
ANALISI INDICE NASDAQ

MAGGIO 2012 - APRILE 2017

Table of Contents

1) Presentazione titolo azionario	1
2) Analisi preliminare grafici	1
3) Test DF per la verifica di stazionarietà	4
4) Analisi della distribuzione dei rendimenti	6
5) Analisi del correlogramma empirico	8
6) Stima del modello ARIMA	10
7) Analisi dei residui	12
8) Test per la presenza di effetti ARCH	15
9) Selezione e stima modello GARCH	18
10) Previsione	22
11) Conclusioni	25

Michele Sordo: il caso della time serie dell'indice $^{\wedge}$ NDX100.

1) Presentazione titolo azionario

Il NASDAQ-100 è un indice di borsa delle maggiori 100 imprese non-finanziarie quotate nel mercato borsistico NASDAQ. È un indice ponderato; il peso delle diverse società che lo compongono è basato sulla loro capitalizzazione di mercato, con alcune regole per tener conto delle influenze delle componenti maggiori. Non comprende società finanziarie, e include alcune società estere. Questi due fattori lo differenziano dall'indice S&P 500. (Fonte: <https://it.wikipedia.org/wiki/NASDAQ-100>) La serie analizzata contiene dati a frequenza giornaliera degli ultimi 5 anni: dal 1 Maggio 2012 al 28 Marzo 2017 è presente un totale di 1257 osservazioni. L'indice è quotato 5 giorni a settimana, dal lunedì al venerdì.

2) Analisi preliminare grafici

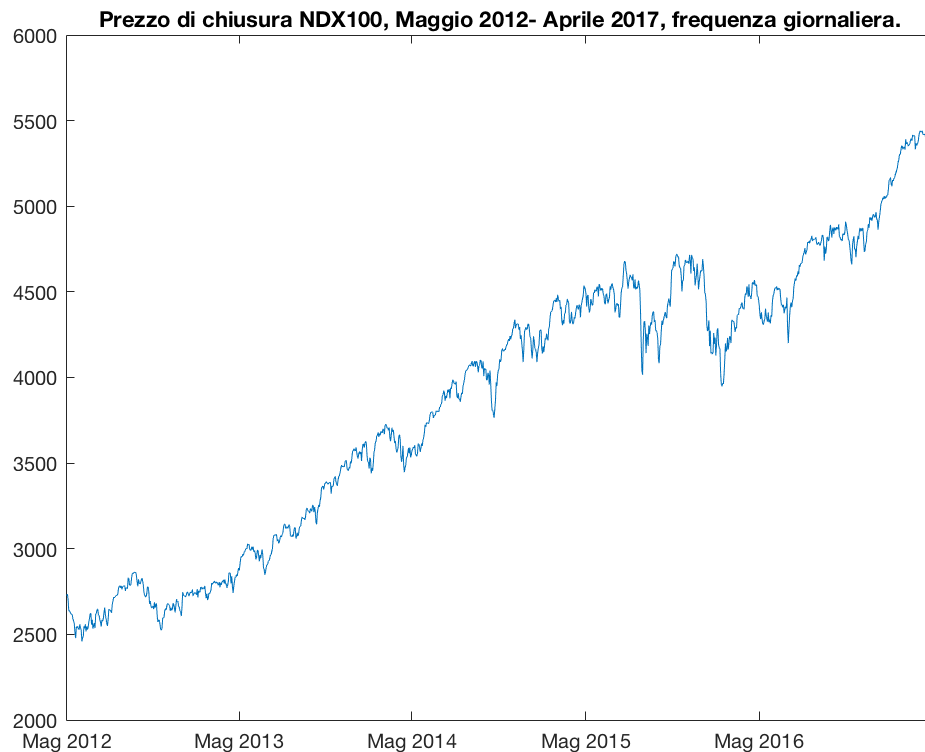
Analisi preliminare della serie storica.

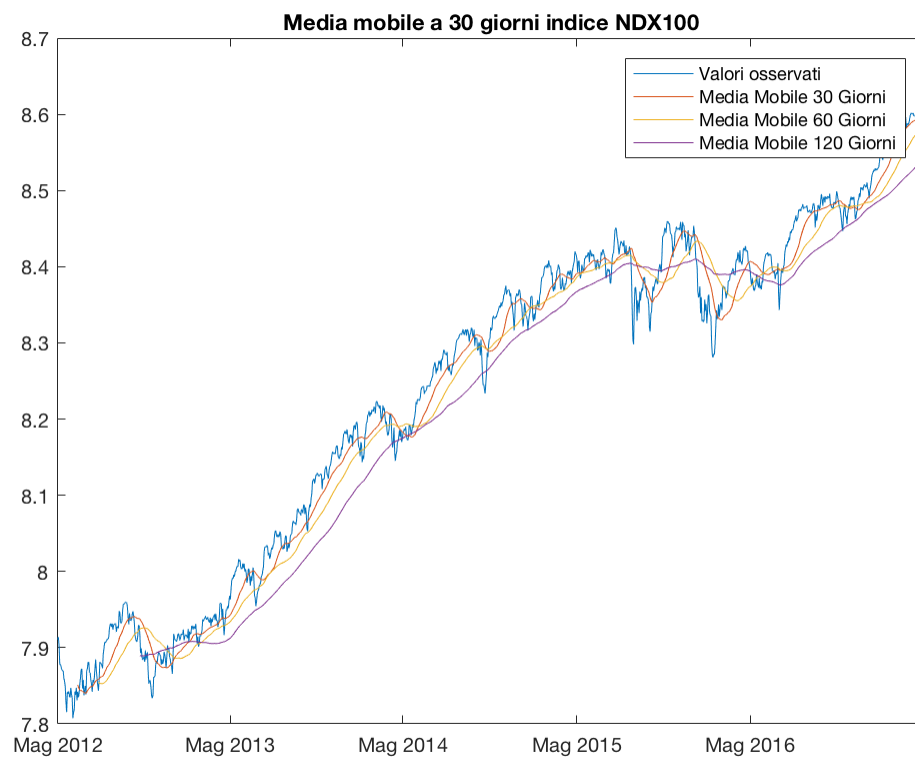
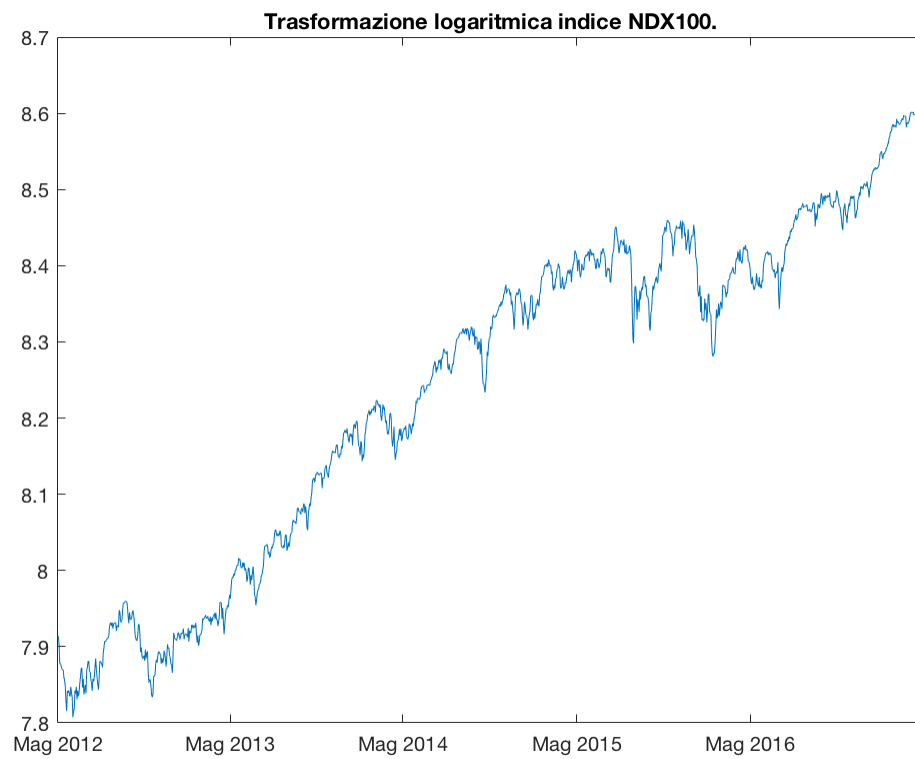
```
load('NDX.mat');           % caricamento dei dati nel workspace
y = NDX;                   % definizione della serie
T = length(y);             % definizione di T numero di osservazioni
t = (1:T);                 % definizione del vettore temporale

figure;                    % apertura nuova figura vuota
plot(y);                   % plot grafico della serie
h1 = gca;                  % definizione assi cartesiani
h1.XLim = [0,T];           % definizione asse delle ascisse
h1.XTick = [1 251 501 753 1004]; % definizione etichette asse
                             ascisse
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015', 'Mag
2016'}; % definizione titolo grafico
title('Prezzo di chiusura NDX100, Maggio 2012- Aprile 2017, frequenza
giornaliera.');
```

```
ly = log(y); % definizione serie trasformata
logaritmica
figure;
plot(ly);
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015', 'Mag
2016'};
title('Trasformazione logaritmica indice NDX100.');
```

```
ma = tsmovavg(ly,'s',30,1); % calcolo media mobile a 30 giorni
ma2 = tsmovavg(ly,'s',60,1);
ma3 = tsmovavg(ly,'s',120,1);
figure;
plot(t,ly,t,ma,t,ma2,t,ma3);
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015', 'Mag
2016'};
legend('Valori osservati','Media Mobile 30 Giorni','Media Mobile 60
Giorni','Media Mobile 120 Giorni')
title('Media mobile a 30 giorni indice NDX100');
```





I primi due grafici illustrano rispettivamente l'andamento della serie storica dell'indice NDX100 e della sua rispettiva trasformata logaritmica. Il terzo grafico è utilizzato per indicare l'andamento della media mobile a trenta, sessanta e 120 giorni della trasformata logaritmica, strumento utilizzato per l'analisi degli investimenti borsistici. La media mobile è calcolata sommando i rispettivi ritardi delle osservazioni, nel nostro caso 30 (60 e 120), e dividendo la sommatoria per il numero delle osservazioni. In questo modo, di giorno in giorno, i valori superiori a 30 (60 o 120) vengono sostituiti dalle nuove misurazioni, indicando un andamento della serie.

L'indicatore di media mobile viene utilizzato nell'ambito dell'analisi tecnica per evidenziare le tendenze sottostanti l'andamento dei prezzi eliminando l'effetto di fluttuazioni minori (transitorie). Le medie mobili possono essere calcolate con diverse metodologie. Le principali metodologie di calcolo sono le seguenti: medie mobili aritmetiche, medie mobili esponenziali e medie mobili ponderate. Eventuali cambiamenti di direzione della media mobile, oppure un incrocio tra la media mobile e la serie dei prezzi possono rappresentare segnali di acquisto o vendita, mentre la media mobile in sè rappresenta spesso un supporto o una resistenza. Fonte (<http://www.borsaitaliana.it/bitApp/glossary.bit?target=GlossaryDetail&word=Media%20Mobile>).

Questa prima analisi grafica permette di rilevare la mancanza di stazionarietà in senso debole della serie, in quanto la media presenta picchi anche distanti tra loro e la varianza si presenta contenuta in determinati periodi e più alta in altri. Al fine di verificare questa ipotesi utilizziamo il test di Dickey-Fuller.

3) Test DF per la verifica di stazionarietà

Il test di Dickey-Fuller verifica l'ipotesi nulla di presenza di una radice unitaria all'interno del processo generatore della serie storica. Se l'ipotesi nulla non viene rifiutata è necessario trasformare la serie, al fine di eliminare l'effetto della radice unitaria. Se l'ipotesi nulla viene rifiutata la serie può essere considerata stazionaria, a meno di particolari fattori. Matlab include di default 20 ritardi della variabile dipendente nella regressione ausiliaria utilizzata dalla funzione, non include un drift.

```
[h,pValue,stat] = adftest(ly);           % calcolo della statistica DF
table(h,stat,pValue)                     % output del test
if h == 1
    fprintf('Rifiuto l''ipotesi nulla');
else
    fprintf('Non rifiuto l''ipotesi nulla');
end                                       % esito del test
```

ans =

<i>h</i>	<i>stat</i>	<i>pValue</i>
<i>false</i>	<i>2.1213</i>	<i>0.99194</i>

Non rifiuto l'ipotesi nulla

L'output della funzione mostra rispettivamente: del test mostra rispettivamente

- H: l'esito dell'ipotesi alternativa (true: rifiuto l'ipotesi nulla, false: non rifiuto l'ipotesi nulla);
- STAT: valore statistica test associata;
- P-VALUE: p-value associato.

Nel nostro caso l'esito è "h = False", quindi non posso rifiutare l'ipotesi nulla di presenza di radice unitaria nella serie storica. Di conseguenza calcolo la serie delle differenze prime e ripeto il test per verificarne la stazionarietà in senso debole.

```
dy = diff(ly); % calcolo delle differenze prime
figure;
plot(dy);
h2 = gca;
h2.XLim = [0,T];
h2.XTick = [1 251 501 753 1004];
h2.XTickLabel = {'Mag 2012', 'Mag 2013', 'Mag 2014', 'Mag 2015', 'Mag 2016'};
title('Differenze prime NDX100');

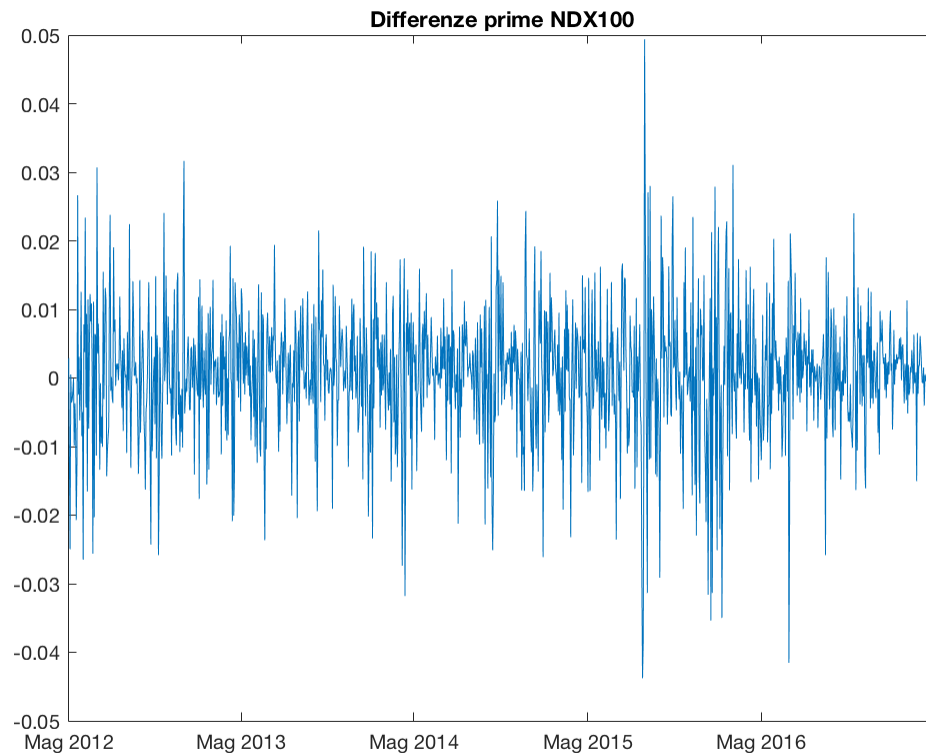
[h,pValue,stat] = adftest(dy);
table(h,stat,pValue)
if h == 1
    fprintf('Rifiuto l''ipotesi nulla');
else
    fprintf('Non rifiuto l''ipotesi nulla');
end
```

*Warning: Test statistic #1 below tabulated critical values:
minimum p-value = 0.001 reported.*

ans =

<i>h</i>	<i>stat</i>	<i>pValue</i>
<i>true</i>	<i>-34.491</i>	<i>0.001</i>

Rifiuto l'ipotesi nulla



L'esito del test Dickey-Fuller, in questo caso, è positivo quindi rifiuto l'ipotesi di presenza di una seconda radice unitaria e traggio la conclusione che la serie storica dell'indice NDX100 si possa considerare stazionaria in senso debole nelle differenze prime. Il grafico associato alla serie delle differenze prime conferma che la serie è almeno stazionaria in media e mostra che la varianza cambia nei diversi periodi, risultando più marcata in alcuni casi e più contenuta in altri. Al fine dell'analisi della variabilità verrà affrontata nella sezione 8. Consideriamo quindi la serie come "integrata di ordine 1", ossia $I(1)$.

4) Analisi della distribuzione dei rendimenti

La serie delle differenze prime corrisponde alla serie dei rendimenti dell'indice NDX100. Al fine di scegliere e stimare il modello migliore si utilizza l'approccio di Box-Jenkins, definendo il processo stocastico di generazione della serie storica.

```
Media = mean(dy);
Varianza = var(dy);
Curtosi = kurtosis(dy);
Asimmetria = skewness(dy);
table(Media, Varianza, Curtosi, Asimmetria)
figure
subplot(2,2,1)
histogram(dy,60)
title('Istogramma serie dei rendimenti')

[f,dyi] = ksdensity(dy);
subplot(2,2,2)
plot(dyi,f)
```

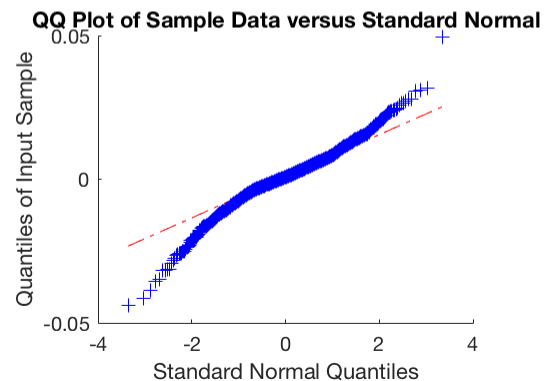
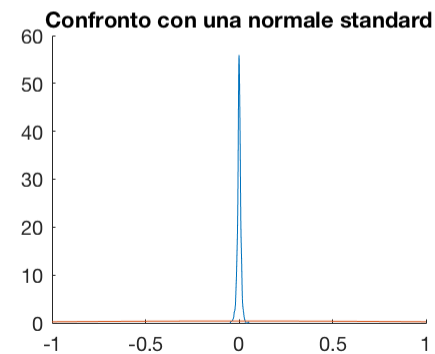
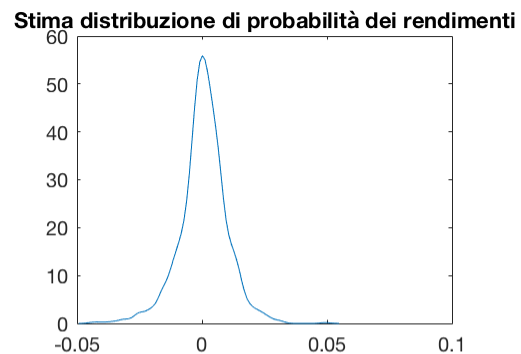
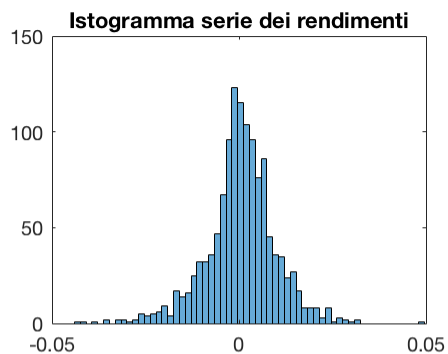
```
title('Stima distribuzione di probabilità dei rendimenti')
```

```
subplot(2,2,3)  
hold on  
plot(dy,f)  
x = (-1:.1:1);  
norm = normpdf(x,0,1);  
plot(x,norm)  
hold off  
title('Confronto con una normale standard')
```

```
subplot(2,2,4)  
qqplot(dy)
```

ans =

<i>Media</i>	<i>Varianza</i>	<i>Curtosi</i>	<i>Asimmetria</i>
<hr/>	<hr/>	<hr/>	<hr/>
0.00057058	8.9686e-05	5.2899	-0.32627



I momenti calcolati mostrano:

- una media prossima allo zero, pari a "0.00057058".

- una varianza prossima allo zero, pari a "8.9686e-05".

La distribuzione è leptocurtica e asimmetrica verso sinistra. Il confronto tra la stima della distribuzione di probabilità dei rendimenti ed il confronto con la normale standard mostra che, sebbene le osservazioni siano concentrate intorno alla media, la distribuzione presenta pesanti code da entrambe le parti, con una particolare accentuazione verso sinistra. L'analisi è confermata dal qq-plot, dove si nota la presenza di asimmetria e la conferma dell'ipotesi delle code più pesanti.

5) Analisi del correlogramma empirico

Al fine di identificare il processo generatore della serie tramite l'approccio di Box-Jenkins si calcolano e si rappresentano attraverso i rispettivi correlogrammi empirici:

- ACF: funzione di autocorrelazione
- PACF: funzione di autocorrelazione parziale.

I correlogrammi empirici presentando una prima analisi grafica della relazione tra la serie e le osservazioni.

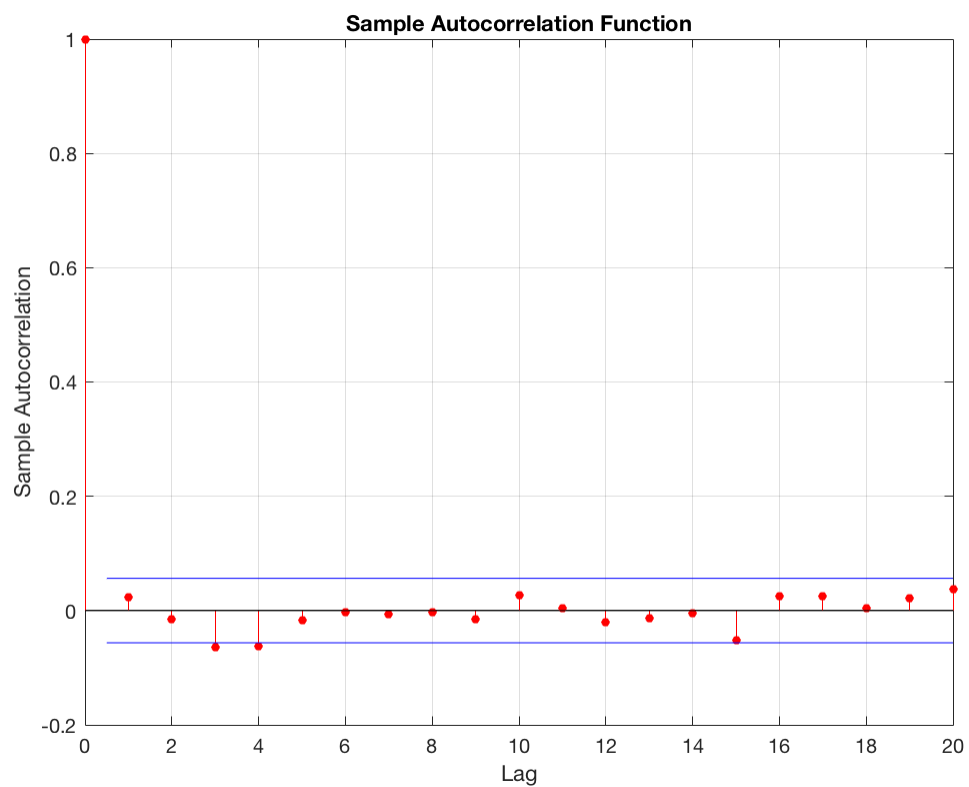
```
K = (0:20)'; % definizione vettore ritardi
ACF = autocorr(dy); % calcolo ACF
PACF = parcorr(dy); % calcolo PACF
table(K,ACF,PACF)
figure
autocorr(dy) % plot correlogramma ACF
figure
parcorr(dy) % plot correlogramma PACF
```

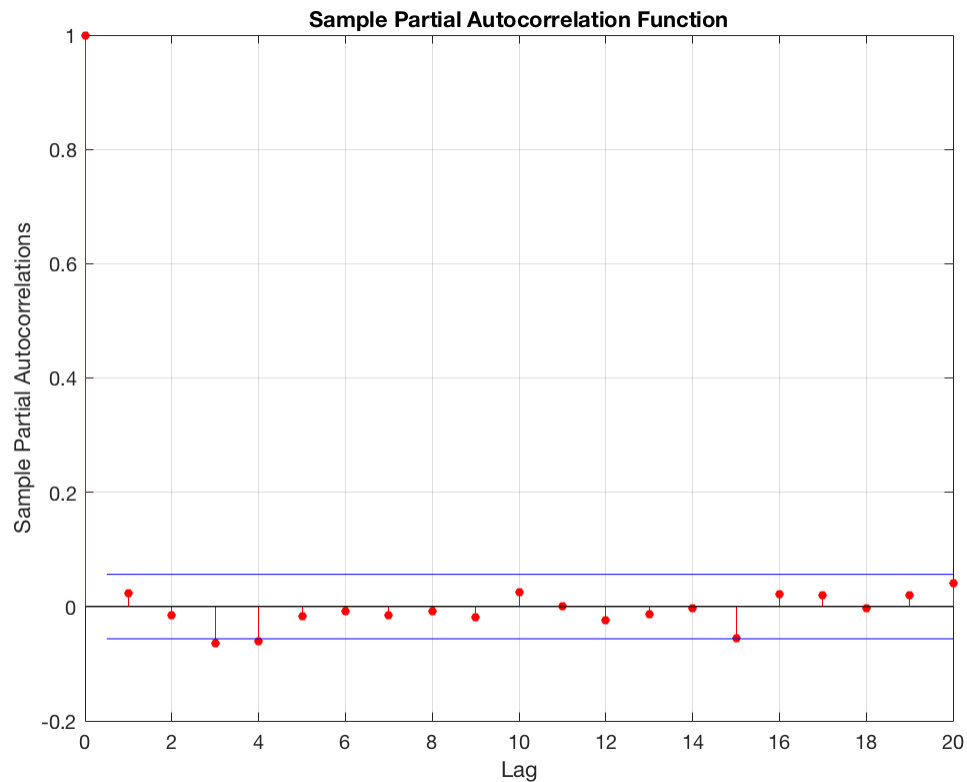
ans =

<i>K</i>	<i>ACF</i>	<i>PACF</i>
0	1	1
1	0.022853	0.022853
2	-0.015352	-0.01589
3	-0.064379	-0.063775
4	-0.06288	-0.060681
5	-0.016464	-0.016268
6	-0.0028925	-0.0085681
7	-0.0064258	-0.015068
8	-0.0023477	-0.008392
9	-0.0155	-0.018985
10	0.026972	0.02482
11	0.0038611	-0.00017703
12	-0.021	-0.024464
13	-0.013686	-0.013015
14	-0.0046123	-0.0022273
15	-0.05162	-0.054756
16	0.024571	0.021842
17	0.025596	0.020668
18	0.0048912	-0.0036883

ANALISI INDICE NASDAQ
MAGGIO 2012 - APRILE 2017

19	0.021881	0.019745
20	0.037173	0.040337





I correlogrammi empirici non permettono di derivare direttamente il modello stocastico migliore per l'analisi della serie storica ma permettono di notare i potenziali effetti dei ritardi, osservando per quali K essi superino l'intervallo di confidenza. Nel nostro caso i ritardi che mostrano potenziali effetti sono il terzo e il quarto. Tramite questa informazioni si procede a stimare il modello ARIMA migliore e più parsimonioso.

6) Stima del modello ARIMA

Questa sezione presenta i risultati dell'analisi sulle stime dei diversi modelli ARIMA con le differenti specificazioni. Per trovare la specificazione che meglio possa descrivere il processo stocastico generatore della serie storica è stato utilizzato un metodo "trial and error". Per motivi di spazio e chiarezza si allega il procedimento contenente i confronti tra le differenti serie analizzate a fine analisi. I risultati più soddisfacenti si trovano per i modelli ARIMA(4,1,0) e ARIMA(0,1,4), ottenuti includendo unicamente il terzo ed il quarto ritardo, in quanto i precedenti possono essere considerati non significativamente diversi da zero.

```
Mdl = arima('ArLags',[3 4],'D',1,'MaLags',[]);
[AR4EstMdl, AR4EstParamCov, AR4logL, AR4info] = estimate(Mdl,ly);

Mdl = arima('ArLags',[],'D',1,'MaLags',[3 4]);
[MA4EstMdl, MA4EstParamCov, MA4logL, MA4info] = estimate(Mdl,ly);
```

ARIMA(4,1,0) Model:

Conditional Probability Distribution: Gaussian

Standard

t

ANALISI INDICE NASDAQ
MAGGIO 2012 - APRILE 2017

<i>Parameter</i>	<i>Value</i>	<i>Error</i>	<i>Statistic</i>
Constant	0.0006686	0.000274536	2.43538
AR{3}	-0.0629839	0.0222096	-2.83589
AR{4}	-0.0614709	0.0223005	-2.75648
Variance	8.86253e-05	2.50313e-06	35.4058

ARIMA(0,1,4) Model:

Conditional Probability Distribution: Gaussian

<i>Parameter</i>	<i>Value</i>	<i>Standard Error</i>	<i>t Statistic</i>
Constant	0.000569214	0.000245826	2.31552
MA{3}	-0.0644678	0.0226558	-2.84553
MA{4}	-0.0635488	0.0231999	-2.73918
Variance	8.88093e-05	2.50972e-06	35.3862

Al fine di scegliere tra il processo ARIMA(4,1,0) ed il processo ARIMA(0,1,4) è utile ricorrere ai test AIC e BIC.

```
[AR4aic,AR4bic]=aicbic(AR4logL,2,T);           % calcolo AIC BIC
ARIMA(4,1,0)
[MA4aic,MA4bic]=aicbic(MA4logL,2,T);           % calcolo AIC BIC
ARIMA(0,1,4)

table(AR4aic,AR4bic,MA4aic,MA4bic)

if AR4aic < MA4aic && AR4bic < MA4bic
    fprintf('Modello ARIMA(4,1,0) è preferibile.');
```

```
end

ans =

    AR4aic    AR4bic    MA4aic    MA4bic
    -----    -----    -----    -----
    -8155.1    -8144.8    -8155.3    -8145
```

Modello ARIMA(0,1,4) è preferibile.

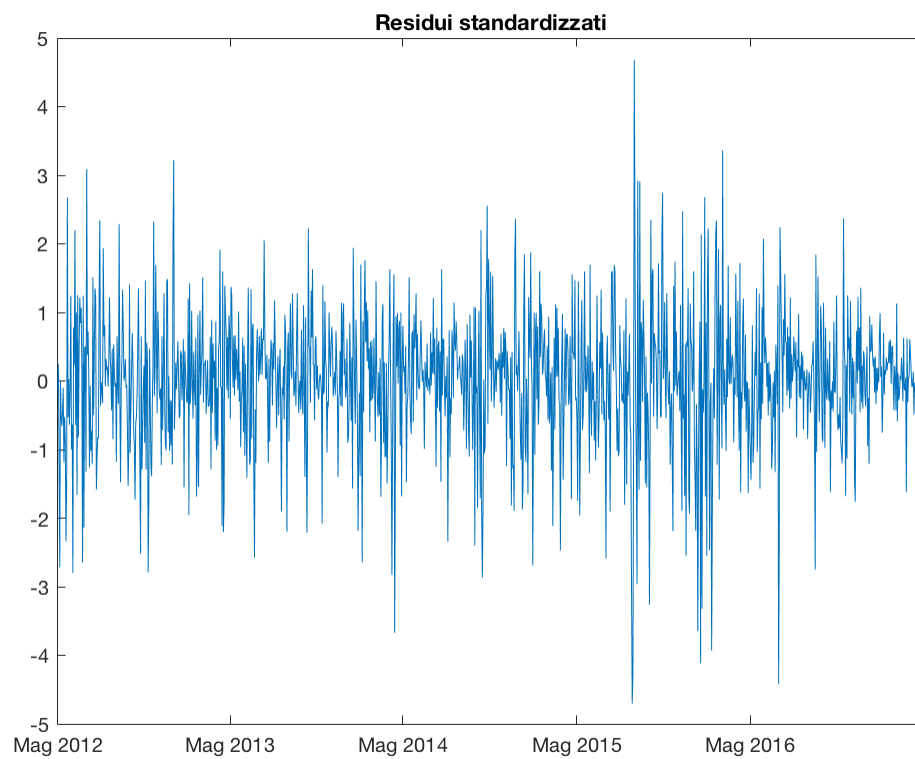
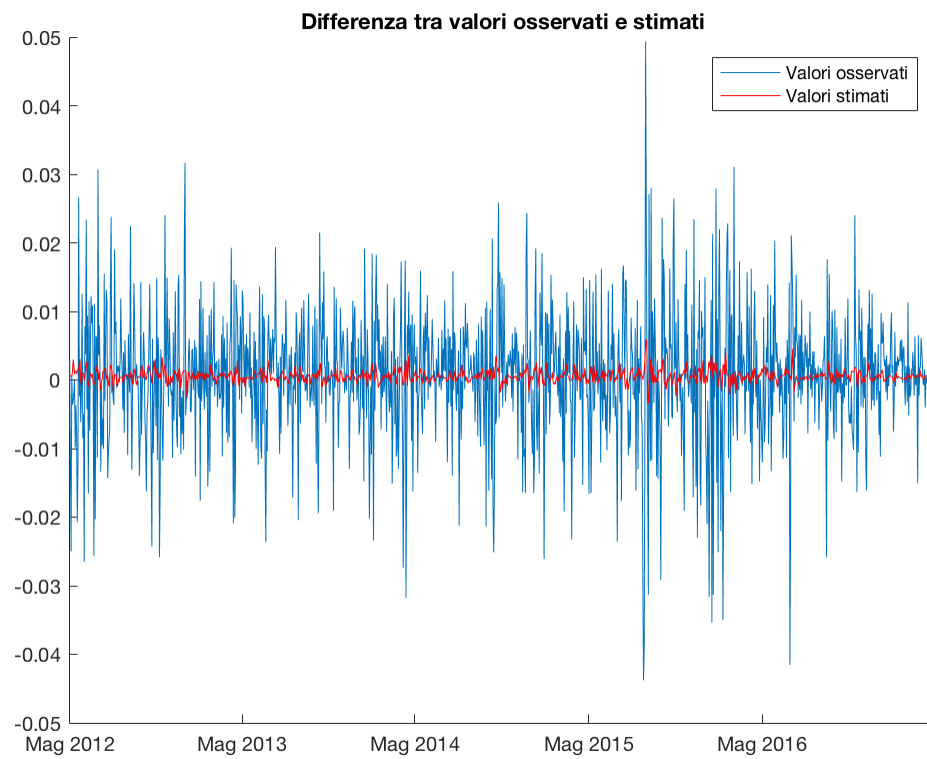
Dopo aver aggiunto e eliminato i diversi ritardi ed aver confrontato le due migliori ipotesi ottenute tramite il calcolo e confronto dei rispettivi valori dei criteri informativi di Akaike e Bayesiano, posso concludere la miglior specificazione disponibile per per il processo stocastico generatore della serie delle differenze prime. La specificazione trovata è un modello ARIMA(0,1,4), dove i coefficienti intermedi 1 e 2 sono uguali a zero.

7) Analisi dei residui

Si procede valutando la qualità del modello scelto, analizzando la distribuzione dei residui della regressione.

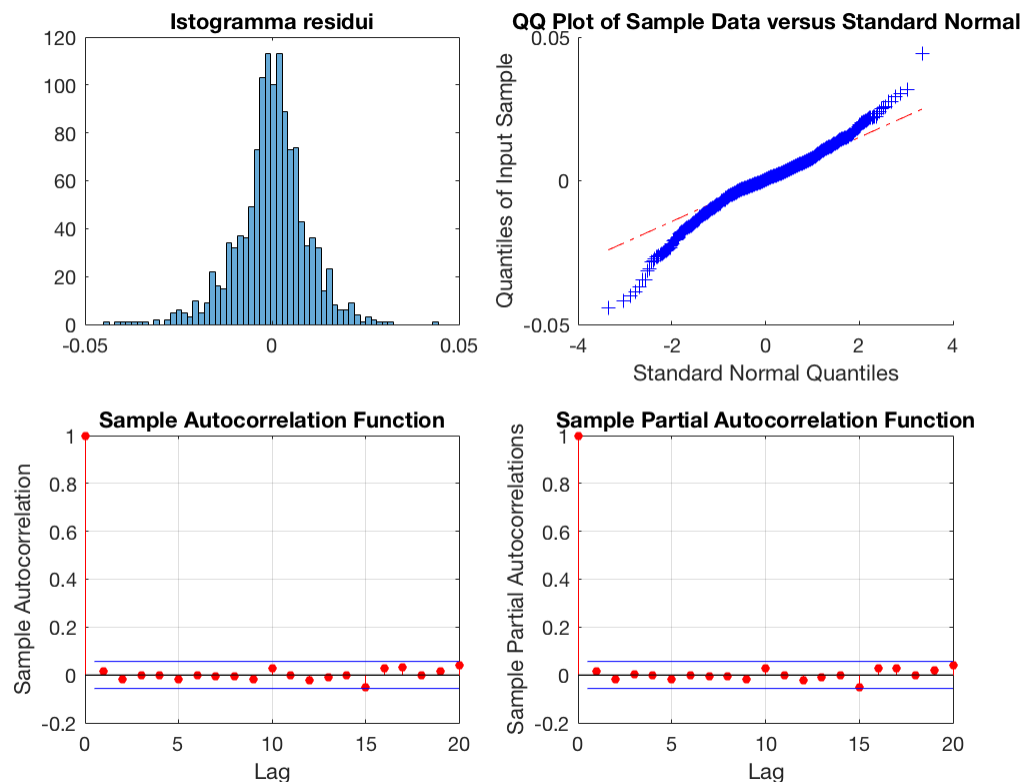
```
res = infer(MA4EstMdl,ly);
sy = (dy-res(2:end));
figure
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015', 'Mag
2016'};
hold on
plot(dy)
plot(sy,'r')
legend('Valori osservati','Valori stimati')
title('Differenza tra valori osservati e stimati')
hold off

figure
plot(res./sqrt(MA4EstMdl.Variance))           % plot dei residui
standardizzati
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015', 'Mag
2016'};
title('Residui standardizzati')
```



Dal grafico delle differenze tra valori osservati e stimati e dal grafico dei residui standardizzati si può notare che la variabilità che effettivamente il modello spiega è piuttosto bassa: la conclusione è che il nostro modello riesca a stimare in modo consistente i parametri che influenzano la serie ma che la spiegazione della maggior parte dei suoi movimenti si dovrebbe ricercare in variabili esogene o elementi particolari che caratterizzano l'origine della serie.

```
figure
subplot(2,2,1)
histogram(res,60)                                % istogramma dei residui
title('Istogramma residui')
subplot(2,2,2)
qqplot(res)                                       % qqplot dei residui
subplot(2,2,3)
autocorr(res)                                    % correlogramma
subplot(2,2,4)
parcorr(res)
```



Come la distribuzione dei valori osservati anche la distribuzione dei residui standardizzati è leptocurtica, asimmetrica verso sinistra e la maggior parte delle osservazioni sono concentrate intorno alla media. L'analisi è confermata anche in questo caso dal qq-plot, dove si nota la presenza di asimmetria a sinistra e la conferma dell'ipotesi delle code più pesanti.

Si procede quindi a confermare questa ipotesi con il test di Jung-Box:

```
[h,p,Qstat,crit] = lbqtest(res,'lags',[1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20]);
names = {'K';'H';'pvalue';'Qstat';'Crit'};
```

```
table(K(2:end),h',p',Qstat',crit','VariableNames',names) %  
creazione tabella output
```

ans =

<i>K</i>	<i>H</i>	<i>pvalue</i>	<i>Qstat</i>	<i>Crit</i>
1	false	0.55337	0.35132	3.8415
2	false	0.70408	0.70173	5.9915
3	false	0.87275	0.70195	7.8147
4	false	0.95102	0.70246	9.4877
5	false	0.95309	1.1114	11.07
6	false	0.98037	1.1262	12.592
7	false	0.99194	1.1546	14.067
8	false	0.99679	1.185	15.507
9	false	0.99595	1.642	16.919
10	false	0.991	2.4917	18.307
11	false	0.99589	2.4917	19.675
12	false	0.99377	3.2215	21.026
13	false	0.99644	3.3331	22.362
14	false	0.99832	3.3344	23.685
15	false	0.96293	6.8014	24.996
16	false	0.95791	7.677	26.296
17	false	0.94478	8.8523	27.587
18	false	0.96306	8.8526	28.869
19	false	0.97052	9.1718	30.144
20	false	0.93389	11.444	31.41

Il test di Jung-Box verifica l'ipotesi nulla di congiunta uguaglianza a zero delle prime K autocorrelazioni dei residui. Nella tabella precedente è illustrato l'output del test per le prime 20 statistiche calcolate su tale ipotesi. Come si nota l'esito è "False" per tutte le prime 20 autocorrelazioni dei residui (colonna H) e non possiamo quindi rifiutare l'ipotesi nulla che esse siano tutte congiuntamente nulle.

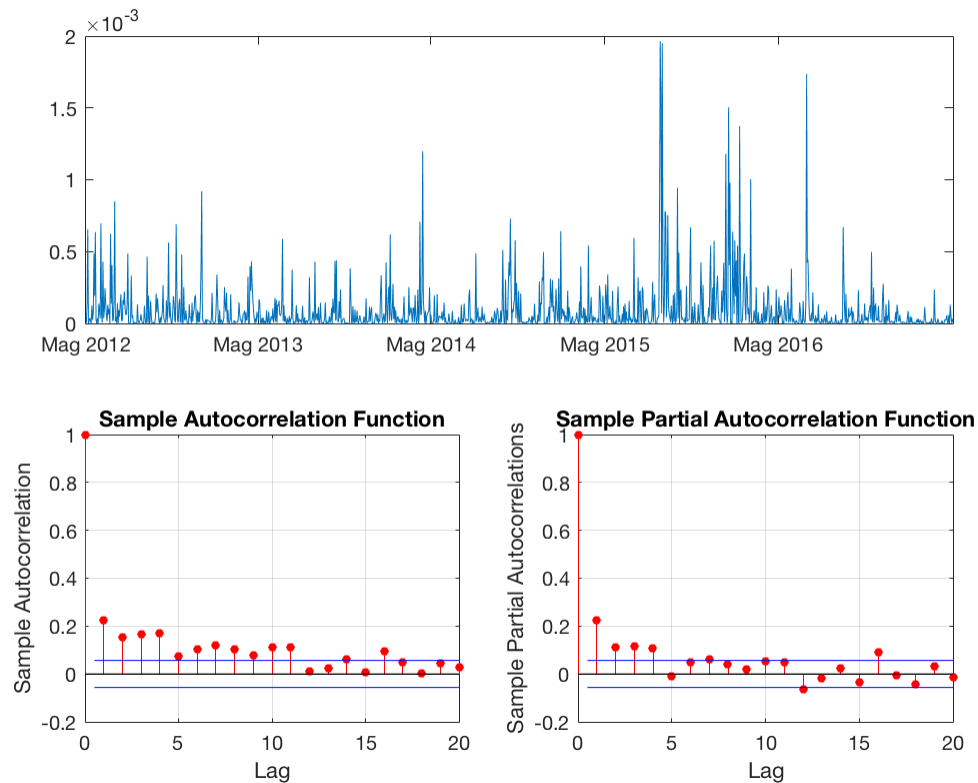
L'esito di questa analisi indica che il nostro modello ARIMA(0,1,4) è adatto a stimare un modello per l'indice NDX100 dal 1 Maggio 2012 al 28 Aprile 2017.

8) Test per la presenza di effetti ARCH

Durante l'analisi precedente abbiamo più volte notato la possibile presenza di eteroschedasticità negli errori. Al fine di comprenderne meglio la dinamica, si procede con un'analisi sulla serie dei residui al quadrato, verificata con i test di Ljung-Box e Engle.

```
figure;  
subplot(2,2,[1,2]);  
plot((res-mean(res)).^2);  
h1 = gca;  
h1.XLim = [0,T];  
h1.XTick = [1 251 501 753 1004];  
h1.XTickLabel = {'Mag 2012','Mag 2013','Mag 2014','Mag 2015','Mag  
2016'};  
subplot(2,2,3)
```

```
autocorr((res).^2);  
subplot(2,2,4)  
parcorr((res).^2);
```



I correlogrammi mostrano la presenza di autocorrelazione degli errori al quadrato, quindi di eteroschedasticità condizionale autoregressiva.

```
fprintf('1 RITARDO\n\n');  
fprintf('-- Test di Ljung-Box --\n');  
[h,p,Qstat,crit] = lbqtest(res.^2,'lags',10);  
table(h,Qstat,crit,p,'rownames',{'lbqtest'})  
if h == 1  
    fprintf('Rifiuto l''ipotesi nulla\n');  
else  
    fprintf('Non rifiuto l''ipotesi nulla\n');  
end  
fprintf('-- Test di Engle--\n');  
[h,p] = archtest(res,'lags',10);  
table(h,p,'rownames',{'Archtest'})  
if h == 1  
    fprintf('Rifiuto l''ipotesi nulla\n');  
else  
    fprintf('Non rifiuto l''ipotesi nulla\n');  
end
```

1 RITARDO

-- Test di Ljung-Box --

ans =

	<i>h</i>	<i>Qstat</i>	<i>crit</i>	<i>p</i>
	—	—	—	—
<i>lbqtest</i>	<i>true</i>	<i>238.94</i>	<i>18.307</i>	<i>0</i>

Rifiuto l'ipotesi nulla

-- Test di Engle--

ans =

	<i>h</i>	<i>p</i>
	—	—
<i>Archtest</i>	<i>true</i>	<i>0</i>

Rifiuto l'ipotesi nulla

Sia il test di Ljung-Box che il test di Engle rifiutano l'ipotesi nulla di indipendenza degli errori al quadrato al primo ritardo.

```
fprintf('10 RITARDI\n\n');
fprintf('-- Test di Ljung-Box --\n');
[h,p,Qstat,crit] = lbqtest(res.^2,'lags',10);
table(h,Qstat,crit,p,'rownames',{'lbqtest'})
if h == 1
    fprintf('Rifiuto l''ipotesi nulla\n');
else
    fprintf('Non rifiuto l''ipotesi nulla\n');
end
fprintf('-- Test di Engle--');
[h,p] = archtest(res,'lags',10);
table(h,p,'rownames',{'Archtest'})
if h == 1
    fprintf('Rifiuto l''ipotesi nulla\n');
else
    fprintf('Non rifiuto l''ipotesi nulla\n');
end
```

10 RITARDI

-- Test di Ljung-Box --

ans =

	<i>h</i>	<i>Qstat</i>	<i>crit</i>	<i>p</i>
	—	—	—	—
<i>lbqtest</i>	<i>true</i>	<i>238.94</i>	<i>18.307</i>	<i>0</i>

Rifiuto l'ipotesi nulla

-- Test di Engle--

ans =

	<i>h</i>	<i>p</i>
	—	—
<i>Archtest</i>	<i>true</i>	<i>0</i>

Rifiuto l'ipotesi nulla

Anche i test di Ljung-Box ed Engle, ripetuti su 10 ritardi, rifiutano l'ipotesi nulla di indipendenza degli errori al quadrato.

9) Selezione e stima modello GARCH

Date le informazioni riportate dai test precedenti, in questa sezione si tenta di trovare un modello per scrivere l'andamento eteroschedastico della varianza condizionale. Al fine di trovare un modello sufficientemente buono sono stati stimati diversi modelli ARCH, GARCH ed EGARCH, specificando di volta in volta le combinazioni dei ritardi. Il modello migliore risultato da questa analisi è il modello GARCH(1,1).

```
Mdl = garch(1,1);
VarMdl = estimate(Mdl,res);
```

GARCH(1,1) Conditional Variance Model:

Conditional Probability Distribution: Gaussian

<i>Parameter</i>	<i>Value</i>	<i>Standard Error</i>	<i>t Statistic</i>
-----	-----	-----	-----
<i>Constant</i>	<i>6.18045e-06</i>	<i>1.61814e-06</i>	<i>3.81948</i>
<i>GARCH{1}</i>	<i>0.801731</i>	<i>0.0259723</i>	<i>30.8686</i>
<i>ARCH{1}</i>	<i>0.130128</i>	<i>0.0173336</i>	<i>7.50726</i>

Entrambi i coefficienti del modello sono significativamente diversi da zero. Si prosegue con l'analisi dei residui standardizzati tramite la varianza condizionata generata dal modello GARCH(1,1). Sia il test di Ljung-Box che il test di Engle sono stati eseguiti da 1 a 10 ritardi, ottenendo lo stesso risultato finale:

```
cv = infer(VarMdl,res);      % generazione della varianza condizionata
stres = (res./sqrt(cv));    % standardizzazione dei residui
[h,p,Qstat,crit] = lbqtest(cv.^2,'lags',10);
table(h,Qstat,crit,p,'rownames',{'lbqtest'})
if h == 1
    fprintf('Rifiuto l''ipotesi nulla');
else
    fprintf('Non rifiuto l''ipotesi nulla');
end
[h,p] = archtest(cv,'lags',10);      %Test di Engle
table(h,p,'rownames',{'Archtest'})
if h == 1
    fprintf('Rifiuto l''ipotesi nulla');
else
    fprintf('Non rifiuto l''ipotesi nulla');
end
```

```
ans =
```

	<i>h</i>	<i>Qstat</i>	<i>crit</i>	<i>p</i>
	—	—	—	—
<i>lbqtest</i>	<i>true</i>	4081.5	18.307	0

Rifiuto l'ipotesi nulla

```
ans =
```

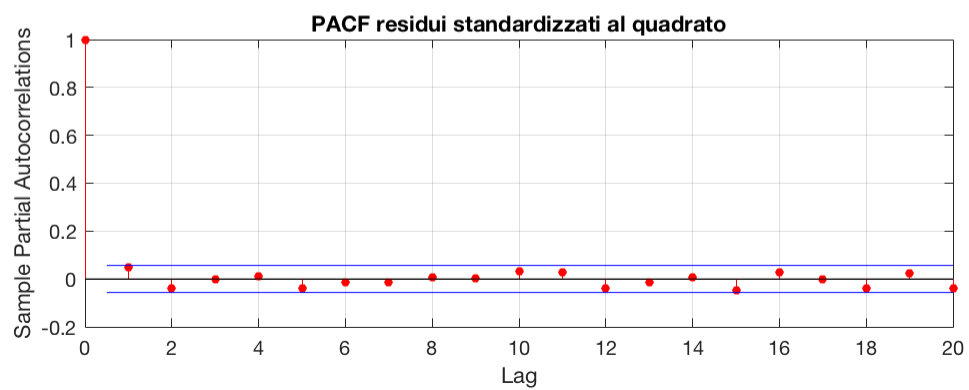
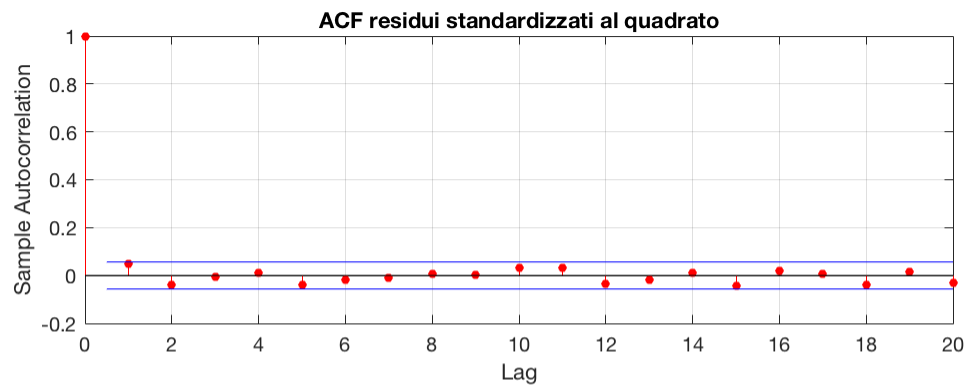
	<i>h</i>	<i>p</i>
	—	—
<i>Archtest</i>	<i>true</i>	0

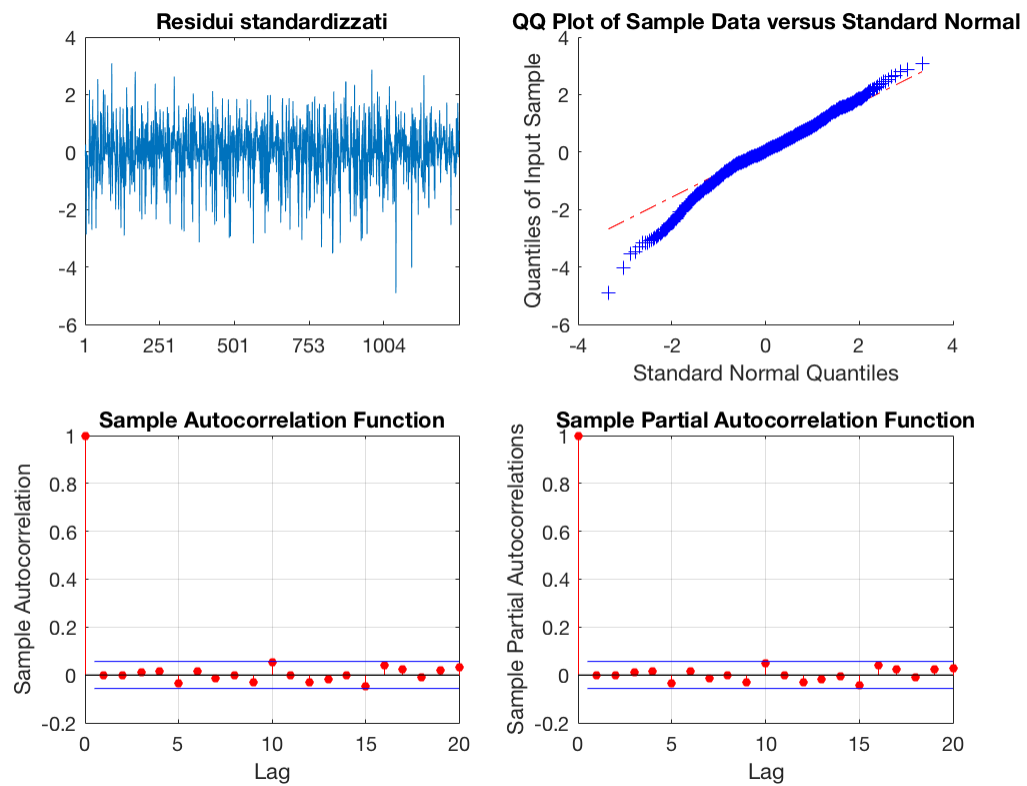
Rifiuto l'ipotesi nulla

Tutti i modelli ARCH, GARCH ed EGARCH, con tutti i diversi ritardi selezionati, rifiutano i test di Ljung-Box ed Engle. Il modello GARCH(1,1) è stato quindi selezionato perché migliore di tutti i modelli ARCH, più parsimonioso tra i modelli GARCH e perché presentava correlogrammi migliori rispetto al rispettivo modello EGARCH.

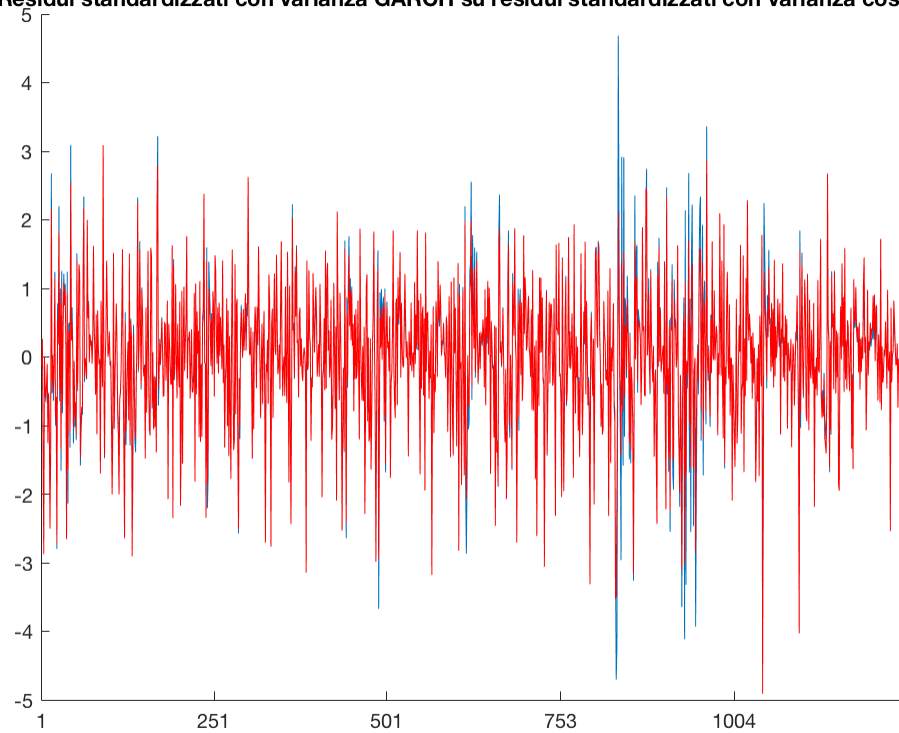
```
figure;
subplot(2,1,1)
autocorr((stres).^2);
title('ACF residui standardizzati al quadrato');
subplot(2,1,2)
parcorr((stres).^2);
title('PACF residui standardizzati al quadrato');
figure
subplot(2,2,1)
plot(stres)
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
title('Residui standardizzati');
subplot(2,2,2)
qqplot(stres)
subplot(2,2,3)
autocorr(stres)
subplot(2,2,4)
parcorr(stres)
figure
h1 = gca;
h1.XLim = [0,T];
h1.XTick = [1 251 501 753 1004];
hold on
plot(res./sqrt(MA4EstMdl.Variance))
plot(stres,'r');
hold off
title('Residui standardizzati con varianza GARCH su residui
standardizzati con varianza costante');
figure
```

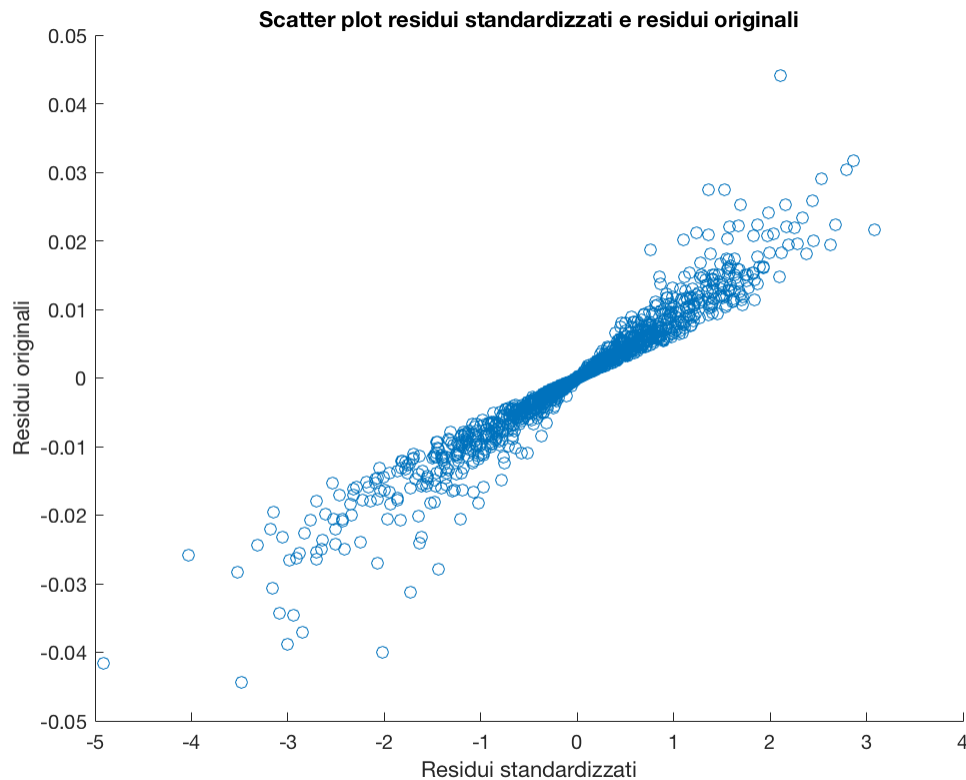
```
h1 = gca;  
h1.XLim = [0,T];  
h1.XTick = [1 251 501 753 1004];  
scatter(stres,res);  
ylabel('Residui originali');  
xlabel('Residui standardizzati');  
title('Scatter plot residui standardizzati e residui originali')
```





Residui standardizzati con varianza GARCH su residui standardizzati con varianza costante





I correlogrammi dei residui standardizzati al quadrato mostrano un miglioramento della situazione, in quanto non sono visibili ritardi problematici.

I quattro grafici successivi illustrano che la standardizzazione mediante varianza condizionata ha un effetto quasi impercettibile sul risultato. Negli ultimi due grafici in particolare si può notare l'inefficacia del modello GARCH, considerazione che ci porta a concludere che il fenomeno dell'eteroschedasticità condizionale autoregressiva è piuttosto contenuto ed i cambiamenti di variabilità sono forse da ricercarsi in elementi esogeni al modello.

10) Previsione

In questa ultima sezione si utilizza il modello scelto al fine di prevedere il valore atteso condizionato e la varianza attesa condizionata, valutandone la precisione. Al fine della previsione si dividono le osservazioni in due gruppi, il primo contenente le osservazioni fino a 30 giorni dall'ultima misurazione, il secondo contenente le ultime 30 osservazioni al fine di testare la previsione.

```
[yf, vf] = forecast(MA4EstMdl,30,'Y0',ly(1:1227));           % valore  
               atteso condizionato  
  
figure  
h1 = gca;  
h1.XLim = [0,T];  
h1.XTick = [1 251 501 753 1004];  
h1.XTickLabel = {'Giu 2011','Giu 2012','Giu 2013','Giu 2014','Giu  
                2015'};  
hold on
```

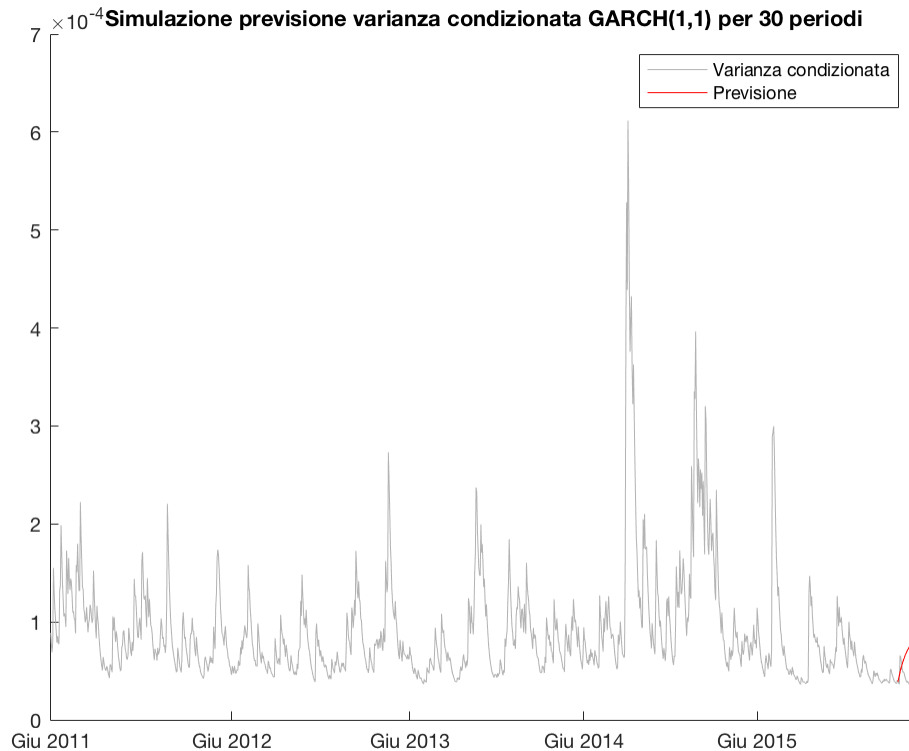
```
h2 = plot(ly, 'Color', [.7, .7, .7]);  
h3 = plot(1228:1257, yf, 'b', 'LineWidth', 2);  
h4 = plot(1228:1257, yf + 1.96*sqrt(vf), 'r:', 'LineWidth', 2);  
plot(1228:1257, yf - 1.96*sqrt(vf), 'r:', 'LineWidth', 2);  
legend([h2 h3 h4], 'Valori osservati', 'Previsione', 'Intervallo di  
confidenza al 95%', 'Location', 'NorthWest');  
title('Simulazione previsione per 30 periodi e intervalli di  
confidenza al 95%')  
hold off
```



Il grafico mostra la previsione del valore atteso condizionato alla serie dei logaritmi del prezzo di chiusura dell'indice NDX100. L'intervallo di confidenza si espande velocemente a causa della non stazionarietà della serie, mostrando quanto una previsione su un orizzonte temporale maggiore di 30 giorni sia pressoché inutile. Nel breve periodo la misurazione risulta comunque non molto precisa: anche se nel nostro caso i valori della previsioni sono molto buoni, non è da scartare l'ipotesi che sia un caso piuttosto che un'ottima precisione, in quanto l'intervallo di confidenza risulta molto ampio.

```
cvf = forecast(VarMdl, 30, 'Y0', res(1:1227)); % varianza  
condizionata  
figure  
h1 = gca;  
h1.XLim = [0, T];  
h1.XTick = [1 263 521 773 1024];  
h1.XTickLabel = {'Giu 2011', 'Giu 2012', 'Giu 2013', 'Giu 2014', 'Giu  
2015'};  
hold on  
plot(1:1257, cv(1:1257), 'b', 'Color', [.7, .7, .7]);
```

```
plot(1228:1257,cv2,'r');  
legend('Varianza condizionata','Previsione');  
title('Simulazione previsione varianza condizionata GARCH(1,1) per 30  
periodi');
```



Il grafico della previsione della varianza condizionata mostra la relativa utilità del modello per descriverne gli effetti: i coefficienti contenuti ne causano una veloce convergenza della previsione futura verso il valore atteso, convergenza ingiustificata a confronto con le ampie oscillazioni della varianza mostrate dal grafico.

```
Mdl = arima('ArLags',[],'D',1,'MaLags',[3 4]); % simulazione  
rendimenti con varianza EGARCH  
EstMdl2 = estimate(Mdl,dy);  
yf2 = forecast(EstMdl2,100,'Y0',dy(1:1157));  
cvf2 = forecast(VarMdl,100,'Y0',res(1:1157));  
figure  
h1 = gca;  
h1.XLim = [0,T];  
h1.XTick = [1 251 501 753 1004];  
h1.XTickLabel = {'Giu 2011','Giu 2012','Giu 2013','Giu 2014','Giu  
2015'};  
hold on  
h2 = plot(dy,'Color',[.7,.7,.7]);  
h3 = plot(1158:1257,yf2,'b','LineWidth',2);  
h4 = plot(1158:1257,yf2 + 1.96*sqrt(cvf2),'r:','LineWidth',2);  
plot(1158:1257,yf2 - 1.96*sqrt(cvf2),'r:','LineWidth',2);  
legend([h2 h3 h4],'Valori osservati','Previsione','Intervalli di  
confidenza al 95%');
```

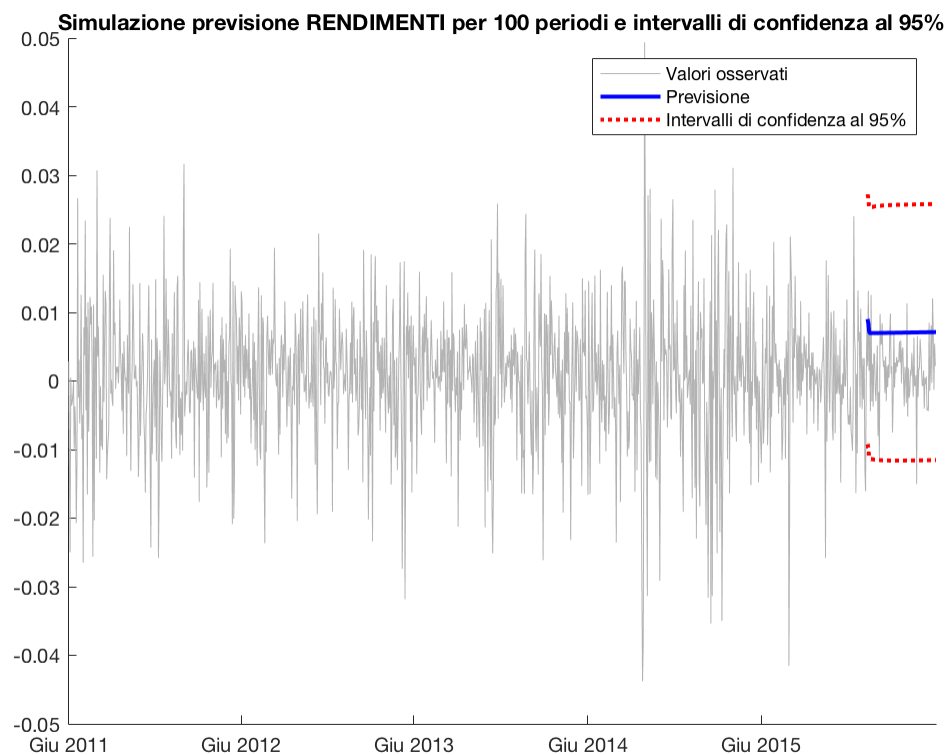


```
title('Simulazione previsione RENDIMENTI per 100 periodi e intervalli  
di confidenza al 95%')  
hold off
```

ARIMA(0,1,4) Model:

Conditional Probability Distribution: Gaussian

<i>Parameter</i>	<i>Value</i>	<i>Standard Error</i>	<i>t Statistic</i>
<i>Constant</i>	<i>1.56857e-06</i>	<i>0.000339706</i>	<i>0.00461743</i>
<i>MA{3}</i>	<i>-0.0475005</i>	<i>0.0267602</i>	<i>-1.77504</i>
<i>MA{4}</i>	<i>-0.0454313</i>	<i>0.0266154</i>	<i>-1.70696</i>
<i>Variance</i>	<i>0.000174718</i>	<i>5.23886e-06</i>	<i>33.3504</i>



Nel terzo grafico struttiamo le informazioni di entrambi i modelli per stimare una previsione del valore atteso futuro dei rendimenti dell'indice, con degli intervalli di confidenza corretti per la varianza condizionata prevista. Il risultato è che anche la previsione dei rendimenti converge quasi subito al suo valore atteso. Come si può notare, inoltre, l'intervallo di confidenza risulta comunque relativamente ridotto rispetto le oscillazioni precedenti.

11) Conclusioni

L'analisi effettuata sull'indice NDX100 ci permette di trarre alcune conclusioni:

- Il modello, escludendo i primi due ritardi, può essere descritto da un processo ARIMA(0,1,4), contenente due soli ritardi. Questa ipotesi è confermata empiricamente grazie alla previsione sulla serie ma può essere poco realistica considerata la teoria dei mercati finanziari.
- Questa relativamente semplice modellizzazione della serie storica può essere indotta dal fatto che negli ultimi 5 anni il trend del mercato sia stato crescente senza troppo grandi oscillazioni e bolle speculative.
- L'effetto dell'eteroschedasticità condizionale autoregressiva non può essere considerata la causa principale dell'eteroschedasticità presente nella serie storica, i modelli ARCH/GARCH/EGARCH risultano quindi relativamente inefficaci nella previsione e poco esplicativi riguardo il fenomeno descritto.
- Entrambe le considerazioni finali possono essere indotte dal fatto che le principali variazioni della serie siano causate da fattori esterni al mercato analizzato, quindi esogene alla nostra analisi.

close all

Published with MATLAB® R2016b