1. **RICHIAMI LEZIONE PRECEDENTE**

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

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

ts.plot(accessori.auto)

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

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

ts.plot(vendite.moto)

1. **METODO DI HOLT-WINTERS**

par(mfrow=c(1,1))

HW.auto = HoltWinters(accessori.auto)

plot(HW.auto, predict(HW.auto,12))

HW.moto = HoltWinters(vendite.moto)

plot(HW.moto, predict(HW.moto,12))

1. **METODI REGRESSIVI PER LE SERIE STORICHE**

Nota. Il software R offre tantissime possibilità automatizzate per costruire modelli di natura regressiva, però ispirati alle categorie AR, ARMA, ecc. Un comando molto semplice ed efficace, per chi volesse provare, è ar.ols

ar.ols(accessori.auto)

predict(ar.ols(accessori.auto))

ar.ols(vendite.moto)

predict(ar.ols(vendite.moto))

Il difetto sono i troppi coefficienti, col pericolo di overfitting. Vediamo a mano.

> length(Y)

[1] 129

y=Y[13:129]; x1=Y[12:128]; x12=Y[1:117]

Reg = lm(y~x1+x12)

summary(Reg)

a1 = Reg$coefficient[2]; a12 = Reg$coefficient[3]; b = Reg$coefficient[1]

(controllare con Reg)

P=1:141; P[1:129]=Y

for (i in 1:12) {

P[129+i]=a1\*P[129+i-1]+a12\*P[129+i-12]+b

}

Pplus = ts(P, frequency =12, start=c(2000,1))

plot(Pplus)

(si ricordi plot(vendite.moto)

par(mfrow=c(1,2))

plot(HW.moto,predict(HW.moto,12))

plot(Pplus)

Esercizio: ripetere l’analisi con la serie X.

1. **PRIME ANALISI DEI RESIDUI.**

residui.auto = residuals(HW.auto)

plot(residui.auto)

acf(residui.auto)

white.noise=rnorm(length(residui.auto))

acf(white.noise)

(a titolo di confronto)

par(mfrow=c(1,2))

acf(residui.auto); acf(white.noise)

hist (residui.auto,20); hist (X,20)

così abbiamo una percezione grafica della variabilità originaria e di quella dei residui

1-var(residui.auto)/var(X)

Quanto sono gaussiani i residui? (vedi anche esercizio in fondo)

qqnorm(residui.auto); qqnorm(white.noise)

1. **ALTRO ESEMPIO**

residui.moto = residuals(HW.moto)

plot(residui.moto)

white.noise=rnorm(length(residui.moto))

par(mfrow=c(1,2))

acf(residui.moto); acf(white.noise)

hist (residui.moto,20); hist (Y,20)

1-var(residui.moto)/var(Y)

qqnorm(residui.moto); qqnorm(white.noise)

1. **INCERTEZZA SULLA PREVISIONE**

…………….. Premessa: intervallo di valori di una gaussiana, al 90% ……………….

X= rnorm(10000, 10, 1)

hist(X,50)

m= mean(X); s=sd(X)

qnorm(0.05,m,s); qnorm(0.95,m,s)

m+s\*qnorm(0.05); m+s\*qnorm(0.95)

quantile(X,0.05); quantile(X,0.95)

…………….. Fine premessa ……………….

**Esercizio**. Relativamente ai dati vendite.moto:

1. Trovare un intervallo parametrico in cui, al 90%, prevediamo cada il primo valore futuro.
2. Trovare l’analogo non parametrico,
3. Tracciare graficamente due bande, entro le quali prevediamo stiano i valori al 90%, relativi ai 12 mesi successivi.

Soluzione.

HW.moto = HoltWinters(vendite.moto)

residui.moto = residuals(HW.moto)

m= mean(residui.moto); s=sd(residui.moto)

predict(HW.moto,1)+qnorm(0.05,m,s); predict(HW.moto,1)+qnorm(0.95,m,s)



predict(HW.moto,1)+ quantile(residui.moto,0.05)

predict(HW.moto,1)+ quantile(residui.moto,0.95)

plot(HW.moto,predict(HW.moto ,12))

lines(predict(HW.moto,12)+ quantile(residui.moto,0.05),col= " green ")

lines(predict(HW.moto,12)+ quantile(residui.moto,0.95), col= " green ")

plot(predict(HW.moto ,12))

lines(predict(HW.moto,12)+ quantile(residui.moto,0.05),col= " green ")

lines(predict(HW.moto,12)+ quantile(residui.moto,0.95), col= " green ")

**Nota**: sembrano valori alti, bande larghe. Si confronti però con il grafico dei residui

plot(residui.moto)

plot(vendite.moto)

**Nota**: realisticamente, andrebbero considerati solo i residui più recenti (che sono più piccoli), che darebbero bande un po’ più strette.

1. **APPROFONDIMENTI SULL’INCERTEZZA SULLA PREVISIONE: DEVIAZIONI STAGIONALI**

Osservando

residui.moto

vediamo con chiarezza gli anni ed i mesi. Possiamo allora creare una matrice avente per colonne i mesi ed per righe gli anni:

A=matrix(nrow=9, ncol=12)

for (i in 1:9) {

A[i,]= residui.moto[((i-1)\*12+1):(i\*12)]

}

Ora calcoliamo le deviazioni standard mese per mese. Utilizziamo solo gli ultimi mesi:

Sig = 1:12

for (j in 1:12) {

Sig[j]= sd(A[3:9,j])

}

ts.plot(Sig)

Sono moto diversificati. Quindi probabilmente ne vale la pena. Vediamo le bande:

par(mfrow=c(1,2))

ts.plot(Sig)

ts.plot(predict(HW.moto,15)[4:15])

lines(predict(HW.moto,15)[4:15]+ qnorm(0.05,0,Sig),col= " red ")

lines(predict(HW.moto,15)[4:15]+ qnorm(0.95,0,Sig),col= " red ")

1. **DEVIAZIONI TIPICHE CON fPCA (ANCORA DA SVOLGERE)**

Vogliamo eseguire una PCA della tabella A, ovvero dei residui annuali. Purtroppo

princomp(A)

non funziona perché ci sono meno anni che mesi. Accontentiamoci allora del profilo del primo semestre:

Arid =matrix(nrow=9, ncol=6)

Arid = A[,1:6]

PCA = princomp(Arid)

par(mfrow=c(1,1))

biplot(PCA)

e1=PCA$loadings[,1]

e2=PCA$loadings[,2]

par(mfrow=c(1,2))

ts.plot(e1); ts.plot(e2)

ts.plot(e1)

ts.plot(predict(HW.moto,15)[4:9])

lines(predict(HW.moto,15)[4:9]+e1\*PCA$sdev[1]\*qnorm(0.95), col= " red ")

lines(predict(HW.moto,15)[4:9]-e1\*PCA$sdev[1]\*qnorm(0.95),col= " red ")

ts.plot(e2)

ts.plot(predict(HW.moto,15)[4:9])

lines(predict(HW.moto,15)[4:9]+e2\*PCA$sdev[2]\*qnorm(0.95), col= " red ")

lines(predict(HW.moto,15)[4:9]-e2\*PCA$sdev[2]\*qnorm(0.95),col= " red ")

summary(PCA)