Analisi classica delle serie storiche

Approccio classico e approccio moderno

- **Processo generatore della serie storica osservata**

$$Y_t = f(t) + u_t$$

- $f(t)$: **componente *sistematica*, puramente deterministica**
- **'Legge' di evoluzione temporale del fenomeno**
- u_t : **componente *stocastica***
- **Entrambe le componenti sono <u>inosservabili</u>**

Approccio classico e approccio moderno (2)

- **Disturbi puramente casuali (processo a componenti incorrelate):** ***white noise*** ε_t

$$\begin{aligned} E(\varepsilon_t) &= 0 && \forall t \\ Var(\varepsilon_t) &= \sigma_\varepsilon^2 && \forall t \\ E(\varepsilon_r \varepsilon_s) &= 0 && r \neq s \end{aligned}$$

Approccio classico

$$u_t \equiv \varepsilon_t$$

Approccio moderno: u_t processo a componenti correlate

$$u_t = g(Y_{t-1}, Y_{t-2}, \ldots, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \varepsilon_t$$

Le componenti di una serie storica

<u>**Trend**</u> T_t

- **Tendenza di fondo, riferita ad un lungo periodo di tempo**
- **Dinamica regolare, legata all'evoluzione strutturale del sistema**

<u>**Ciclo**</u> C_t

- **Componente legata all'alternarsi di fasi di espansione e di recessione del sistema economico**
- **Fluttuazioni di periodo ampio (in genere, qualche anno)**

Le componenti di una serie storica (2)

<u>**Stagionalità**</u> S_t

- **Componente periodica i cui effetti si esauriscono nell'arco dell'anno**
- **Influenza di fattori climatici, sociali, ecc., che si ripetono nello stesso periodo dell'anno**

Modelli di combinazione delle componenti

Modello additivo

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

Modello moltiplicativo

$$Y_t = T_t C_t S_t \varepsilon_t$$

Modello misto

$$Y_t = T_t C_t S_t + \varepsilon_t$$

Da un modello moltiplicativo a quello additivo

$$\log Y_t = \log T_t + \log C_t + \log S_t + \log \varepsilon_t$$

Analisi del trend mediante funzioni matematiche

$$Y_t = f(t) + \varepsilon_t$$

- **Per semplicità, $f(t)$ solo trend**
- **Approccio parametrico: approssimazione di $f(t)$ mediante una opportuna funzione del tempo**
- **Consideriamo 2 casi**

1. **$f(t)$ lineare o linearizzabile nei parametri**
2. **$f(t)$ non lineare e non linearizzabile**

Modello lineare a *k* variabili

$$Y_t = \beta_1 + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + u_t, \qquad t = 1, \ldots, n$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_t \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & \cdots & X_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2t} & \cdots & X_{kt} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_t \\ \vdots \\ u_n \end{bmatrix}$$

- **R^2 non decresce all'aumentare del numero delle variabili esplicative**

$$Y_t = b_1 + b_2 X_{2t} + \cdots + b_k X_{kt} + e_t \quad \Rightarrow \quad R^2_{(k)}$$

$$Y_t = \tilde{b}_1 + \tilde{b}_2 X_{2t} + \cdots + \tilde{b}_k X_{kt} + \tilde{b}_{k+1} X_{k+1t} + \tilde{e}_t \quad \Rightarrow \quad R^2_{(k+1)}$$

$$\boxed{R^2_{(k+1)} \geq R^2_{(k)}}$$

Coefficiente di determinazione corretto: $\overline{R}^2$

$$R^2 = 1 - \frac{\mathbf{e'e}/n}{\mathbf{y'Ay}/n} \qquad \Rightarrow \qquad \overline{R}^2 = 1 - \frac{\mathbf{e'e}/(n-k)}{\mathbf{y'Ay}/(n-1)}$$

$$\begin{aligned} \overline{R}^2 &= 1 - \frac{n-1}{n-k}\left(1 - R^2\right) \\ &= \frac{1-k}{n-k} + \frac{n-1}{n-k} R^2 \end{aligned}$$

$$\overline{R}^2_{(k+1)} \underset{\leq}{\geq} \overline{R}^2_{(k)}$$

Trend lineare o linearizzabile

$$f(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_q t^q$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1^q \\ 1 & 2 & \cdots & 2^q \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \cdots & n^q \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_q \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

- **L'ordine q del polinomio dipende dal comportamento di fondo della serie**

$Y_t = \alpha_0 + \varepsilon_t$	**Trend costante**
$Y_t = \alpha_0 + \alpha_1 t + \varepsilon_t$	**Trend lineare**
$Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \varepsilon_t$	**Trend parabolico**

Scelta del grado del polinomio

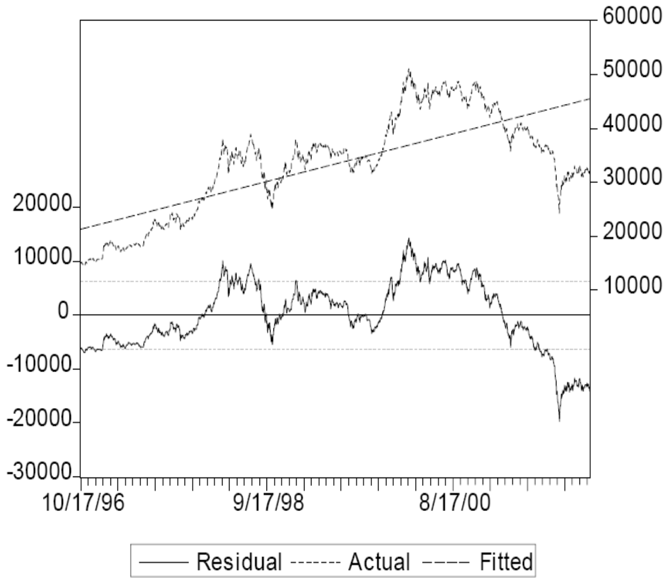
- Il polinomio stimato va usato con cautela a fini previsivi
- q non deve essere troppo elevato
- L'aggiunta di una nuova v. esplicativa non provoca un aumento della devianza residua $\Rightarrow$ $R_r^2 \leq R_{r+1}^2$
- La stessa relazione non vale per $\overline{R}^2$
- Confronto tra $\overline{R}_r^2$ e $\overline{R}_{r+1}^2$ per la scelta di q:
  - ci si arresta al grado r $(q=r)$ se $\overline{R}_r^2 \geq \overline{R}_{r+1}^2$
  - si prosegue nella ricerca se $\overline{R}_r^2 < \overline{R}_{r+1}^2$
- Attenzione alla significatività dei coefficienti associati alle potenze di t

Trend lineare su MIB30 (17/10/94-14/1/2002)

Dependent Variable: MIB30
Method: Least Squares
Date: 02/21/02 Time: 16:13
Sample: 10/17/1994 1/14/2002
Included observations: 1891

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10587.57	257.3893	41.13447	0.0000
TRENDLIN	18.58325	0.235661	78.85593	0.0000

R-squared	0.766999	Mean dependent var	28167.33
Adjusted R-squared	0.766876	S.D. dependent var	11586.18
S.E. of regression	5594.148	Akaike info criterion	20.09789
Sum squared resid	5.91E+10	Schwarz criterion	20.10375
Log likelihood	-19000.55	F-statistic	6218.258
Durbin-Watson stat	0.006870	Prob(F-statistic)	0.000000

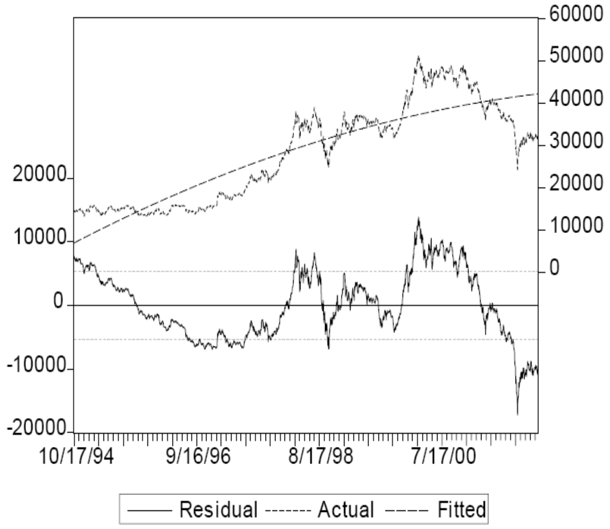


Trend quadratico su MIB30 (17/10/94-14/1/2002)

Dependent Variable: MIB30
Method: Least Squares
Date: 02/21/02 Time: 16:13
Sample: 10/17/1994 1/14/2002
Included observations: 1891

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6952.631	369.8020	18.80096	0.0000
TRENDLIN	30.10446	0.902741	33.34784	0.0000
QUAD	-0.006089	0.000462	-13.18059	0.0000

R-squared	0.786632	Mean dependent var	28167.33
Adjusted R-squared	0.786406	S.D. dependent var	11586.18
S.E. of regression	5354.690	Akaike info criterion	20.01092
Sum squared resid	5.41E+10	Schwarz criterion	20.01971
Log likelihood	-18917.32	F-statistic	3480.288
Durbin-Watson stat	0.007499	Prob(F-statistic)	0.000000

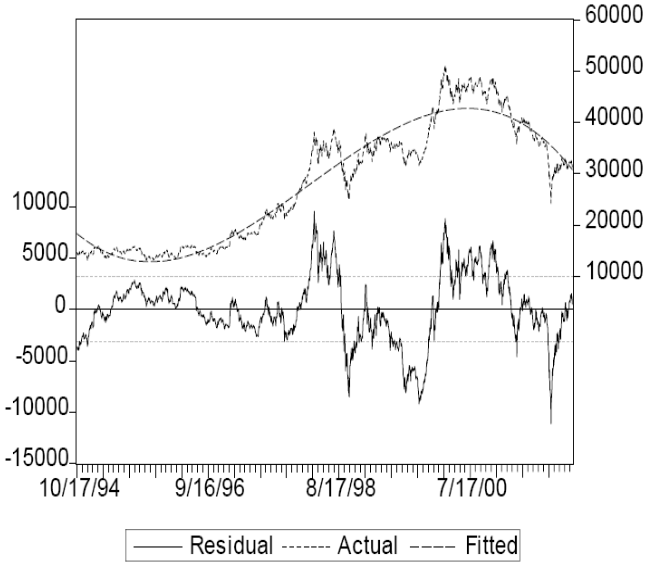


Trend cubico su MIB30 (17/10/94-14/1/2002)

Dependent Variable: MIB30
Method: Least Squares
Date: 02/21/02 Time: 16:58
Sample: 10/17/1994 1/14/2002
Included observations: 1891

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18356.80	294.1706	62.40187	0.0000
TRENDLIN	-42.13094	1.346152	-31.29733	0.0000
QUAD	0.089334	0.001653	54.03932	0.0000
CUB	-3.36E-05	5.74E-07	-58.54085	0.0000

R-squared	0.924234	Mean dependent var	28167.33
Adjusted R-squared	0.924113	S.D. dependent var	11586.18
S.E. of regression	3191.709	Akaike info criterion	18.97661
Sum squared resid	1.92E+10	Schwarz criterion	18.98834
Log likelihood	-17938.39	F-statistic	7672.838
Durbin-Watson stat	0.021064	Prob(F-statistic)	0.000000

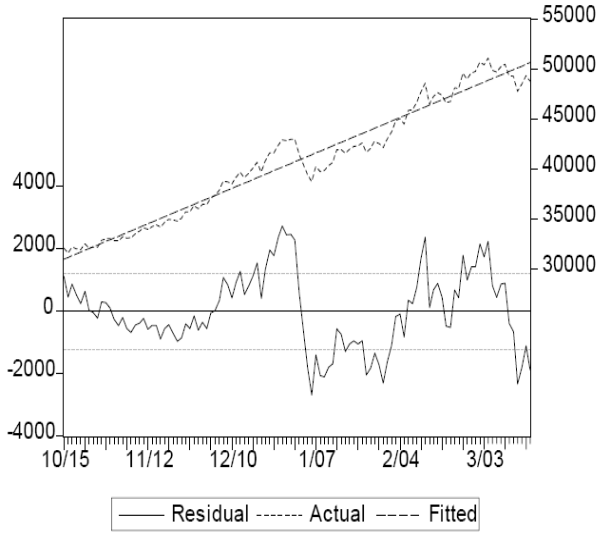


Trend lineare su MIB30 (17/10/99-20/3/2000)

Dependent Variable: MIB30
Method: Least Squares
Date: 02/21/02 Time: 15:37
Sample: 10/15/1999 3/20/2000
Included observations: 112

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-200221.5	4857.200	-41.22159	0.0000
TRENDLIN	177.1593	3.569150	49.63628	0.0000

R-squared	0.957261	Mean dependent var	40803.79
Adjusted R-squared	0.956872	S.D. dependent var	5880.395
S.E. of regression	1221.192	Akaike info criterion	17.07074
Sum squared resid	1.64E+08	Schwarz criterion	17.11928
Log likelihood	-953.9613	F-statistic	2463.760
Durbin-Watson stat	0.281784	Prob(F-statistic)	0.000000


55000
50000
45000
40000
35000
30000
4000
2000
0
-2000
-4000
10/15
11/12
12/10
1/07
2/04
3/03
Residual
Actual
Fitted