# Analisi dei residui di SE, SET, HW

Fase 1: caricamento e assegnazione frequency

X = scan("clipboard")

11849 1316 4712 800 5097 3270 5390 2135 5962 5795

9271 6864 4247 7961 7191 4970 5012 2929 7363 4907

4700 8219 8674 8263 4294 6097 9115 8924 12561 8626

9559 1706 7405 8057 6463 7595 6702 11052 8422 10019

9594 6443 12052 3535 7962 12876 10614 6469 9396 8421

10895 16583 13724 15362 10740 6999 8168 11241 12206 9600

13865 11190 11306 9760 16678 17245 17059 6331 17635 25429

24124 8025 16168 23902 15618 24579 23023 16311 18848 8986

19745 21513 17378 17839 20271 13917 22520 30279 33373 27728

29400 12759 32293 39832 21975 13304 21511 24382 27790 19053

30140 18024 39378 9351 29202 30944 30984 51181 25809 23035

33939 26385 37051 27823 26570 27587 23434 40944 30355 37954

21883 28773 37127 41490 40786 48097 42233 31002 44818 52006

44235 47597 41311 29891 65657 46764 56831 45358 42840 22724

54484 60902 59253 30692 54913 49463 61164 53140 53932 61780

76395 25181 52878 59362 36356 22600 69768 62622 48934 43957

46525 47652 68435 23165 65990 60518 38106 21206

Y = scan("clipboard")

70.6 121.8 165.1 159.9 210.3 188.8 191.1 41.9 141.7 96.4

67.7 41.1 90.7 113.7 137.1 113.0 117.7 122.6 112.0 19.6

85.7 74.0 71.3 48.3 64.7 101.1 117.9 113.0 139.8 123.8

128.7 27.1 89.9 85.7 72.6 48.0 71.1 94.9 112.1 119.2

123.5 117.0 129.0 23.8 111.9 96.8 82.1 58.2 85.8 110.2

142.2 133.2 142.8 140.9 131.2 28.7 101.5 82.1 69.7 56.1

76.7 86.3 115.8 132.1 147.3 139.9 131.2 41.0 105.5 86.6

86.3 51.3 80.3 120.5 172.4 133.0 169.1 168.3 146.8 38.0

105.1 97.6 88.1 63.7 96.9 120.6 154.7 137.3 166.5 158.9

150.1 35.9 119.0 103.6 89.6 60.6 92.1 118.1 140.5 157.8

157.6 153.3 151.3 17.2 106.7 94.1 70.9 57.9 61.6 83.1

113.0 109.2 121.5 122.4 123.2 18.3 103.1 68.2 57.1 33.1

59.3 100.1 115.2 116.0 126.0 109.2 105.1 10.0 75.0

X.fr = ts(X, frequency=12, start=c(1995,1))

Y.fr = ts(Y, frequency=12, start=c(2000,1))

Fase 2: esecuzione di SE, SET, HW, loro previsioni passate e future:

SE = HoltWinters(X,beta=FALSE ,gamma=FALSE)

plot(SE)

SET = HoltWinters(X, gamma=FALSE)

plot(SET)

HW = HoltWinters(X.fr)

plot(HW)

plot(SE, predict(SE,12))

plot(SET, predict(SET,12))

plot(HW, predict(HW,12))

Come estrarre i valori numerici della previsione sul passato?

?HoltWinters

prev.SE = SE$fitted[,1]

par(mfrow=c(1,2)); plot(SE); plot(prev.SE)

prev.SET = SET$fitted[,1]

par(mfrow=c(1,2)); plot(SET); plot(prev.SET)

prev.HW = HW$fitted[,1]

par(mfrow=c(1,2)); plot(HW); plot(prev.HW)

?HoltWinters offre anche SSE. Proviamo:

c(SE$SSE, SET$SSE, HW$SSE)

> c(SE$SSE, SET$SSE, HW$SSE)/10000000000

[1] 1.4278340 1.9086986 0.9980452

Però non sappiamo se sono affidabili, cioè se sono calcolati a partire dalle stesse finestre. Inoltre il secondo può essere inutilmente affetto dagli errori in fase di inizializzazione. Va quindi fatto a mano.

Fase 3: calcolo dei residui.

e = p – x.

Il problema sta nella lunghezza delle serie prev.xxx

length(X)

length(prev.SE)

length(prev.SET)

length(prev.HW)

E’ naturale, ricordando che più il metodo è elaborato più la fase di inizializzazione è elaborata.

Quindi, se calcoliamo prev.xxx[n]-X[n], sbagliamo.

residui.SE = prev.SE[(167-60):167]-X[(168-60):168]

residui.SET = prev.SET[(166-60):166]-X[(168-60):168]

residui.HW = prev.HW[(156-60):156]-X[(168-60):168]

par(mfrow=c(1,2)); ts.plot(residui.SE); ts.plot(residui.SET)

par(mfrow=c(1,2)); ts.plot(residui.SET); ts.plot(residui.HW)

c(sd(residui.SE), sd(residui.SET), sd(residui.HW))

Apparentemente paradossale, ma conferma che SE non è un sottocaso di SET e che la strategia di catturare il trend può venir ingannata da fluttuazioni che appaiono come trend.

Fase 4: analisi dei residui e intervallo di confidenza sulla previsione.

par(mfrow=c(1,1))

hist(residui.SE,20)

quantile(residui.SE,0.05)

quantile(residui.SE,0.95)

> quantile(residui.SE,0.05)

 5%

-23715.17

> quantile(residui.SE,0.95)

 95%

27652.67

>

Li si confronti con l’istogramma.

L’intervallo (asimmetrico, empirico) di confidenza al 90% per i residui è [-23715.17 , 27652.67 ]

L’intervallo (asimmetrico, empirico) di confidenza al 90% per la previsione del primo mese successivo è:

[45292.56-23715.17 , 45292.56+27652.67 ]

perché, con predict(SE,1),

> predict(SE,1)

Time Series:

Start = 169

End = 169

Frequency = 1

 fit

[1,] 45292.56

>

Vediamo con HW.

hist(residui.HW,20)

Fingiamo che sia abbastanza gaussiano, per esercitarci.

m= mean(residui.HW)

sd= sd(residui.HW)

qnorm(0.05,m,sd)

qnorm(0.95,m,sd)

> qnorm(0.05,m,sd)

[1] -19511.47

> qnorm(0.95,m,sd)

[1] 18352.62

>

predict(HW,1)

> predict(HW,1)

 Jan

2009 59226.11

>

L’intervallo (simmetrico, parametrico) di confidenza al 90% per la previsione del primo mese successivo è:

[59226.11-19511.47, 59226.11+18352.62]

Nota:

quantile(residui.HW,0.05)

quantile(residui.HW,0.95)

qnorm(0.05,m,sd)

qnorm(0.95,m,sd)

Se non è richiesta una grande precisione, sono paragonabili.