta al tempo  $t$  ( $y_{t+1|t}$ ), infatti:

$$y_{t+1|t} = A_t + \mathbf{h}'\mathbf{s}_{t+1|t} \quad (1.17)$$

$$MSE_{t+1} = \mathbf{h}'\mathbf{P}_{t+1|t}\mathbf{h} + R \quad (1.18)$$

dove  $MSE_{t+1}$  è l'errore quadratico medio della previsione  $y_{t+1|t}$ .

Se nel modello 1.13-1.14  $w_t$  e  $\mathbf{v}_t$  sono white-noise gaussiani, incorrelati tra di loro e con lo stato iniziale  $\mathbf{s}_1$ , anch'esso gaussiano, con qualche passaggio algebrico-statistico (vedi), si dimostra che:

$$\mathbf{s}_{t+1|t} = \mathbf{F}\mathbf{s}_{t|t-1} + \mathbf{k}_t(y_t - A_t - \mathbf{h}'\mathbf{s}_{t|t-1}) \quad (1.19)$$

$$\mathbf{P}_{t+1|t} = \mathbf{F}\mathbf{P}_{t|t-1}\mathbf{F}' - \mathbf{k}_t(\mathbf{h}'\mathbf{P}_{t|t-1}\mathbf{h} + R)\mathbf{k}_t' + \mathbf{Q} \quad (1.20)$$

dove:

$$\mathbf{k}_t = \mathbf{F}\mathbf{P}_{t|t-1}\mathbf{h}(\mathbf{h}'\mathbf{P}_{t|t-1}\mathbf{h} + R)^{-1} \quad (1.21)$$

Le equazioni 1.21, 1.19 e 1.20 sono note in letteratura come *Filtro di Kalman* (*KF*) e costituiscono un sistema ricorsivo di equazioni per aggiornare in maniera ottimale le previsioni di  $\mathbf{s}_{t+1|t}$  e di  $\mathbf{P}_{t+1|t}$  (vedi Figura 1.2). Il vettore  $\mathbf{k}_t$  indica, nella 1.19, di quanto le previsioni  $\mathbf{s}_{t+1|t}$  vadano corrette in alto o in basso a seconda che quelle precedenti  $\mathbf{s}_{t|t-1}$  abbiano determinato previsioni inferiori o superiori al dato effettivo  $y_t$ . Poichè la correzione avviene in virtù di questa informazione aggiuntiva,  $\mathbf{k}_t$  viene detto *guadagno*.

Per applicare il *KF* occorre, ovviamente, conoscere:

1. i valori di  $\mathbf{s}_{1|0}$  e  $\mathbf{P}_{1|0}$ ;
2. i valori in  $A_t$ ,  $\mathbf{h}$ ,  $\mathbf{F}$ ,  $R$  e  $\mathbf{Q}$ .

Nella realtà, però,  $A_t$ ,  $\mathbf{h}$ ,  $\mathbf{F}$ ,  $R$  e  $\mathbf{Q}$  sono solitamente caratterizzati da parametri incogniti e anche i valori iniziali  $\mathbf{s}_{1|0}$  e  $\mathbf{P}_{1|0}$  non sono noti.

I due paragrafi successivi tratteranno il problema dei valori iniziali  $\mathbf{s}_{1|0}$  e  $\mathbf{P}_{1|0}$  (*inizializzazione del filtro*) e la stima dei valori incogniti in  $A_t$ ,  $\mathbf{h}$ ,  $\mathbf{F}$ ,  $R$  e  $\mathbf{Q}$  (*stima dei parametri*).

### 1.3.2 Inizializzazione del filtro

Poiché inizialmente ( $t = 1$ ) non esiste un'informazione pregressa da sfruttare ( $I_0 = \emptyset$ ), il valore atteso dello stato iniziale e la sua varianza non possono che essere non condizionati:

$$\mathbf{s}_{1|0} = E[\mathbf{s}_1|I_0] = E[\mathbf{s}_1] \quad (1.22)$$

$$\mathbf{P}_{1|0} = \text{Var}[\mathbf{s}_1|I_0] = \text{Var}[\mathbf{s}_1] \quad (1.23)$$

La determinazione/stima di tali momenti avviene, però, in maniera diversa a seconda che il processo degli stati  $\mathbf{s}_t$  sia stazionario o meno.

**Stati stazionari.** In caso di stazionarietà degli stati  $\mathbf{s}_t$ <sup>4</sup> si ha che:

$$E[\mathbf{s}_1] = E[\mathbf{s}_{t-1}] = E[\mathbf{s}_t] \quad (1.24)$$

$$\text{Var}[\mathbf{s}_1] = \text{Var}[\mathbf{s}_{t-1}] = \text{Var}[\mathbf{s}_t] \quad (1.25)$$

Da tali relazioni, si ricavano le seguenti equazioni:

$$(\mathbf{F} - \mathbf{I})E[\mathbf{s}_1] = \mathbf{0} \quad (1.26)$$

$$\text{Var}[\mathbf{s}_1] = \mathbf{F}\text{Var}[\mathbf{s}_1]\mathbf{F}' + \mathbf{Q} \quad (1.27)$$

le cui soluzioni forniscono le formule per l'*inizializzazione* del *KF*. Nel caso,  $\mathbf{F}$  e  $\mathbf{Q}$  contengano parametri incogniti, le formule 1.26 e 1.27 utilizzeranno le stime di questi parametri (vedi paragrafo successivo).

**Stati non stazionari.** Se gli stati  $\mathbf{s}_t$  non sono stazionari, allora sia  $E[\mathbf{s}_1]$  che  $\text{Var}[\mathbf{s}_1]$  non sono determinabili. In questo caso sono possibili solo delle ipotesi su  $\mathbf{s}_{1|0}$  e su  $\mathbf{P}_{1|0}$ . Tuttavia, se per  $\mathbf{s}_{1|0}$  ipotesi verosimili sono in genere fattibili, più complesso è fare ipotesi su  $\mathbf{P}_{1|0}$ ; un approccio cautelativo è tipicamente il seguente:

$$\mathbf{P}_{1|0} = \kappa\mathbf{I} \quad (1.28)$$

---

<sup>4</sup>Cosa che avviene se tutti gli autovalori di  $\mathbf{F}$  sono  $< |1|$ ; in caso contrario il processo  $\mathbf{s}_t$  risulta non stazionario.

dove  $\kappa$  è un numero estremamente grande da essere operativamente simile ad infinito (ad.es  $\kappa = 10^7$ ). Tale soluzione viene detta in gergo tecnico-statistico *diffuse prior* perchè formalizza la massima incertezza possibile.

**Covergenza del filtro.** Un dubbio, che legittimamente può emergere circa l'inizializzazione, riguarda la possibilità di commettere degli errori che avranno un effetto deleterio sulle previsioni future. Fortunatamente, alcune rassicurazioni possono essere fatte in tal senso:

- Adottando l'inizializzazione 1.27, in caso di  $\mathbf{s}_t$  stazionari, si osserva la convergenza di  $\mathbf{P}_{t|t-1}$  a valori costanti e di conseguenza anche  $\mathbf{k}_t$  converge a valori costanti. Inoltre, le previsioni  $\mathbf{s}_{t|t-1}$ , convergono a valori che prescindono dal valore iniziale di  $\mathbf{s}_{1|0}$ .
- In caso di *diffuse prior* ( $\mathbf{s}_t$  non stazionari) le precedenti convergenze non sono garantite, tuttavia nella maggiorparte dei modelli *SSF* per processi  $y_t$  univariati reali tali convergenze sono osservate, sebbene la convergenza avvenga più lentamente.

Pertanto, se il modello è noto (nel senso che non vi sono parametri da stimare), l'inizializzazione del filtro non è in genere un problema, perchè l'errore commesso nelle assegnazioni iniziali viene "riassorbito" più o meno velocemente.

Il discorso è diverso nel caso i parametri debbano essere stimati, perchè l'inizializzazione del filtro deve avvenire prima della stima e influisce sull'esito della stessa. Di questo problema ne parleremo nel paragrafo successivo.

## 1.4 Stima dei parametri

Nella pratica, almeno una varianza in  $R$  e/o in  $\mathbf{Q}$  è ignota, ma spesso anche qualche elemento in  $A_t$ ,  $\mathbf{h}$ ,  $\mathbf{F}$  è un parametro ignoto. Se definiamo  $\Theta$  l'insieme di tutti i parametri incogniti (comprese anche le varianze incognite degli errori), la stima di  $\Theta$  nel *KF* avviene con il metodo della massima verosimiglianza.

Sotto l'assunzione che gli errori  $w_t$  e  $\mathbf{v}_t$  e lo stato  $\mathbf{s}_1$  siano gaussiani, la funzione di log-verosimiglianza può essere scritta come:

$$\ell = -1/2[T \log(2\pi) + \sum_{t=1}^T \log \Sigma_t + \sum_{t=1}^T (y_t - y_{t|t-1})^2 / \Sigma_t] \quad (1.29)$$

dove  $\Sigma_t = \mathbf{h}'_t \mathbf{P}_{t|t-1} \mathbf{h}_t$ .

Ragionando sulla 1.29, si comprende che il legame tra  $\ell$  e i parametri  $\Theta$  non è lineare e la massimizzazione di  $\ell$  deve essere risolta per via numerica; per cui

occorre fissare dei valori iniziali per i parametri ignoti. Non solo: occorre fissare anche il valore iniziale di  $\mathbf{s}_{1|0}$  ( $\mathbf{P}_{1|0}$  dipenderà dai parametri iniziali se gli stati sono stazionari, altrimenti si adotta l’inizializzazione *diffuse prior*).

Questa circostanza costituisce inevitabilmente un punto di debolezza della procedura di stima, perchè a seconda dei valori iniziali assegnati ai parametri e allo stato  $\mathbf{s}_{1|0}$  si possono ottenere soluzioni finali differenti. Addirittura il processo numerico di stima può non sortire alcuna soluzione perchè qualche “intoppo” (non convergenza, impossibilità ai calcoli, etc.) si è verificato.

Un discorso approfondito sui metodi numerici di massimizzazione della verosimiglianza va oltre la presente illustazione, tuttavia è quasi lapalissiano che per minimizzare i possibili “intoppi” è bene partire da valori iniziali il più possibile vicini a quelli “veri”. In molti casi un’analisi preliminare del fenomeno, certamente partendo dal grafico di  $y_t$ , ma non fermandosi solo a quello, aiuta a formulare delle ipotesi circa i valori iniziali. La procedura di stima diventa così un perfezionamento (ottimizzazione) di quello che il buon senso, la conoscenza e l’intelligenza (requisito mai inutile!) sono in grado di formulare.

Certo, spesso non è facile e quindi si devono fare più tentativi, talora rivedere la definizione del modello, perchè un modello non realistico è più difficile da stimare. Ci vuole pazienza e soprattutto la convinzione che l’individuazione (definizione, stima, convalida) di un buon modello è una ricerca che merita un po’ di (lieve) sofferenza!

Infine, due considerazioni:

1. la 1.29 è una verosimiglianza esatta (non condizionata), il che rende la *SSF*, grazie al *KF*, una forma preferibile per la stima di modelli di serie storiche<sup>5</sup>;
2. nel caso di inizializzazione *diffuse prior*, si conviene di non calcolare la verosimiglianza su tutto l’arco temporale, ma dopo un certo periodo  $d$ , perchè i primi addendi in 1.29 potrebbero gravare troppo su  $\ell$  a causa dell’inizializzazione adottata<sup>6</sup>.

## 1.5 Smoothing

Spesso la *SSF* permette di strutturare il processo allo studio in componenti che hanno un significato preciso secondo una certa teoria (si pensi alla scomposizione in “trend+stagionalità+accidentalità”). In questi casi, la stima delle componenti

<sup>5</sup>Si pensi ai modelli a *Media Mobile* per i quali la massimizzazione della verosimiglianza esatta è impraticabile sulla forma originale del modello.

<sup>6</sup>Per lo più  $d$  è posto pari al numero di componenti strutturali in  $\mathbf{s}_t$ .

può avere grande importanza nella comprensione dei “meccanismi” che regolano il processo.

La formula 1.19 fornisce già una stima delle componenti latenti  $\mathbf{s}_t$ , ma tale stima è condizionata all’informazione pregressa  $I_{t-1}$ . Tuttavia, se l’obiettivo della stima è l’interpretazione e non la previsione, meglio sarebbe, se la stima utilizzasse tutta l’informazione a disposizione. In altre parole questa stima dovrebbe essere il risultato di un filtraggio/smoothing di tutti i dati. Tale filtraggio può essere dedotto a posteriori, ovvero dopo aver calcolato dal primo ( $t = 1$ ) all’ultimo periodo ( $t = T$ ) le previsioni  $y_{t|t-1}$  e gli elementi del filtro  $\mathbf{s}_{t|t-1}$ ,  $\mathbf{P}_{t|t-1}$  e  $\mathbf{k}_t$ .

In questa sede non daremo dettaglio del modo con cui vengono ottenute le stime “smussate”  $\mathbf{s}_{t|T}$ <sup>7</sup>; diremo solo che tali stime sono ottenute tramite una procedura a ritroso, che, dal tempo  $T$  al tempo 1, corregge le previsioni “one-step”  $\mathbf{s}_{t|t-1}$  andando a considerare anche l’informazione che esse non hanno precedentemente utilizzata.

## 1.6 Previsione con Filtri di Kalman

Con le formule 1.19-1.21 si arriva fino alle previsioni  $\mathbf{s}_{T+1|T}$  e  $\mathbf{P}_{T+1|T}$ , dalle quali si derivano facilmente le previsioni di  $y_{T+1}$ :

$$y_{T+1|T} = A_t + \mathbf{h}'\mathbf{s}_{T+1|T} \quad (1.30)$$

$$MSE_{T+1} = \mathbf{h}'\mathbf{P}_{T+1|T}\mathbf{h} + R \quad (1.31)$$

La prima formula fornisce una previsione puntuale di  $y_{T+1}$  date le osservazioni note al tempo  $T$ ; con l’aggiunta della seconda si possono fare anche delle previsioni intervallari a diversi livelli di confidenza  $1 - \alpha$ :

$$y_{T+1|T} \pm z_{\alpha/2}\sqrt{MSE_{T+1}} \quad (1.32)$$

dove  $z_{\alpha/2}$  è il percentile  $(1 - \alpha/2)$  della normale standard.

Se si vuole la previsione ottima di  $y_{T+2}$ , fatta al tempo  $T$ , occorre tener presente che:

$$\begin{aligned} \mathbf{s}_{T+2|T} &= E[\mathbf{s}_{T+2}|I_T] \\ &= E[\mathbf{F}\mathbf{s}_{T+1} + \mathbf{v}_{T+2}|I_T] \\ &= \mathbf{F}\mathbf{s}_{T+1|T} \end{aligned} \quad (1.33)$$

---

<sup>7</sup>Nella misura in cui i calcoli vengono generalmente fatti da un software (per noi Gretl), non è strettamente necessario conoscere le formule del calcolo, quanto piuttosto la logica dello stesso. Tuttavia, per i più curiosi e amanti delle formule si può consultare Hamilton, 1994.



e che:

$$\begin{aligned}
\mathbf{P}_{T+2|T} &= \text{Var}[\mathbf{s}_{T+2}|I_T] \\
&= \text{Var}[\mathbf{F}\mathbf{s}_{T+1} + v_{T+2}|I_T] \\
&= \mathbf{F}\mathbf{P}_{T+1|T}\mathbf{F} + \mathbf{Q}
\end{aligned} \tag{1.34}$$

E' facile vedere che le formule 1.33 e 1.34 sono una semplificazione delle formule 1.19 e 1.20 nel caso  $\mathbf{k}_{T+1} = \mathbf{0}$ . Questa semplificazione è del tutto logica se si considera che  $\mathbf{k}_t$ , rappresenta il guadagno derivante dal disporre dell'informazione  $y_t$ . Al tempo  $T + 1$  e per i tempi successivi tale guadagno non sussiste più, perchè le osservazioni note arrivano solo fino al tempo  $T$ . Pertanto, dal tempo  $T + 2$  in avanti il *KF* si risolve nelle formule 1.33 e 1.34: è su queste che si costruiscono in maniera ricorsiva le successive previsioni  $\mathbf{s}_{T+h|T}$  e  $\mathbf{P}_{T+h|T}$  per  $h = 2, 3, \dots$ . Alla fine, non è difficile verificare che per  $h = 2, 3, \dots$  si ha:

$$\mathbf{s}_{T+h|T} = \mathbf{F}^{h-1}\mathbf{s}_{T+1|T} \tag{1.35}$$

$$\mathbf{P}_{T+h|T} = \mathbf{F}\mathbf{P}_{T+1|T}\mathbf{F}' + \sum_{j=0}^{h-2} \mathbf{F}^j\mathbf{Q}\mathbf{F}'^j \tag{1.36}$$

La 1.36 evidenzia come l'MSE relativo a  $\mathbf{s}_{T+h|T}$  cresca al crescere dell'orizzonte temporale di previsione (come è logico che sia!).

Le previsioni puntuale ed intervallare di  $y_{T+h}$  saranno dunque:

$$y_{T+h|T} = A_t + \mathbf{h}'\mathbf{s}_{T+h|T} \tag{1.37}$$

$$y_{T+h|T} \pm z_{\alpha/2}\sqrt{MSE_{T+h}} \tag{1.38}$$

essendo:

$$MSE_{T+h} = \mathbf{h}'\mathbf{P}_{T+h|T}\mathbf{h} + R \tag{1.39}$$

## 1.7 Filtro di Kalman e modello non gaussiano

Nel derivare le formule del Filtro di Kalman si è assunto che gli errori di disturbo  $w_t$  e  $\mathbf{v}_t$ , nonché lo stato iniziale  $\mathbf{s}_1$ , siano gaussiani. Tale assunzione può in molti casi essere ragionevole, ma in altri contesti no. In ambito finanziario, ad esempio, gli errori di disturbo, derivanti dalla non corretta interpretazione delle informazioni, presentano distribuzioni con code elevate, che contrastano con l'assunzione di gaussianità<sup>8</sup>.

<sup>8</sup>Questo è dovuto al fatto che l'emotività degli investitori determina scarti "di una certa entità" più frequentemente rispetto alla normalità gaussiana.

Cosa succede se il modello non è gaussiano? Succede che il secondo membro dell'equazione lineare 1.19 non corrisponde più a  $E[\mathbf{s}_{t+1}|I_t]$ , ovvero il predittore ottimo di  $\mathbf{s}_{t+1}$ : in altre parole il Filtro di Kalman non è più il predittore ottimo (in gergo MMSE<sup>9</sup>) di  $\mathbf{s}_{t+1}$ . Tuttavia il Filtro di Kalman rimane il predittore più efficiente tra quelli lineari (in gergo MMSLE<sup>10</sup>).

Per quanto concerne la stima dei parametri, se il modello non è gaussiano, il secondo membro della 1.29 non corrisponde alla log-verosimiglianza, la cui corretta espressione diventa di complessa formalizzazione. La massimizzazione della 1.29 costituisce un approccio subottimale, noto in letteratura come *Quasi Massima Verosimiglianza* (QML), che, tuttavia, nella maggior parte dei casi fornisce stime vicine a quelle di massima verosimiglianza.

Concludendo: ignorare il fatto che gli errori di disturbo non sono gaussiani (ovvero fingere che lo siano), in genere, non inficia la convenienza dell'uso del Filtro di Kalman, la cui praticità rimane superiore alla sub-ottimalità.

---

<sup>9</sup>Minimum Mean Squared Estimator.

<sup>10</sup>Minimum Mean Squared Linear Estimator.

## Capitolo 2

### Caso Studio 1: Analisi del PIL Italia 1981-2010

Questo primo caso studio riguarda la crescita reale del PIL annuo dell'Italia dal 1981 al 2010. A tale scopo si sono considerati i PIL annui espressi in prezzi del 2000, così che la loro dinamica rifletta esclusivamente la crescita reale della produttività. Come si può notare dal grafico della serie storica (Figura 2.1), i dati sono caratterizzati da una tendenza generale alla crescita, sebbene nel 2008 e soprattutto nel 2009 si siano registrati dei cali a causa della crisi economica, che in quegli anni ha iniziato a manifestarsi.

In prima battuta supponiamo che il PIL annuo sia caratterizzato da un ciclo-trend ( $m_t$ ) rettilineo, ovvero con deriva costante pari a  $\beta$ :

$$\begin{aligned}PIL_t &= m_t + \varepsilon_t \\m_{t+1} &= m_t + \beta\end{aligned}\tag{2.1}$$

con  $\varepsilon_t \sim NID(\sigma^2)$ .

Tale assunzione è chiaramente irrealistica, perchè nega a priori, contro ogni evidenza, la possibilità di cicli economici, ovvero di periodi in cui il trend è crescente, seguiti da periodi in cui il trend è decrescente (o meno cresce). Pertanto, come tra poco sarà evidente, anche il modello che ne deriva dovrà, inevitabilmente, manifestare dei macroscopici limiti!

Poichè il modello 2.1 non è altro che un modello di regressione lineare:

$$PIL_t = \alpha + \beta \cdot t + \varepsilon_t^1\tag{2.2}$$

i suoi parametri possono essere stimati con il classico metodo dei *minimi quadrati ordinari* (OLS); la Tabella 2.1 riporta il risultato di tale stima.

---

<sup>1</sup>Con  $\alpha = m_0$ .

Figura 2.1: *PIL Italia 1981–2010*

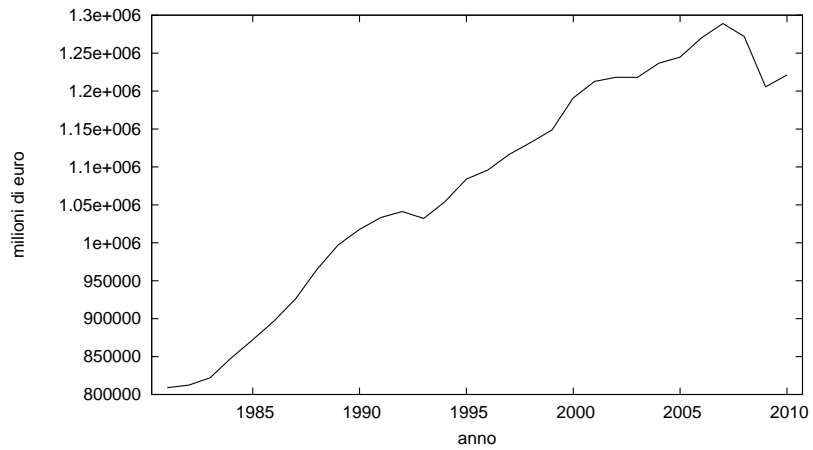
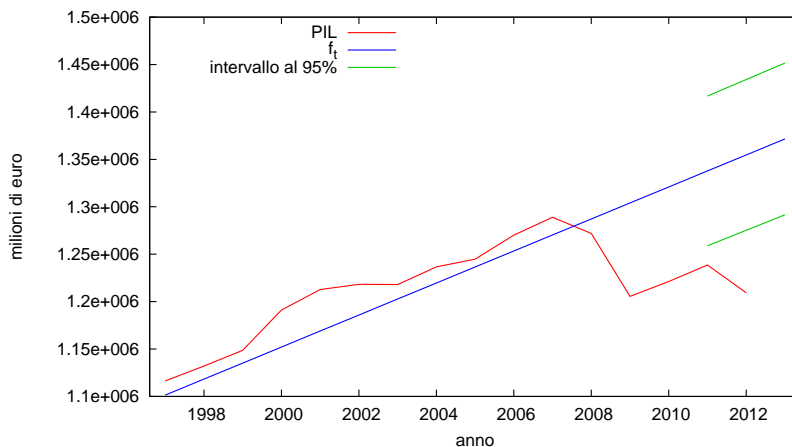


Tabella 2.1: *OLS, PIL Italia 1981–2010 (T = 30)*

	Stima	Errore Std.	rapporto $t$	p-value
$\alpha$	-2.99e+07	1.72e+06	-17.41	0.0000
$\beta$	16884.4	761.836	22.1627	0.0000
Somma quadr. residui	3.65e+10	E.S. della regressione	36116.97	
$R^2$	0.946070	$R^2$ corretto	0.944143	
Log-verosimiglianza	-356.3688	Criterio di Akaike	716.7376	
Criterio di Schwarz	719.5400	Hannan–Quinn	717.6341	

Figura 2.2: Previsioni modello regressivo



Con questo modello la previsione dei PIL successivi al 2010 si ottiene proiettando (prolungando) il trend rettilineo  $m_t$  oltre il 2010, come riportato in Figura 2.2. E' facile notare come l'intervallo di previsione al 95% per gli anni 2011 e 2012 non riesca a comprendere i dati effettivi di quegli anni (confronto *ex post*). Con un solo dato fuori intervallo, poteva trattarsi di un dato anomalo, ma due dati consecutivi evidenziano abbastanza chiaramente l'inadeguatezza del modello e delle assunzioni che vi stanno dietro!

Come i tempi attuali tristemente ci ricordano, il PIL di un paese è soggetto a periodi di maggiore e minore (talora anche negativa!) crescita. Per contemplare questa variabilità del trend, è necessario modificare il modello 2.1 inserendo una relazione che permetta alla deriva del trend di variare, come nel seguente modello:

$$\begin{aligned}
 PIL_t &= m_t + \varepsilon_{1,t} \\
 m_{t+1} &= m_t + \beta_t + \varepsilon_{2,t} \\
 \beta_{t+1} &= \beta_t + \varepsilon_{3,t}
 \end{aligned}
 \tag{2.3}$$

con  $\varepsilon_{1,t} \sim NID(\sigma_1^2)$ ,  $\varepsilon_{2,t} \sim NID(\sigma_2^2)$  e  $\varepsilon_{3,t} \sim NID(\sigma_3^2)$ .

Questo modello presuppone la presenza di tre differenti tipi di shock nel fenomeno allo studio: il primo ( $\varepsilon_{1,t}$ ) ha carattere estemporaneo e non incide sul PIL futuro; il secondo ( $\varepsilon_{2,t}$ ) incide sul livello futuro del fenomeno, ma non sulla sua tendenza a crescere o decrescere; il terzo ( $\varepsilon_{3,t}$ ), invece, modifica tale tendenza caratterizzando in alcuni casi una vera e propria inversione di tendenza. Poichè la variabilità di

questi ultimi due shock è solitamente contenuta<sup>2</sup>, il trend risulta essere *localmente lineare*, anche se nel lungo periodo non lo è affatto<sup>3</sup> Anche in questo caso ci troviamo di fronte ad una semplificazione della realtà, ma sicuramente molto più verosimile della precedente!

Il modello 2.3 è chiaramente riconducibile alla SSF 1.13-1.14, infatti:

$$A_t = 0 \quad \mathbf{h}' = [1 \quad 0] \quad \mathbf{s}_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$w_t = \varepsilon_{1,t} \quad R = \sigma_1^2 \quad \mathbf{v}_t = \begin{bmatrix} \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_3^2 \end{bmatrix}$$

I parametri incogniti del modello sono dunque  $\sigma_1$ ,  $\sigma_2$  e  $\sigma_3$ , cioè le varianze degli scarti accidentali  $\varepsilon_{1,t}$ ,  $\varepsilon_{2,t}$  e  $\varepsilon_{3,t}$ ; i risultati della loro stima sono riportati nella Tabella 2.2. Come si può notare il parametro  $\sigma_2$  non è significativamente diverso

Tabella 2.2: *Stima del modello 2.3*

	stima	Errore Std.	z	p-value
$\sigma_1$	8499,39	4618,20	1,840	0,0657
$\sigma_2$	1318,84	82734,5	0,0159	0,9873
$\sigma_3$	12952,6	3254,01	3,981	0,0000
Log-verosimiglianza	-339,53	Criterio di Akaike	685,05	
Criterio di Schwarz	689,25	Hannan-Quinn	686,39	

da zero, mentre  $\sigma_3$  lo è. Effettivamente, se lo shock ha effetti che si ripercuotono sul futuro, la ripercussione è generalmente sulla capacità di crescita!

Ovviamente, se il parametro  $\sigma_2$  non è significativo, lo scarto  $\varepsilon_{2,t}$  va escluso dalla seconda relazione del modello 2.3 e quest'ultimo va ristimato; la tabella 2.3 riporta i risultati di questa seconda stima. Il parametro  $\varepsilon_{3,t}$  rimane significativamente diverso da zero e questo avvalorava l'assunzione di una deriva del trend variabile nel tempo. Il modello con trend localmente lineare è sicuramente migliore rispetto al modello regressivo e questo è confermato da una log-verosimiglianza maggiore e da indicatori di Akaike, Schwarz e Hannan-Quinn inferiori. Anche il confronto ex post con i dati effettivi del 2011 e 2012 è migliore dal momento che rientrano entrambi nell'intervallo di previsione al 95% (Figura 2.3).

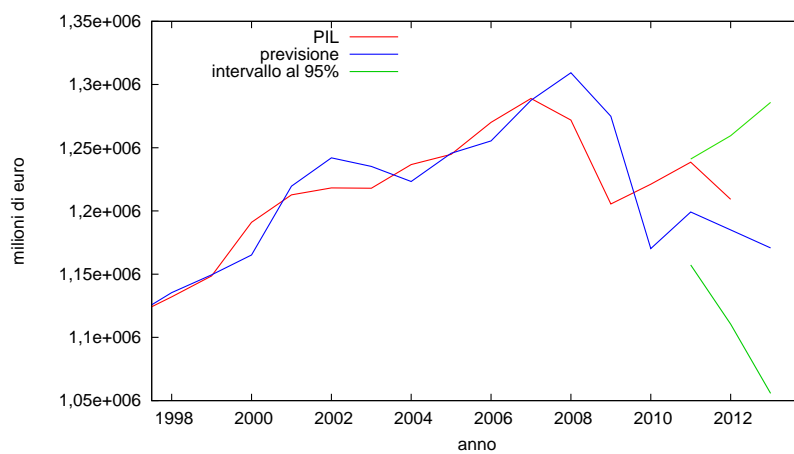
<sup>2</sup>Non perchè non possano verificarsi shock di rilevante entità, ma perchè questi sarebbero l'eccezione piuttosto che la regola!

<sup>3</sup>.

Tabella 2.3: *Stima del modello 2.3 modificato*

	stima	Errore Std.	z	p-value
$\sigma_1$	8356,49	1176,68	7,102	0,0000
$\sigma_3$	13312,0	1686,19	7,895	0,0000
Log-verosimiglianza	-339,53		Criterio di Akaike	683,05
Criterio di Schwarz	685,86		Hannan-Quinn	683,95

Figura 2.3: *Previsioni modello a trend local. lineare*



## Capitolo 3

### Caso Studio 2: Analisi dell'indice FTSE-MIB

Questo secondo caso studio ha un duplice obiettivo: il primo, di natura tecnica, di dimostrare nella pratica come i modelli ARIMA possano essere resi nella SSF e stimati mediante il filtro di Kalman; il secondo, di natura concettuale, di dare un supporto teorico alla modellistica ARIMA utilizzata.

Oggetto dell'analisi è l'indice FTSE-Mib della Borsa di Milano, considerato nel periodo dal settembre 2012 al giugno 2013. Come si può anche notare dal grafico della serie storica (Figura 3.1), tale indice non è un processo stazionario, ma, come per la maggiorparte delle quotazioni finanziarie, è stazionaria la differenza del suo logaritmo  $fm_t = \log(FM_t)$ <sup>1</sup>.

L'analisi del correlogramma delle differenze  $fm_t - fm_{t-1}$  suggerisce l'adozione di un modello Media Mobile di ordine 1 per tali differenze:

$$(fm_t - fm_{t-1}) = \delta + \varepsilon_t + \theta\varepsilon_{t-1} \quad (3.1)$$

con  $\varepsilon_t \sim NID(\sigma^2)$ .

Il modello 3.1 è riconducibile alla SSF 1.13-1.14, con le seguenti associazioni:

$$\begin{aligned} A_t = \delta \quad \mathbf{h}' = [1 \quad \theta] \quad \mathbf{s}_t = \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ w_t = 0 \quad R = 0 \quad \mathbf{v}_{t+1} = \begin{bmatrix} \varepsilon_{t+1} \\ 0 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

I parametri incogniti del modello sono dunque  $\sigma$ ,  $\delta$  e  $\theta$ ; i risultati della loro stima (massima verosimiglianza effettiva!) sono riportati in Tabella 3.1: Due osserva-

---

<sup>1</sup>Tale differenza approssima la variazione relativa del prezzo del FTSE-Mib.



Figura 3.1: *FTSE-Mib 9/12–6/13*

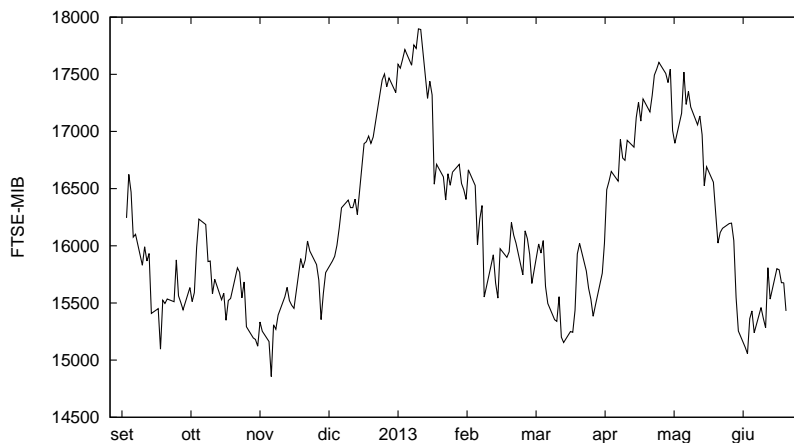


Tabella 3.1: *Stima del modello 3.1*

	Coefficiente	Errore Std.	$z$	p-value
$\sigma$	0,01416	0,00068	20,73	0,0000
$\delta$	-0,0003	0,00084	-0,3356	0,7372
$\theta_1$	-0,1333	0,06649	-2,004	0,0450
Log-verosimiglianza		598,490	Criterio di Akaike	-1190,979
Criterio di Schwarz		-1180,938	Hannan–Quinn	-1186,920

zioni meritano di essere fatte: (i) il parametro  $\delta$  non è significativamente diverso da zero e quindi può essere estromesso dal modello; (ii) il parametro  $\theta$  è negativo e il suo valore assoluto è compreso nell'intervallo  $[0, 1]$ . Entrambe queste circostanze hanno una rilevanza ai fini dello sviluppo di una teoria sottostante il modello appena stimato. Infatti, il modello 3.1 può essere riscritto come:

$$fm_t = fm_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (3.2)$$

con  $\varepsilon_t \sim NID(\sigma^2)$ ;  $0 \leq \theta \leq 1$ .

La previsione un tempo avanti diventa, allora:

$$\begin{aligned} fm_{t+1|t} &= fm_t - \theta \varepsilon_t \\ &= fm_t - \theta(fm_t - fm_{t|t-1}) \end{aligned} \quad (3.3)$$

essendo  $\varepsilon_t = fm_t - fm_{t|t-1}$ .

La relazione 3.3 esprime un'importante teoria comportamentale del mercato. In-

fatti ci dice che: ogni variazione di prezzo rispetto alle attese è considerata, in parte, frutto di entusiasmi/pessimismi eccessivi circa il futuro; il prezzo attuale non viene quindi considerato una esatta previsione del valore atteso futuro del mercato e pertanto tale previsione va corretta al ribasso/rialzo<sup>2</sup>.

Questa teoria comportamentale può essere meglio formalizzata dal seguente modello a livello locale:

$$\begin{aligned} fm_t &= f_t + err_t \\ f_{t+1} &= f_t + new_{t+1} \end{aligned} \quad (3.4)$$

dove:

- $f_t$  è il prezzo atteso corretto al tempo  $t$ , cioè il prezzo che al tempo  $t$  è atteso per il tempo  $t + 1$  se tutte le informazioni sul futuro, note al tempo  $t$ , venissero scontate correttamente;
- $err_t$  è l'errore derivante dallo scontare correttamente tutte le informazioni disponibili; si assume che  $err_t \sim NID(\sigma_e^2)$ ;
- $new_t$  è l'aggiornamento corretto, ma a priori impreveduto, dovuto a nuove informazioni; si assume che  $new_t \sim NID(\sigma_n^2)$ .

In pratica, il modello 3.4 ci dice che il prezzo attuale di mercato viene determinato dall'incontro di domanda e offerta sulla base delle informazioni/convinzioni per il tempo successivo: chi pensa che il prezzo salirà cercherà di comprare facendo salire il prezzo, chi pensa il contrario venderà facendo scendere il prezzo; alla fine il prezzo al tempo  $t$  è la sintesi di tutte le aspettative circa il tempo  $t + 1$ . Se le aspettative fossero in media corrette si avrebbe  $fm_t = f_t$ , ma non necessariamente  $fm_{t+1} = f_t$ , perchè al tempo  $t + 1$  potrebbero sopraggiungere delle novità sul futuro  $new_{t+1}$  che modificano le attese  $f_{t+1}$  e quindi il prezzo  $fm_{t+1}$ .

Dalla formalizzazione 3.4 discende (vedi formule 1.17 e 1.19) il seguente sistema di previsione:

$$fm_{t+1|t} = f_{t+1|t} \quad (3.5)$$

$$f_{t+1|t} = f_{t|t-1} + k_t(fm_t - fm_{t|t-1}) \quad (3.6)$$

Poichè si verifica che  $k_t \rightarrow 1 - \theta$ , sostituendo tale risultato nelle equazioni 3.5 e 3.6 si ottiene:

$$\begin{aligned} fm_{t+1|t} &= fm_{t|t-1} + (1 - \theta)(fm_t - fm_{t|t-1}) \\ &= fm_t - \theta(fm_t - fm_{t|t-1}) \end{aligned}$$

---

<sup>2</sup>I prezzi qui considerati sono propriamente log-prezzi, perchè si prestano meglio a modellizzazioni di tipo lineare. Tuttavia, stante la relazione monotona tra prezzi e log-prezzi, la teoria vale concettualmente in entrambi i casi.

cioè la previsione un tempo avanti del modello ARIMA(0,1,1) (vedi formula 3.3). Tale risultato è generalizzabile e mostra come un modello a livello locale è equivalente ad un modello ARIMA(0,1,1).

Gli operatori non sono in grado di quantificare esattamente quanto della differenza  $fm_t - fm_{t|t-1}$  è riconducibile all'errore di valutazione delle informazioni e quanto alle novità, ma stimano in maniera *forfettaria* l'errore ponendolo pari a una quota  $\theta$  di tale differenza.

**Implicazioni del coefficiente  $\theta$ .** Se  $\theta$  è la quota della differenza  $fm_t - fm_{t|t-1}$  imputata “forfettariamente” all'errore  $err_t$ , allora  $\theta$  sarà tanto maggiore quanto maggiore è la varianza dell'errore ( $\sigma_e^2$ ) rispetto a quella della novità ( $\sigma_n^2$ ). Tale deduzione trova conferma nella relazione matematica che lega  $\theta$  al rapporto tra le predette varianze:

$$\theta = \frac{2}{2 + \rho + \sqrt{\rho^2 + 4\rho}} \quad (3.7)$$

dove  $\rho$  è pari al rapporto segnale/errore  $\sigma_n^2/\sigma_e^2$ .

L'applicazione della formula 3.7 ai risultati di stima del modello 3.4 (Tabella 3.2) fornisce infatti un valore di  $\theta$  pari a 0,129, sostanzialmente vicino al valore ottenuto nella stima diretta del modello ARIMA(0,1,1)<sup>3</sup> (vedi Tabella 3.1).

Tabella 3.2: *Stima del modello 3.4*

	Coefficiente	Errore Std.	z	p-value
$\sigma_e$	0,00504	0,00124	4,071	0,0000
$\sigma_n$	0,01221	0,00107	11,37	0,0000
Log-verosimiglianza	586,267	Criterio di Akaike	-1168,533	
Criterio di Schwarz	-1161,839	Hannan–Quinn	-1165,827	

Il valore di  $\theta$  oscilla dunque tra 0 e 1 a seconda che il rapporto segnale/errore tenda a  $+\infty$  o a 0.

Quando il rapporto segnale/errore è nullo ( $\rho = 0$  e  $\theta = 1$ ), vi è solo errore, ma nessun segnale effettivo di cambiamento ( $new_t = 0$ ):

$$fm_t = f + err_t \quad (3.8)$$

con  $err_t \sim NID(\sigma_e^2)$ . E' chiaramente un caso limite per cui il prezzo  $fm_t$  è determinato solo da informazioni/convinzioni “umorali” che non forniscono (o

<sup>3</sup>Non si è ottenuto lo stesso risultato perchè la stima, in entrambi i casi, avviene mediante una procedura iterativa che viene troncata dall' algoritmo di stima quando si ottiene una certa convergenza, che è differente nei due modelli.

forse oscurano) segnali/novità sul futuro. In questa situazione di massima inefficienza del prezzo (perchè non contiene nessun valido segnale), non vi è motivo per aggiornare le attese sul prezzo

$$fm_{t+1|t} = fm_{t|t-1} \quad (3.9)$$

Per contro, se non esiste errore ( $\rho \rightarrow +\infty$  e  $\theta = 0$ ), allora il prezzo attuale  $fm_t$  sconta correttamente tutte le informazioni sul futuro; il modello di prezzo è quindi un *random walk*:

$$fm_t = fm_{t-1} + new_t \quad (3.10)$$

con  $new_t \sim NID(\sigma_n^2)$ . La previsione al tempo successivo diventa quindi:

$$fm_{t+1|t} = fm_t \quad (3.11)$$

La relazione 3.11 esprime in termini statistici la *teoria dell'efficienza dei mercati finanziari*, secondo la quale il prezzo attuale sconta efficientemente tutte le informazioni sul futuro è pertanto è la media condizionata corretta del prezzo al periodo successivo.

Il valore di  $\theta$  ottenuto dall'analisi dell'indice FTSE-Mib ( $\approx 0,13$ ) ci dice che il mercato azionario corrispondente è *quasi efficiente* nel senso che il valore dell'indice in ogni tempo sconta “quasi correttamente” tutte le informazioni relative al futuro!

**Avvertenze!** Riprendiamo il modello 3.2:

$$fm_t = fm_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \quad (3.12)$$

con  $\varepsilon_t \sim NID(\sigma^2)$ .

Questo modello non è proprio corretto! Infatti, in ambito finanziario, gli scarti  $\varepsilon_t$  tra prezzi effettivi e prezzi attesi non sono di regola distribuiti normalmente, perchè presentano generalmente code elevate (leptocurtosi) e asimmetria negativa. Questo avviene anche nel caso illustrato (curtosi normalizzata=0.84; asimmetria=-0.3) e pertanto l'assunzione di normalità di  $\varepsilon_t$  non può essere soddisfatta.

Nel paragrafo 1.7 si è visto che tale circostanza comporta che il previsore fornito dal Filtro di Kalman non sia ottimale in assoluto, ma solo tra i previsori lineari. Anche la stima dei parametri non è più una stima di massima verosimiglianza e per tale ragione è detta di *quasi massima verosimiglianza (QML)*.

A dirla tutta, gli errori  $\varepsilon_t$  non sono nemmeno serialmente indipendenti, ma, sinceramente, non è questo il problema! Infatti, anche se le assunzioni standard per il Filtro di Kalman fossero tutte soddisfatte, il modello 3.12 non è poi granchè per fare previsioni: con  $\theta = -0.13$  il modello è quasi un modello di *random-walk*! L'utilità di tale modello risiede piuttosto nell'uso che in questo capitolo si è fatto:

modellizzare una teoria, misurare il grado di efficienza del mercato.

Modelli veramente buoni per prevedere quotazioni, se esistono, non sono noti in letteratura! Non potrebbe essere altrimenti: nel momento che fossero noti (ammesso che chi li ha individuati voglia renderli noti, piuttosto che utilizzarli per arricchirsi!), perderebbero immediatamente la loro utilità, perchè il mercato sconterebbe immediatamente tale capacità predittiva nel prezzo attuale. Pertanto la differenza tra prezzo futuro e prezzo attuale continuerebbe a essere determinata (in toto o quasi, a secondo del grado di efficienza del mercato) da ciò che non si è riusciti a prevedere. Se così stanno le cose per il prezzo (o il rendimento) atteso delle azioni, diverso è il discorso per la variabilità attesa di prezzi e rendimenti, in gergo *volatilità*. Ma questo lo vedremo nel capitolo successivo.

## Capitolo 4

### Caso Studio 3: Analisi della volatilità dei rendimenti finanziari

L'efficienza (o la quasi efficienza) dei mercati finanziari fa sì che i prezzi attuali scontino del tutto (o quasi tutto) quello che è possibile sapere sul futuro, così che le differenze tra prezzo futuro e prezzo attuale rispecchino solo (o quasi) ciò che non è prevedibile dal mercato nel suo complesso (cioè in media).

In termini formali l'efficienza dei mercati implica che:

$$r_t = p_t - p_{t-1} = \varepsilon_t \quad (4.1)$$

dove:

- $p_t$  è il log-prezzo di un titolo finanziario o di un portafoglio titoli;
- $r_t$  è il log-rendimento<sup>1</sup> del titolo rispetto al tempo precedente;
- $\varepsilon_t$  è un errore a media nulla, varianza costante, incorrelato nel tempo.

Un tale modello comporta inevitabilmente che la previsione del rendimento (log-rendimento) futuro sia banalmente nulla:

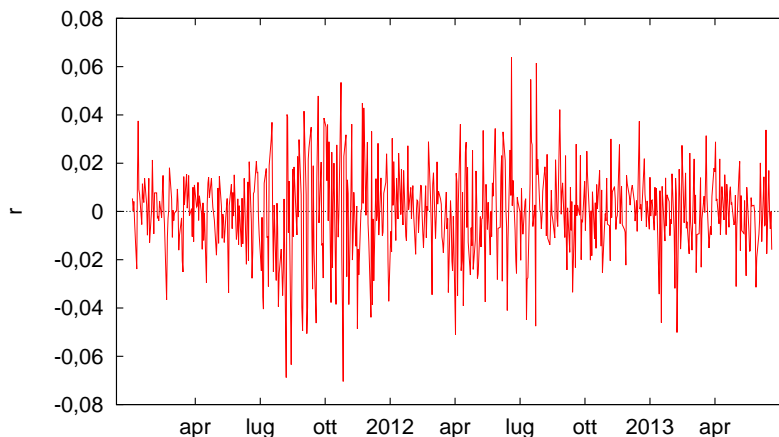
$$r_{t+1|t} = E[r_{t+1}|I_t] = 0 \quad (4.2)$$

Tuttavia, è noto che i rendimenti finanziari, per quanto incorrelati (o poco correlati), non siano distribuiti normalmente e quindi non necessariamente indipendenti nel tempo. Infatti, le varianze dei rendimenti, in gergo volatilità, sono tipicamente dipendenti tra di loro nel tempo. Tale dipendenza si coglie abbastanza chiaramente nella figura 4.1 che riporta l'andamento dei rendimenti del FTSE-Mib

---

<sup>1</sup>Il log-rendimento  $r_t$ , come noto, è una comoda approssimazione del rendimento effettivo  $R_t = (P_t - P_{t-1})/P_{t-1}$ .

Figura 4.1: Rendimenti del FTSE-Mib 1/11–6/13



da inizio 2011 al giugno 2013: la volatilità non è costante, vi sono periodi in cui le oscillazioni sono ampie (alta volatilità) e altri in cui sono minori (bassa volatilità). Questo fenomeno, noto in letteratura come *volatilità a grappoli* (*volatility clustering*), ci evidenzia come la volatilità in ogni tempo  $t$  sia legata a quella del periodo precedente  $t - 1$ .

Esistono diversi modelli che formalizzano questa dipendenza; tra questi i più noti sono sicuramente i modelli GARCH. Purtroppo l'utilizzo del Filtro di Kalman per i modelli GARCH non è particolarmente agevole, mentre si presta bene per una modellizzazione alternativa della volatilità, nota in letteratura come *stochastic volatility model*.

Nel modello di volatilità stocastica dei rendimenti si assume che la volatilità  $\sigma_t^2$  del rendimento dipenda in ogni tempo da quella del tempo precedente a meno di un errore. Più precisamente la dipendenza (a meno di un errore) è lineare sui logaritmi delle varianze, come si evince dal modello:

$$\begin{aligned} r_t &= \sigma_t z_t \\ \log \sigma_t^2 &= \gamma + \phi \log \sigma_{t-1}^2 + \eta_t \end{aligned} \quad (4.3)$$

con  $z_t \sim NID(0, 1)$  e  $\eta_t \sim NID(0, \sigma_\eta^2)$ .

Adottando le seguenti trasformazioni:

$$\begin{aligned} \alpha &= E[\log z_t^2] + \gamma / (1 - \phi) \\ h_t &= \log \sigma_t^2 - \gamma / (1 - \phi) \\ \zeta_t &= \log z_t^2 - E[\log z_t^2] \end{aligned}$$

con qualche passaggio algebrico il modello 4.3 diventa equivalente al seguente:

$$\begin{aligned} \log r_t^2 &= \alpha + h_t + \zeta_t \\ h_t &= \phi h_{t-1} + \eta_t \end{aligned} \quad (4.4)$$

con  $\zeta_t \sim IID(0, 4.93)^2$  e  $\eta_t \sim NID(0, \sigma_\eta^2)$ .

Questo modello fa al caso nostro, perchè è riconducibile alla forma 1.13-1.14 con le seguenti associazioni:

$$\begin{aligned} A_t &= \alpha & \mathbf{h}' &= [1] & \mathbf{s}_t &= [h_t] & \mathbf{F} &= [\phi] \\ w_t &= \zeta_t & R &= 4,93 & \mathbf{v}_t &= [\eta_t] & \mathbf{Q} &= [\sigma_\eta^2] \end{aligned}$$

e permette di risalire dalla componente latente  $h_t$  alla volatilità  $\sigma_t^2$  con la seguente relazione:

$$\sigma_t^2 = \exp(h_t + \alpha + 1,27) \quad (4.5)$$

Tuttavia, per come è stato definito, lo scarto aleatorio  $\zeta_t$  non ha distribuzione gaussiana: pertanto l'utilizzo del Filtro di Kalman per la stima dei parametri  $\alpha$ ,  $\phi$  e  $\sigma_\eta^2$  diventa un metodo QML (vedi paragrafo 1.7), mentre le previsioni  $h_{t+1|t}$  sono ottime tra i previsori lineari.

Il modello 4.4 è stato applicato ai rendimenti (log-rendimenti) del FTSE-Mib da inizio 2011 al giugno 2013; i risultati sono riportati in Tabella 4.1. Nella Figura 4.2, è invece riportata la serie dei  $\sigma_{t|T}$  "smussati" ottenibile dalla serie degli  $h_{t|T}$  "smussati"<sup>3</sup>; tale grafico appare coerente con quello dei rendimenti in Figura 4.1, nel senso che i periodi con maggiore volatilità coincidono con quelli di maggiore oscillazione dei rendimenti.

Tabella 4.1: *Stima del modello 4.4*

	Coefficiente	Errore Std.	z	p-value
$\sigma_\eta$	0,1116	0,0483	2,312	0,0208
$\alpha$	-9,5714	0,3266	-29,30	0,0000
$\phi$	0,9862	0,0112	88,19	0,0000
Log-verosimiglianza	-1399,1	Criterio di Akaike	2804,2	
Criterio di Schwarz	2817,6	Hannan-Quinn	2809,4	

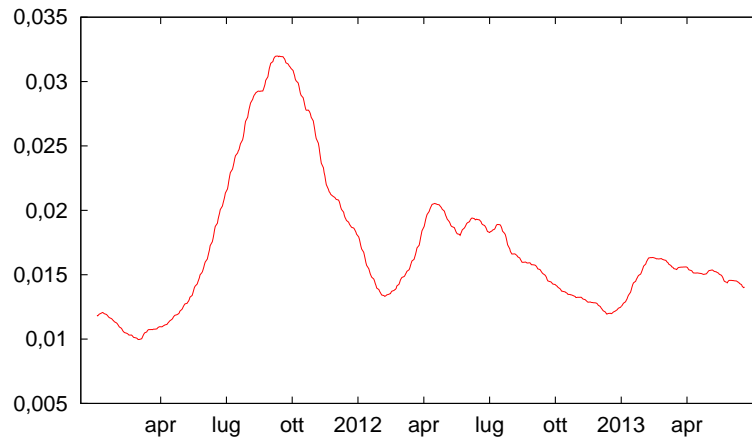
**Distribuzioni a code alte.** Se il modello è correttamente specificato, le stime a posteriori  $z_{t|T} = r_t / \sigma_{t|T}$  dovrebbero risultare serialmente incorrelate e gaussiane.

<sup>2</sup>Se  $z_t \sim NID(0, 1)$ , allora:  $E[\log z_t^2] = -1.27$  e  $Var[\log z_t^2] = 4.93$ ; si veda Harvey 1994.

<sup>3</sup> $\sigma_{t|T} = \exp[(h_{t|T} + \alpha + 1,27)/2]$



Figura 4.2:  $\sigma_{t|T}$ , periodo 1/11–6/13



Spesso, tuttavia, le  $z_{t|T}$  continuano a presentare code elevate (leptocurtosi), così da richiedere la modifica:  $z_t \sim t_\nu$  ( $t$ -Student di parametro  $\nu$ ).

In questo caso cambiano  $E[\log z_t^2]$  e  $Var[\log z_t^2]$ , i cui valori si possono calcolare in maniera esatta con formule abbastanza complesse, oppure in maniera approssimata, ad esempio, con le seguenti:

$$E[\log z_t^2] = -1.27 + 0,978\nu^{-0,978} \quad (4.6)$$

$$Var[\log z_t^2] = 4.93 + 2,52\nu^{-1,06} \quad (4.7)$$

## Capitolo 5

### Caso Studio 4: Analisi della stagionalità nei prezzi all'ingrosso dell'elettricità in Italia

Il 31 marzo del 2004 è partito in Italia un mercato all'ingrosso per produttori e grossisti di energia elettrica nell'ambito di una liberalizzazione dei mercati energetici che è in atto da diversi anni nel nostro paese. Nel mercato a pronti è quotato per ogni ora del giorno un prezzo unico a livello nazionale, denominato PUN; la media dei PUN orari di ogni singolo giorno da luogo al PUN medio giornaliero. Il PUN presenta caratteristiche formali molto interessanti:

1. presenta marcati andamenti periodici di natura stagionale;
2. presenta talora picchi in alto (spike) ed in basso (jump) molto accentuati;
3. presenta periodi di alta volatilità seguiti da periodi di bassa volatilità.

Questo caso studio illustrerà la prima caratteristica con riferimento ai prezzi medi giornalieri. La stagionalità più marcata nel PUN giornaliero è la periodicità settimanale, ma sono presenti anche periodicità mensili e annuali. Un modello abbastanza semplice per il PUN giornaliero è il seguente:

$$\begin{aligned} pun_t &= m_t + g_t + \varepsilon_t \\ m_{t+1} &= m_t + \varepsilon_{m,t} \\ g_{t+1} &= -(g_t + g_{t-1} + \dots + g_{t-5}) + \varepsilon_{g,t} \end{aligned} \quad (5.1)$$

con  $\varepsilon_t \sim NID(0, \sigma^2)$ ,  $\varepsilon_{m,t} \sim NID(0, \sigma_m^2)$  e  $\varepsilon_{g,t} \sim NID(0, \sigma_g^2)$ .

In pratica con tale modello si ipotizza che il prezzo giornaliero sia caratterizzato da:

- un livello medio  $m_t$  che differisce da quello del giorno precedente per effetto di un'innovazione  $\varepsilon_{m,t}$  (random walk);
- un effetto "giorno della settimana"  $g_t$  che si compensa con quelli dei sei giorni precedenti a meno di un'innovazione  $\varepsilon_{g,t}$ ;
- uno scarto inatteso, del tutto casuale,  $\varepsilon_t$ .

Il modello 5.1 è riconducibile alla SSF 1.13-1.14, con le seguenti associazioni:

$$\begin{aligned}
 A_t &= 0 & \mathbf{h}' &= [ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 ] \\
 \mathbf{s}_t &= \begin{bmatrix} m_t \\ g_t \\ g_{t-1} \\ g_{t-2} \\ g_{t-3} \\ g_{t-4} \\ g_{t-5} \end{bmatrix} & \mathbf{F} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\
 w_t &= \varepsilon_t & R &= 0 \\
 \mathbf{v}_t &= \begin{bmatrix} \varepsilon_{m,t} \\ \varepsilon_{g,t} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \mathbf{Q} &= \begin{bmatrix} \sigma_m^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_g^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

Con il modello 5.1 si è analizzato il logaritmo del PUN giornaliero ( $pun_t$ ) nel quadriennio 2008-11; i risultati della stima del modello sono riportati nella Tabella 5.1, mentre nella Figura 5.1 è riportato il grafico dello smoothing di  $g_t$ . Si può notare che tutte e tre le deviazioni standard  $\sigma$ ,  $\sigma_m$  e  $\sigma_s$  sono significative e quindi non solo il prezzo è stocastico, ma anche il livello medio e l'effetto giorno della settimana. Inoltre, poichè i dati sono logaritmi, tali deviazioni standard rappresentano delle variazioni percentuali medie di ogni componente: il livello medio  $m_t$  varia mediamente del 3,9% (rispetto al valore atteso), l'effetto giornaliero varia meno dello 0,5%. Tuttavia, anche se l'effetto giornaliero varia pochissimo, non è periodicamente costante, anzi, come la Figura 5.1 mostra, sembra restringersi nel tempo. Per tener conto di questa evidenza, il modello 5.1 è stato modificato

Tabella 5.1: *Stima del modello 5.1*

	Coefficiente	Errore Std.	z	p-value
$\sigma$	0,0858	0,0015	58,68	0,0000
$\sigma_m$	0,0397	0,0015	25,76	0,0000
$\sigma_g$	0,0034	0,0005	6,65	0,0000
Log-verosimiglianza	1176,25	Criterio di Akaike		-2346,51
Criterio di Schwarz	-2330,37	Hannan-Quinn		-2340,51

inserendo un parametro di “contrazione”  $\alpha$ :

$$\begin{aligned}
 pun_t &= m_t + g_t + \varepsilon_t \\
 m_{t+1} &= m_t + \varepsilon_{m,t} \\
 g_{t+1} &= -\alpha(g_t + g_{t-1} + \dots + g_{t-5}) + \varepsilon_{g,t}
 \end{aligned} \tag{5.2}$$

con  $\varepsilon_t \sim NID(0, \sigma^2)$ ,  $\varepsilon_{m,t} \sim NID(0, \sigma_m^2)$ ,  $\varepsilon_{g,t} \sim NID(0, \sigma_g^2)$  e  $0 < \alpha < 1$ .  
 In questo caso la matrice di transizione diventa:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha & -\alpha & -\alpha & -\alpha & -\alpha & -\alpha \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

I risultati di stima di questo modello sono riportati nella Tabella 5.2: Come ci si aspettava, il parametro di “contrazione”  $\alpha$  è leggermente inferiore ad 1; questo significa che ogni giorno l’ampiezza dell’effetto giornaliero si restringe leggermente, ma in maniera mediamente costante. Questo risultato “sgrava” un po’ la variabilità aleatoria della stagionalità ( $\sigma_g$  che adesso risulta inferiore rispetto a prima. Il nuovo modello è più performante del precedente, come indicato da valori inferiori nei criteri di Akaike, Schwarz e Hannan-Quinn.

Proviamo ad interpretare i risultati emersi. Le Figure 5.2 e 5.3 evidenziano che i prezzi sono inferiori alla media ( $g_t$  negativi) il sabato e soprattutto la domenica, perchè in quei giorni si consuma meno energia elettrica (e quindi cala la domanda); tuttavia, questo avviene in maniera meno marcata a fine 2011 che non a fine 2008. La spiegazione di tale evidenza esula dagli obiettivi di questa dispensa, ma alcuni fatti sembrano coerenti con essa:

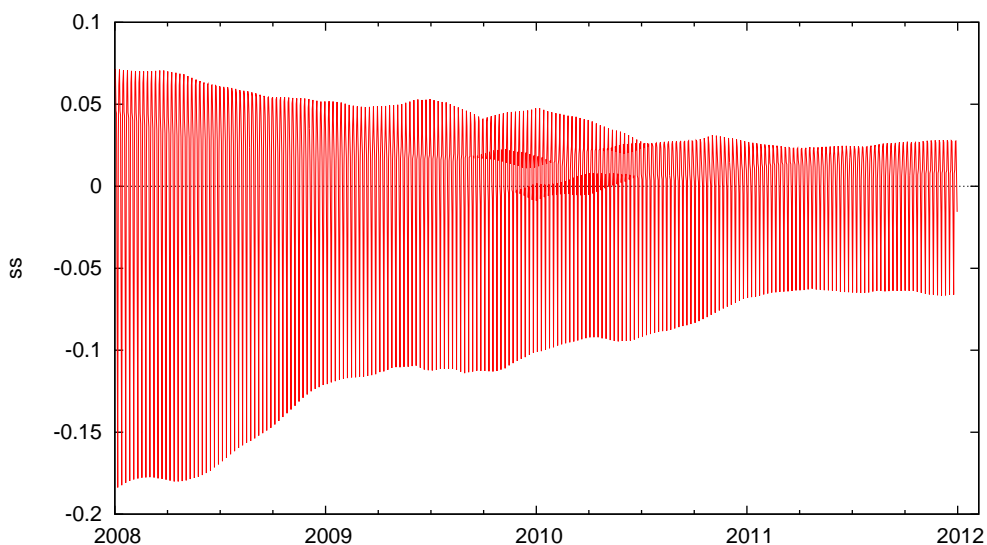


Figura 5.1: Effetto giornaliero sul log-PUN

1. le famiglie tendono a spostare parte dei loro consumi elettrici in periodi in cui il costo all'ingrosso dell'elettricità è minore: alla sera dei giorni feriali e nel fine settimana; questo fa diminuire il costo medio dei giorni feriali ed aumentare quello del sabato e della domenica;
2. fra le molte conseguenze della crisi economica vi è pure un calo del consumo di energia elettrica durante i giorni lavorativi.
3. l'ingresso nel mercato di piccoli e medi produttori di energia elettrica ha reso l'offerta un po' più flessibile, permettendo al prezzo di essere un po' più stabile<sup>1</sup>.

Anche sul fronte dell'offerta si è registrata una migliore organizzazione della produzione e della rete che permette di avere minore divario tra produzione ed offerta nell'arco della settimana.

---

<sup>1</sup>In verità non ancora abbastanza, sicuramente meno rispetto a diversi paesi europei!

Tabella 5.2: *Stima del modello 5.2*

	Coefficiente	Errore Std.	$z$	p-value
$\sigma$	0,0858	0,0008	103,3	0,0000
$\sigma_m$	0,0397	0,0004	104,3	0,0000
$\sigma_g$	0,0029	0,0006	5,03	0,0000
$\alpha$	0,9923	0,0019	527,1	0,0000
Log-verosimiglianza	1186,58	Criterio di Akaike	-2365,15	
Criterio di Schwarz	-2343,64	Hannan-Quinn	-2357,16	

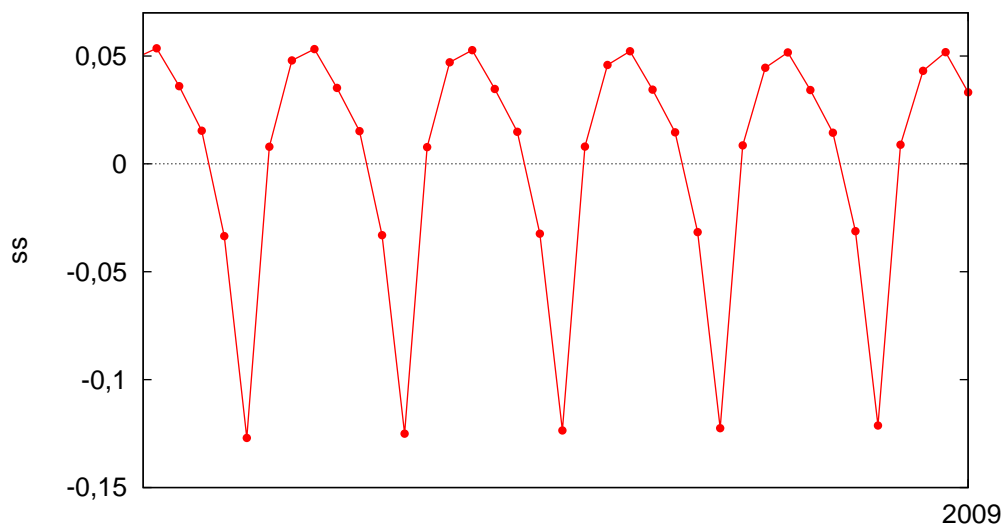


Figura 5.2: Effetto giornaliero nel dicembre 2008

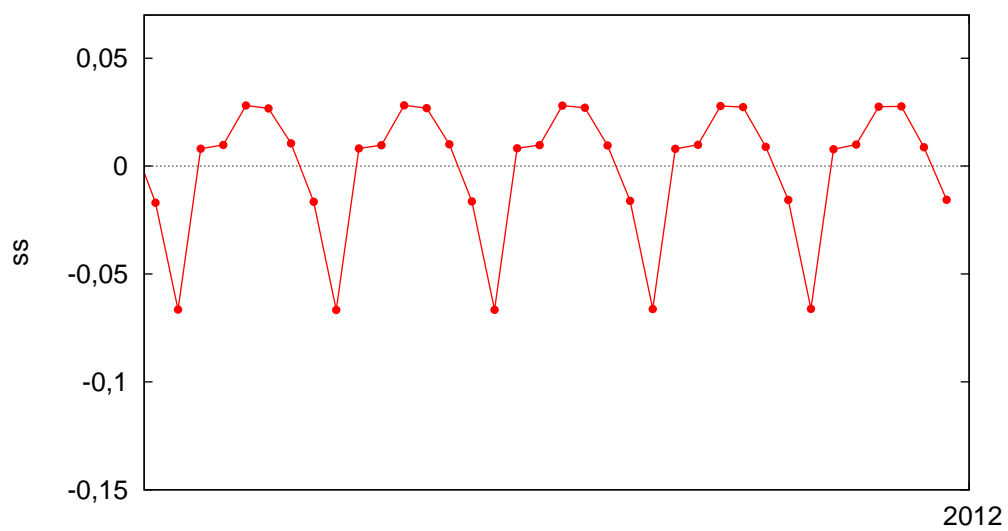


Figura 5.3: Effetto giornaliero nel dicembre 2011

# Appendice A

## Introduzione al Filtro di Kalman in Gretl

Gretl<sup>1</sup> è un software econometrico completamente gratuito, rilasciato con licenza GNU, ideato da Allin Cottrell (Wake Forest University, USA). A lui e a Riccardo Jack Lucchetti (Università Politecnica delle Marche, Italia), seconda anima del progetto, si è aggiunta negli anni una “community” di utilizzatori che con le loro osservazioni e contributi ha permesso a Gretl di essere oggi un potente e sempre più diffuso software econometrico. Il punto di forza di Gretl risiede, oltre nell’essere gratuito, nell’essere “maneggevole ed intuitivo”, non solo per esperti, ma anche per studenti universitari con basi di statistica ed econometria. Gretl presenta, infatti, diverse caratteristiche favorevoli al suo utilizzo:

- si installa facilmente scaricando un file eseguibile dal sito <http://gretl.sourceforge.net/>;
- oltre alla versione base in Inglese, è disponibile in diverse versioni-lingua;
- è corredato da una Guida di Utilizzo dettagliata e ricca di esempi, costantemente aggiornata nella versione Inglese, ma disponibile anche in Italiano;
- fornisce un ampio database di esempi anche dei principali testi econometrici;
- interagisce con R, il più potente tra i software statistici gratuiti;
- è utilizzabile con interfaccia grafica, ovvero permette di importare ed analizzare dati utilizzando menù a tendina.

L’ultima caratteristica, certamente la più gradita per molti utilizzatori, è valida per gran parte delle analisi econometriche, ma al momento non per i Filtri di Kalman.

---

<sup>1</sup> Acronimo di *Gnu Regression, Econometrics and Time-series Library*



Tuttavia Gretl è utilizzabile non solo con interfaccia grafica, ma anche scrivendo *script* (insieme di comandi) all'interno di un'opportuna finestra comandi (percorso: *File* → *Comandi* → *Nuovo*).

I comandi possono ovviamente essere salvati e successivamente richiamati ed eventualmente modificati per nuove applicazioni<sup>2</sup>.

Prima di entrare nel dettaglio, però, una premessa: la trattazione che segue va intesa solo come introduzione all'uso del Filtro di Kalman in Gretl: diversi aspetti tecnici sono da assumere a "a scatola chiusa". Per un'illustrazione più completa si rimanda al capitolo 29 della User Guide (versione in Inglese) di Gretl.

## A.1 Comandi per il filtro

I modelli *univariati-invarianti nel tempo* in Gretl (vedi capitolo 29 della sua User Guide) sono intesi nella seguente forma:

$$\begin{aligned} y_t &= \mathbf{a}'\mathbf{x}_t + \mathbf{h}'\mathbf{s}_t + w_t \\ \mathbf{s}_{t+1} &= \mathbf{c} + \mathbf{F}\mathbf{s}_t + \mathbf{v}_t \end{aligned} \tag{A.1}$$

con  $w_t \sim NID(0, R)$  e  $\mathbf{v}_t \sim NID(0, \mathbf{Q})$ .

La A.1 differisce leggermente dalla SSF trattata nei capitoli precedenti in due punti:

1.  $A_t$  va espressa come combinazione lineare di un vettore di variabili ausiliarie  $\mathbf{x}_t$  (ad es.  $A_t = \delta$  equivale a:  $\mathbf{a} = \delta$  e  $\mathbf{x}_t = 1$ );
2. è contemplato un vettore di costanti  $\mathbf{c}$  nelle equazioni di transizione (ma può essere omissso se  $\mathbf{c} = 0$ ).

In Gretl, la definizione ed applicazione del Filtro di Kalman avviene mediante un blocco di comandi identificato dalle keyword:

```
kalman
    ...
end kalman
```

---

<sup>2</sup>In questa sede, si darà per nota la conoscenza dell'uso di Gretl, almeno nelle operazioni di base quali l'importazione di dati e l'avvio delle applicazioni disponibili a menù, almeno le più elementari da un punto di vista econometrico. Si tratta di una conoscenza facilmente acquisibile in una mezzora di utilizzo da autodidatta del software (a patto di avere qualche base di statistica ed econometria!).

tali comandi servono a definire, oltre la variabile dipendente  $y_t$ , anche  $\mathbf{a}$ ,  $\mathbf{x}_t$ ,  $\mathbf{h}$ ,  $\mathbf{F}$ ,  $\mathbf{c}$ ,  $R$  e  $\mathbf{Q}$  (Tabella A.1).

Tuttavia uno script completo richiede anche comandi preliminari e successivi. Per comprendere come costruire uno script completo si è deciso di illustrare due casi specifici già visti nei casi studio:

1. il modello a livello locale;
2. il modello ARMA.

Tabella A.1: *Comandi*

comando	esempio
obsy	<i>obsy y</i> - indica che la variabile $y$ è la variabile dipendente del modello
obsx	<i>obsx x</i> - indica che il vettore di variabili $x$ è il vettore delle variabili ausiliarie (se omesso è sottointeso $x=1$ )
obsxmat	<i>obsxmat a</i> - indica che $a$ è il vettore dei coefficienti di $x$ (se omesso è sottointeso $a=0$ )
obsymat	<i>obsymat h</i> - indica che $h$ è il vettore dei coefficienti delle componenti strutturali
statemat	<i>statemat F</i> - indica che $F$ è la matrice di transizione
obsvar	<i>obsvar R</i> indica - che $R$ è la varianza dell'errore di misurazione
statevar	<i>statevar Q</i> - indica che $Q$ è la matrice di varianze e covarianze delle innovazioni sulle componenti strutturali
inistate	<i>inistate s1</i> - indica che il vettore $s1$ contiene gli stati iniziali delle componenti strutturali (se omesso è sottointeso $s1=0$ )
inivar	<i>inivar P1</i> - indica che $P1$ è la matrice iniziale $P$ (se omesso viene calcolata o fissata come <i>diffuse prior</i> )
stconst	<i>stconst c</i> - indica che $c$ è il vettore di costanti (se omesso è sottinteso $c=0$ )

### A.1.1 Esempio 1: il modello a livello locale

Come già visto, il modello a livello locale è costituito dalle seguenti equazioni:

$$\begin{aligned} y_t &= m_t + w_t \\ m_{t+1} &= m_t + v_t \end{aligned}$$

dove  $w_t \sim NID(0, \sigma_1^2)$  e  $v_t \sim NID(0, \sigma_2^2)$ .

Tale modello è riconducibile alla forma A.1 con le seguenti associazioni:

$$\begin{aligned} \mathbf{a}' &= [0] & \mathbf{x}_t &= [1] \\ \mathbf{h}' &= [1] & \mathbf{F} &= [1] \\ R &= \sigma_1^2 & \mathbf{Q} &= [\sigma_2^2] \end{aligned}$$

**Script per il solo filtro: varianze note.** Se le varianze sono note, ad esempio:  $R = 25$ ,  $Q = 9$ , la definizione e filtraggio del modello richiede che nella finestra comandi si scrivano (e poi si eseguano) i seguenti comandi:

```
matrix h = {1}
matrix F = {1}
matrix R = {25}
matrix Q = {9}
matrix s1 = {y[1]}

kalman
  obsy y
  obsymat h
  statemat F
  obsvar R
  statevar Q
  inistate s1
end kalman --diffuse

matrix E V S P K
kfilter(&E, &V, &S, &P, &K)
matrix SM=ksmooth()
```

Il primo blocco di comandi serve per definire numericamente gli elementi della forma generale e per definire il valore iniziale  $s_{1|0} = y_1$ .

Il secondo blocco serve per definire ed applicare il filtro alla variabile  $y$  (Tabella A.1). Il suffisso “-diffuse” è necessario in quanto nel modello a livello locale  $m_t$  non è un processo stazionario (è un random walk!)

Il terzo blocco serve per creare matrici/vettori, nelle quali vengono inseriti dati di output<sup>3</sup>. Precisamente:

- E contiene la serie degli errori di previsione  $y_t - y_{t|t-1}$ ;
- V contiene la serie degli  $MSE_t$ ;

---

<sup>3</sup>Il comando: “matrix E V S P K” va eseguito solo la prima volta!

- S contiene la serie degli stati  $\mathbf{s}'_{t|t-1}$ ;
- P contiene la serie dei  $\text{vech}(\mathbf{P}_{t|t-1})$ , ovvero la serie dei vettori contenenti le varianze e le covarianze degli stati  $\mathbf{s}'_{t|t-1}$ .
- K contiene la serie dei  $k_t$ ;
- SM contiene la serie degli stati smussati  $\mathbf{s}'_{t|T}$ .

Infine si avvia tutto l'insieme di comandi "cliccando" sul pulsante di esecuzione (ingranaggi). Oltre ad una finestra riportante l'esito dell'operazione, si aprirà anche la finestra icone di Gretl nella quale compariranno le matrici e vettori create.

**Script per filtro e stima: varianze ignote.** In questo caso R e Q dipendono da parametri che vanno stimati con il metodo della massima verosimiglianza. Tale procedura comporta una soluzione numerica e quindi bisogna innanzitutto fissare dei valori iniziali ai parametri.

Per non assegnare dei valori "assurdi", ci si può aiutare con il grafico a cui si aggiunge una ideale linea del livello  $m_t$ . Così facendo si può avere un'idea di massima circa le deviazioni standard di  $w_t$  (supponiamo 5) e di  $v_t$  (supp. 3). L'insieme dei comandi diventa il seguente:

```

scalar var1=5^2
scalar var2=3^2

matrix h = {1}
matrix F = {1}
matrix R = {var1}
matrix Q = {var2}
matrix s1 = {y[1]}

kalman
  obsy y
  obsymat h
  statemat F
  obsvar R
  statevar Q
  inistate s1
end kalman --diffuse

mle ll = ERR ? NA : $kalman_llt
  R[1,1] = var1
  Q[1,1] = var2

```

```

    ERR = kfilter()
    params var1 var2
end mle

matrix E V S P K
kfilter(&E, &V, &S, &P, &K)
matrix SM=ksmooth()

```

Come si può notare si sono aggiunti un blocco iniziale per i valori iniziali dei parametri e un blocco (mle... end mle) per la stima della verosimiglianza del modello, calcolata dalla funzione \$kalman\_llt. Ai nostri fini, è sufficiente sapere che in tale blocco bisogna scrivere:

- quali sono i parametri da stimare: params var1 var2;
- dove tali parametri si trovano nel modello: R[1,1] = var1 e Q[1,1] = var2.

“Cliccando” sul pulsante di esecuzione (ingranaggi) si ottiene la stima dei parametri e l’applicazione del filtro.

### A.1.2 Esempio 2: il modello ARMA

Iniziamo a considerare il modello ARMA più semplice, l’ARMA(1,1):

$$y_t = \delta + \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \quad (\text{A.2})$$

dove  $\varepsilon_t \sim NID(0, \sigma^2)$ .

Anche se non è proprio immediato, tale modello può essere riscritto come:

$$\begin{aligned} y_t &= \mu + z_t + \theta z_{t-1} \\ z_t &= \phi z_{t-1} + \varepsilon_t \end{aligned} \quad (\text{A.3})$$

dove  $\mu = \delta / (1 - \phi)$  e  $\varepsilon_t \sim NID(0, \sigma^2)$ , o in alternativa come:

$$\begin{aligned} y_t &= z_t + \theta z_{t-1} \\ z_t &= \gamma + \phi z_{t-1} + \varepsilon_t \end{aligned} \quad (\text{A.4})$$

dove  $\gamma = \delta / (1 - \theta)$  e  $\varepsilon_t \sim NID(0, \sigma^2)$ .

In generale ogni modello ARMA può essere visto come un modello MA di un modello AR<sup>4</sup>, ovvero essere scomposto in due equazioni:

---

<sup>4</sup>Ma anche come un modello AR di un modello MA!

- la prima esprime la variabile dipendente  $y_t$  in funzione di una variabile latente  $z_t$  utilizzando la parte MA del modello;
- la seconda esprime la variabile latente con la parte AR del modello.

La rappresentazione A.3 è più comune perchè il parametro  $\mu$  corrisponde alla media del processo, se questo è stazionario. Per contro la rappresentazione A.4 permette la stima dei parametri anche se non ci si è accorti che il processo non è stazionario ( $\phi = 1$ ), cosa che non può avvenire con la A.3 ( $\mu = \delta/0$  è indeterminato).

Assumendo il processo  $y_t$  stazionario, la A.3 è riconducibile alla SSF A.1 con le seguenti associazioni:

$$\mathbf{a}' = [\mu] \quad \mathbf{x}_t = [1]$$

$$\mathbf{h}' = [1 \quad \theta] \quad \mathbf{F} = \begin{bmatrix} \phi & 0 \\ 1 & 0 \end{bmatrix}$$

$$R = 0 \quad \mathbf{Q} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$$

**Stima del modello ARMA.** In questo caso la stima del modello prevede il seguente script:

```

scalar mu=mean(y)
scalar theta=0
scalar phi=0
scalar s2=var(y)

matrix a = {mu}
matrix h = {1; theta}
matrix F = {phi, 0; 1, 0}
matrix R = {0}
matrix Q = {s2, 0; 0, 0}

kalman
  obsy y
  obsxmat a
  obsymat h
  statemat F
  obsvar R
  statevar Q
end kalman

```

```

mle ll = ERR ? NA : $kalman_llt
  a[1,1] = mu
  h[2,1] = theta
  F[1,1] = phi
  Q[1,1] = s2
  ERR = kfilter()
  params mu theta phi s2
end mle

matrix E V S P K
kfilter(&E, &V, &S, &P, &K)
matrix SM=ksmooth()

```

Il primo blocco di questo script riguarda le “inizializzazioni”: non è difficile capire che tali inizializzazioni identificano  $y_t$  come un white-noise intorno alla media  $\mu$ . In genere è un inizio che permette di arrivare ad una stima, ma non sempre è così. In quei casi, occorre con pazienza “provare” altre inizializzazioni.

Nel caso di un modello ARMA(p,q), è sufficiente ricordare che tale modello può essere riscritto nella seguente SSF:

$$\begin{aligned}
y_t &= \mu + z_t + \theta_1 z_{t-1} + \dots + \theta_q z_{t-q} \\
z_t &= \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + \varepsilon_t
\end{aligned}
\tag{A.5}$$

dove  $\mu = \delta / (1 - \phi_1 - \dots - \phi_p)$  e  $\varepsilon_t \sim NID(0, \sigma^2)$ .

## A.2 Comandi per la previsione

Una volta stimati i parametri incogniti ed avviato il filtro è possibile procedere con la previsione dei dati. La script può essere standardizzato nel seguente:

```

dataset addobs 10

kalman
  obsy y
  obsymat h
  obsxmat a
  statemat F
  obsvar R
  statevar Q
end kalman
kfilter(&E, &V, &S, &P, &K)

```

```

matrix h2=2*h*h'
matrix h2[diag]=h.*h
matrix h2=vech(h2)

matrix YP = a+S*h
matrix SE =sqrt(P*h2+R)
series yp = YP
series inf = YP-1.96*SE
series sup = YP+1.96*SE

```

In pratica la previsione comporta le seguenti operazioni:

1. aggiungere tante osservazioni “vuote” quante i periodi di previsione futura (nell’esempio sopra: 10 periodi);
2. riavviare il filtro così da avere anche gli stati futuri  $\mathbf{s}_{T+k|T}$  delle componenti strutturali e i corrispondenti errori quadratici  $\mathbf{P}_{T+k|T}$ ;
3. determinare le previsioni puntuali e intervallari dai dati al punto 2.

Se prima della previsione, vi è stata la stima del modello (come di solito avviene!), non è più necessario ridefinire gli elementi del filtro, perchè questi permangono nella memoria di Gretl con i valori stimati. Di conseguenza si può subito riavviare il filtro.

Nello script precedente il blocco:

```

matrix h2=2*h*h'
matrix h2[diag]=h.*h
matrix h2=vech(h2)

```

ha lo scopo di creare un vettore i cui elementi sono il quadrato polinomiale dei coefficienti in  $\mathbf{h}$ ; se  $\mathbf{h}'=[a, b]$  allora  $\mathbf{h2}'=[a^2, 2ab, b^2]$ .

Tale vettore fa sì che  $\text{vech}(\mathbf{P}_{t+1|t})\mathbf{h2} = \mathbf{h}'\mathbf{P}_{t+1|t}\mathbf{h}$ .

I comandi successivi determinano le previsioni puntuali, gli standard error, gli estremi inferiore e superiore, con confidenza al 95% dell’intervallo di previsione.

### A.2.1 Previsione di un modello ARIMA(1,1,1)

Consideriamo il modello ARIMA(1,1,1):

$$\Delta y_t = \delta + \phi \Delta y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \quad (\text{A.6})$$



Come noto, questo è nient'altro che un modello ARMA(1,1) della differenza prima  $\Delta y_t = y_t - y_{t-1}$ .

Pertanto, in fase di stima i parametri vanno stimati utilizzando la SSF:

$$\begin{aligned}\Delta y_t &= z_t + \theta z_{t-1} \\ z_t &= \gamma + \phi z_{t-1} + \varepsilon_t\end{aligned}\tag{A.7}$$

sapendo che  $\gamma = \delta / (1 - \theta)$ .

Ma un ARIMA(1,1,1) può essere scritto anche come un ARMA(2,1):

$$y_t = \delta + (1 + \phi)y_{t-1} - \phi y_{t-2} + \varepsilon_t\tag{A.8}$$

Quindi, in fase di previsione (dopo aver stimato i parametri), è utile ricorrere alla seguente SSF:

$$\begin{aligned}y_t &= z_t + \theta z_{t-1} \\ z_t &= \gamma + (1 + \phi)z_{t-1} - \phi z_{t-2} + \varepsilon_t\end{aligned}\tag{A.9}$$

perchè con questa è possibile ottenere direttamente stime puntuali ed intervallari su  $y_{T+k}$ .

Il lettore attento potrebbe chiedersi perchè, nella A.9 si è scelta la SSF con la costante  $\gamma$  nella equazione di transizione e non quella con la costante  $\mu$  nell'equazione di misurazione. Come detto precedentemente, la rappresentazione con il parametro  $\mu$  è sempre possibile se il processo è stazionario, ma non se il processo è integrato, come in questo caso.

## Appendice B

# Derivazione analitica del Filtro di Kalman

La derivazione analitica del Filtro di Kalman, che in questa appendice verrà illustrata per il modello a livello locale (SSF più semplice) si basa sul seguente lemma:

**Lemma.** Se  $A, B, C$  sono variabili casuali (o vettori di variabili casuali) gaussiane, allora valgono le seguenti relazioni:

$$E[A|B, C] = E[A|C] + \frac{Cov[A, B|C]}{Var[B|C]}(B - E[B|C]) \quad (B.1)$$

$$Var[A|B, C] = Var[A|C] - \frac{Cov[A, B|C]^2}{Var[B|C]} \quad (B.2)$$

*Dimostrazione.* Le relazioni B.1 e B.2 si deducono dalle relazioni fondamentali circa  $E[Y|X]$  e  $Var[Y|X]$  di una normale bivariata  $X, Y$ :

$$E[Y|X] = E[Y] + \frac{Cov[X, Y]}{Var[X]}(X - E[X]) \quad (B.3)$$

$$Var[Y|X] = Var[Y] - \frac{Cov[X, Y]^2}{Var[X]} \quad (B.4)$$

tenendo conto che in generale vale la seguente relazione tra eventi aleatori:

$$(A|B, C) \equiv (A|C)|(B|C)^1 \quad (B.5)$$

Sostituendo  $A|C$  a  $Y$  e  $B|C$  a  $X$  nelle B.3 e B.4, si ottengono le relazioni del Lemma.

---

<sup>1</sup>Si pensi che  $P(A|B, C) = \frac{P(ABC)}{P(BC)} = \frac{P(AB|C)P(C)}{P(B|C)P(C)} = \frac{P(AB|C)}{P(B|C)} = P[(A|C)|(B|C)]$ .

**Derivazione del filtro.** Consideriamo il *modello a livello locale gaussiano*:

$$\begin{aligned}y_t &= m_t + \varepsilon_t \\m_{t+1} &= m_t + a_t\end{aligned}$$

dove  $\varepsilon_t \sim NID(0, R)$  e  $a_t \sim NID(0, Q)$ .

Siccome  $I_t = \{y_t, I_{t-1}\}$ , non è difficile dedurre che:

$$\begin{aligned}m_{t+1|t} &= E[m_{t+1}|I_t] = E[m_t|I_t] \\ &= E[m_t|y_t, I_{t-1}]\end{aligned}\tag{B.6}$$

$$\begin{aligned}P_{t+1|t} &= \text{Var}[m_{t+1}|I_t] = \text{Var}[m_t|I_t] + Q \\ &= \text{Var}[m_t|y_t, I_{t-1}] + Q\end{aligned}\tag{B.7}$$

Poichè tutte le grandezze in questione sono gaussiane, si possono applicare le relazioni del lemma ai risultati B.6 e B.7.

Più precisamente, applicando B.1 a B.6 si ottiene:

$$\begin{aligned}m_{t+1|t} &= E[m_t|I_{t-1}] + \frac{\text{Cov}[m_t, y_t|I_{t-1}]}{\text{Var}[y_t|I_{t-1}]}(y_t - E[m_t|I_{t-1}]) \\ &= m_{t|t-1} + k_t(y_t - m_{t|t-1})\end{aligned}\tag{B.8}$$

dove:

$$\begin{aligned}k_t &= \frac{\text{Cov}[m_t, y_t|I_{t-1}]}{\text{Var}[y_t|I_{t-1}]} \\ &= \frac{\text{Cov}[m_t, m_t + \varepsilon_t|I_{t-1}]}{\text{Var}[m_t + \varepsilon_t|I_{t-1}]} \\ &= \frac{\text{Var}[m_t|I_{t-1}] + \text{Cov}[m_t, \varepsilon_t|I_{t-1}]}{\text{Var}[m_t|I_{t-1}] + R} \\ &= \frac{P_{t|t-1}}{P_{t|t-1} + R}\end{aligned}\tag{B.9}$$

Inoltre, applicando B.2 a B.7 si ottiene:

$$\begin{aligned}P_{t+1|t} &= \text{Var}[m_t|I_{t-1}] - \frac{\text{Cov}[m_t, y_t|I_{t-1}]^2}{\text{Var}[y_t|I_{t-1}]} + Q \\ &= \text{Var}[m_t|I_{t-1}] - \left(\frac{\text{Cov}[m_t, y_t|I_{t-1}]}{\text{Var}[y_t|I_{t-1}]}\right)^2 \text{Var}[y_t|I_{t-1}] + Q \\ &= P_{t|t-1} - k_t^2(P_{t|t-1} + R) + Q\end{aligned}\tag{B.10}$$

I risultati B.8, B.10 e B.9 costituiscono il Filtro di Kalman per il modello a livello locale.

La derivazione del filtro di Kalman per il modello univariato, invariante nel tempo è solo più complessa da un punto di vista formale, ma segue esattamente la stessa logica!

# Bibliografia

- [1] Commandeur J.J.F, Koopman S.J. (2002), *An introduction to State Space Time Series Analysis*, Oxford University Press, Oxford, GB.
- [2] Durbin J., Koopman S.J. (2001), *Time series analysis by state space methods*, Oxford statistical science series, Oxford, GB.
- [3] Cottrell A., Lucchetti R. (2011), *gretl User's Guide*, Version 1.9.12. <http://gretl.sourceforge.net/>.
- [4] Hamilton J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.
- [5] Harvey A.C. (1989), *Forecasting, structural series models and the Kalman Filter*, Cambridge University Press, Cambridge, GB.
- [6] Harvey A.C., Ruiz E., Shepard N. (1994), Multivariate Stochastic Variance Models, *Review of Economic Studies*, 61, pp. 247-264.
- [7] Lucchetti R. (2011), State Space Methods in gretl, *Journal of Statistical Software*, Volume 41, Issue 11. <http://www.jstatsoft.org/>
- [8] Meinhold R.J, Singpurwalla N.D. (1983), Understanding the Kalman Filter, *The American Statistician*, Vol. 37, No. 2, pp. 123-127.