

A

Splines cúbicos em séries temporais

A aproximação de uma função não-linear pode ser realizada por meio de uma função *spline*, na qual pode-se escolher o grau de suavização dependendo do número de derivadas existentes. Por exemplo, se existir as p primeiras derivadas da função *spline*, dizemos que ela é de ordem p . Sendo p um valor alto, isto implica que a função é muito suave, portanto, valores relativamente baixos de p produzem uma aproximação suavizada. Para p igual a 1 temos uma função *spline* linear, quadrática para p igual a 2 e cúbica para p igual a 3.

Movimentos periódicos são uma importante característica na análise de séries temporais. Efeitos sazonais são comuns em dados semestrais ou mensais. Para intervalos menores de tempo, como por exemplo dados semanais ou diários, os movimentos periódicos podem existir. Comumente, um comportamento periódico ao longo de s intervalos de tempo é modelado com $s - 1$ variáveis dicotômicas ou, alternativamente, com $s - 1$ termos trigonométricos. Duas importantes conseqüências decorrem destas aproximações. Primeiro, movimentos periódicos fixos são raros em séries temporais reais, desenvolvem-se ao longo do tempo. Segundo, se s é grande, o modelo com variáveis dicotômicas ou termos trigonométricos não são muito parcimoniosos. Sendo assim, o uso de *splines* em muito desses casos é uma solução mais apropriada para modelar séries temporais periódicas. Para mais detalhes sobre a metodologia vide Koopman [4].

B**Ações negociáveis simultaneamente no sistema eletrônico e no viva voz**

Código	Empresa	Especificação
BBDC3	BRADESCO	ON
BBDC4	BRADESCO	PN
CMIG3	CEMIG	ON
CMIG4	CEMIG	PN
ELET3	ELETROBRAS	ON
ELET5	ELETROBRAS	PNA
ELET6	ELETROBRAS	PNB
ITAU3	ITAUBANCO	ON
ITAU4	ITAUBANCO	PN
PETR3	PETROBRAS	ON
PETR4	PETROBRAS	PN
TMAR3	TELEMAR N L	ON
TMAR5	TELEMAR N L	PNA
TMAR6	TELEMAR N L	PNB
TNLP3	TELEMAR	ON
TNLP4	TELEMAR	PN
VALE3	VALE R DOCE	ON
VALE5	VALE R DOCE	PNA

Além das ações citadas acima, as opções dessas ações, do Índice Bovespa e do índice IBrX 50, também têm negociação simultânea nos Pregões Eletrônico e Viva Voz.

C Demonstração do teste Engle

Dada a sequência de variáveis aleatórias X_1, \dots, X_n i.i.d. com distribuição exponencial de parâmetro λ . Testaremos as hipóteses:

$$\begin{cases} H_0 : \sigma^2 = \frac{1}{\lambda} = 1 \\ H_A : \sigma^2 \neq 1 \end{cases}$$

A estatística baseada em $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (X_i - 1)^2$.

Fazendo-se $Y_i = (X_i - 1)^2$. Logo, $\hat{\sigma}^2 = \bar{Y}$. Assim temos:

$$\frac{\bar{Y} - E(\bar{Y})}{\sqrt{V(\bar{Y})}} \approx N(0, 1).$$

Sendo, $E(\bar{Y}) = 1$ temos $\frac{\bar{Y} - 1}{\sqrt{V(\bar{Y})}} \approx N(0, 1)$.

Tomando-se o termo do denominador observamos que $\hat{V}(\bar{Y}) = \frac{\sigma_y^2}{N} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N} = \frac{\sum_{i=1}^N ((X_i - 1)^2 - \overline{(X_i - 1)^2})^2}{N} = \frac{DP^2[(X_i - 1)^2]}{N}$.

Resultando em $\sqrt{N} \left(\frac{\hat{\sigma} - 1}{\sigma_v} \right)$, onde $\sigma_v = DP[(X_i - 1)^2]$

D

Programa

D.1

Análise da dados de alta frequência

```
rm(list=ls(all=TRUE));
arq1<-read.table("C:/Documents and
Settings/Savano/Meus
documentos/savano/PUC/dissertacao/dados/Chinese_ju_oc.csv",header=T,sep="");
arq2<-read.table("C:/Documents and Settings/Savano/Meus
documentos/savano/PUC/dissertacao/dados/Chinese_nov_jan.csv",header=T,sep="");
dado<-rbind(arq1,arq2);
remove(arq1);
remove(arq2);
attach(dado);

## mostrando que a s{'e}rie e irregularmente espa{c}ada no tempo ##
amostra<-dado[month==7 & day==18 & hour==11 & minute<=30,];
second.total<-amostra$hour*60*60+amostra$minute*60+amostra$second;
amostra2<-cbind(amostra,second.total);
rm(amostra,second.total);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
/figcap1_irregularidade.eps");
plot(amostra2$second.total,amostra2$transaction.price,xlim=c(39590,41450)
,xlab="Horas",ylab="Pre{c}o negociado",xaxt="n");
axis(1,c(39603,40516,41424),c("11:00:03","11:15:16","11:30:24"));
dev.off();

## discretiza{c}{\~a}o ##
price.change<-numeric(length(transaction.price)-1);
for(i in 1:length(price.change)){
price.change[i]<-transaction.price[i+1]-transaction.price[i]
}
kurtosis<- (1/length(price.change))*sum(((price.change-mean(price.change))/sd(price.change))^4);
skewness<- (1/length(price.change))*sum(((price.change-mean(price.change))/sd(price.change))^3);
variance<-var(price.change);
options(digits=5);
estat.preco<-as.matrix(summary(price.change));
estat.tabela<-rbind(estat.preco,kurtosis,skewness,variance);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
/savano/PUC/dissertacao/programas_R/resultados/figcap1_discretizacao.eps");
hist(price.change,xlab="Centavos",ylab="Freq{\u}{\~e}ncia",br=c(-0.015*0:100,0.015*1:100)
,xlim=c(-0.15,0.15));
dev.off();
```

```

remove(estat.preco,kurtosis,skewness,variance,price.change);

## comportamento diurno ##
amostra.dia<-dado[month==7 & day==18,];
min.inicial<-amostra.dia[amostra.dia$minute<=30,];
min.final<-amostra.dia[amostra.dia$minute>=31,];
aux.inicial<-as.matrix(by(min.inicial$duration,min.inicial$hour,mean));
aux.final<-as.matrix(by(min.final$duration,min.final$hour,mean));
aux.media<-matrix(nrow=length(aux.inicial)+length(aux.final),ncol=1);
for(i in 1:length(aux.inicial)){
aux.media[2*i-1,1]<-aux.inicial[i,1]
}
for(i in 1:length(aux.final)){
aux.media[2*i,1]<-aux.final[i,1]
}
aux.media[1,1]<-NaN
aux.media<-na.exclude(aux.media)
postscript(file="C:/Documents and Settings/Savano/Meus documentos
/savano/PUC/dissertacao/programas_R/resultados/figcap1_diurno.eps");
plot(aux.media,xlab="Hora do dia",ylab="Segundos",xaxt="n");
axis(1,c(2,4,6,8,10),c("10:00","11:00","13:00","14:00","15:00"));
lines(smooth.spline(aux.media));
dev.off();
remove(amostra.dia, min.inicial, min.final,aux.inicial,aux.final,aux.media);

## agregando e dessazonando as dura\c{c}{\~o}es ##
media.semana<- aggregate(duration,list(dummy=minute>30,hour=hour,day=day,month=month,year=year)
,mean);

detach(dado);
attach(media.semana);
aux.dia<-paste(day,month,year,sep="/");
data<-strptime(aux.dia, "%d/%m/%Y");
remove(aux.dia);
dia.semana<-weekdays(data);
detach(media.semana);
dado.agregado<-cbind(dia.semana,media.semana);
remove(dia.semana,media.semana);

## calculando a componente sazonal determin\{'i}stica das dura\c{c}{\~o}es ##
aux.spline <- smooth.spline(dado.agregado$x);
phi <- aux.spline$y;
remove(aux.spline);
agregado.phi <- cbind(dado.agregado,phi);
#write.table(agregado.phi, file = "C:/Documents and Settings/Savano/Meus documentos
/savano/PUC/dissertacao/programas_R/resultados/agregado_phi.txt", sep = ";");
postscript(file="C:/Documents and Settings/Savano/Meus documentos
/savano/PUC/dissertacao/programas_R/resultados/sazao.eps");
plot(phi,type="l",main="");
dev.off();
remove(phi);

## dessazonando as dura\c{c}{\~o}es ##
dummy <- numeric(length(dado$minute));
for(i in 1:length(dummy)) {
if(dado$minute[i]>30) {
dummy[i] <- "TRUE";
}
else {
dummy[i] <- "FALSE";
}
}

```

```

}
}
dado.dummy <- cbind(dado,dummy);
remove(dummy);
dado.sazao<- merge(dado.dummy,agregado.phi);
ind1 <- order(dado.sazao$year,dado.sazao$month,dado.sazao$day,dado.sazao$hour,
              dado.sazao$minute,dado.sazao$second);
dado.aux1 <- dado.sazao[ind1,];
dessa <- dado.aux1$duration/dado.aux1$phi;
remove(dado.sazao);
remove(ind1);
dado.dessa <- cbind(dado.aux1,dessa);
ind1<-which(dado.dessa$hour==9 & dado.dessa$minute==30);
ind2<-which(dado.dessa$hour==15 & dado.dessa$minute>0);
dado.dessa[ind1,] <- NA;
dado.dessa[ind2,] <- NA;
dado.dessa <- na.excluir(dado.dessa);
remove(ind1,ind2);
remove(dado.aux1);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/acf_duracao.eps");
acf(dado.dessa$dessa,lag.max=200,main="");
dev.off();
#write.table(dado.dessa, file = "C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/duracao_dessa.txt",sep="");

## dessasazonando a varia\c{c}{\~a}o de pre\c{c}os ##
price.change<-numeric(length(dado$transaction.price));
price.change[i]<-NA
for(i in 2:length(price.change)){
price.change[i]<-dado$transaction.price[i]-dado$transaction.price[i-1]
}
dado.aux2 <- cbind(dado.dummy,price.change);
remove(price.change);
agregado.dv <- aggregate(dado.aux2$price.change,list(dummy=dado.aux2$minute>30,
            hour=dado.aux2$hour,day=dado.aux2$day,month=dado.aux2$month,year=dado.aux2$year),sd);
agregado.dv <- na.excluir(agregado.dv);
aux.spline <- smooth.spline(agregado.dv$x);
dv <- aux.spline$y;
remove(aux.spline);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/change_comp.eps");
plot(dv,type="l",main="");
dev.off();
agregado.dv <- cbind(agregado.dv,dv);
remove(dv);
dado.aux3 <- merge(dado.aux2,agregado.dv);
remove(dado.aux2);
ind3 <- order(dado.aux3$year,dado.aux3$month,dado.aux3$day,dado.aux3$hour,
              dado.aux3$minute,dado.aux3$second);
dado.aux3 <- dado.aux3[ind3,];
remove(ind3);
price.change.dessa <- dado.aux3$price.change/dado.aux3$dv;
dado.dv <- cbind(dado.aux3,price.change.dessa);
remove(price.change.dessa,dado.aux3);
ind1<-which(dado.dv$hour==9 & dado.dv$minute==30);
ind2<-which(dado.dv$hour==15 & dado.dv$minute>0);
dado.dv[ind1,] <- NA;
dado.dv[ind2,] <- NA;

```

```

dado.dv <- na.exclude(dado.dv);
remove(ind1,ind2);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/acf_change_dessa.eps");
acf(dado.dv$price.change.dessa,main="");
dev.off();
#write.table(dado.dv, file = "C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/price_change_dessa.txt",sep=";");

## dessasazonando a varia\c{c}{\~a}o de pre\c{c}os - valor absoluto ##
price.change<-numeric(length(dado$transaction.price));
price.change[i]<-NA
for(i in 2:length(price.change)){
price.change[i]<-abs(dado$transaction.price[i]-dado$transaction.price[i-1])
}
dado.aux2 <- cbind(dado.dummy,price.change);
remove(price.change);
agregado.dv <- aggregate(dado.aux2$price.change,list(dummy=dado.aux2$minute>30,
           hour=dado.aux2$hour,day=dado.aux2$day,month=dado.aux2$month,year=dado.aux2$year),sd);
agregado.dv <- na.exclude(agregado.dv);
aux.spline <- smooth.spline(agregado.dv$x);
dv <- aux.spline$y;
remove(aux.spline);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/changeabs_comp.eps");
plot(dv,type="l",main="");
dev.off();
agregado.dv <- cbind(agregado.dv,dv);
remove(dv);
dado.aux3 <- merge(dado.aux2,agregado.dv);
remove(dado.aux2);
ind3 <- order(dado.aux3$year,dado.aux3$month,dado.aux3$day,dado.aux3$hour,
           dado.aux3$minute,dado.aux3$second);
dado.aux3 <- dado.aux3[ind3,];
remove(ind3);
price.change.dessa <- abs(dado.aux3$price.change)/dado.aux3$dv;
dado.dv <- cbind(dado.aux3,price.change.dessa);
remove(price.change.dessa,dado.aux3);
ind1<-which(dado.dv$hour==9 & dado.dv$minute==30);
ind2<-which(dado.dv$hour==15 & dado.dv$minute>0);
dado.dv[ind1,] <- NA;
dado.dv[ind2,] <- NA;
dado.dv <- na.exclude(dado.dv);
remove(ind1,ind2);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/acf_changeabs_dessa.eps");
acf(dado.dv$price.change.dessa,lag.max=200,main="");
dev.off();
postscript(file="C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/plot_changeabs_dessa.eps");
plot(dado.dv$price.change.dessa,type="l",ylab="Varia\c{c}{\~a}o do pre\c{c}o de negocia\c{c}{\~a}o",main="");
dev.off();
#write.table(dado.dv, file = "C:/Documents and Settings/Savano/Meus documentos
           /savano/PUC/dissertacao/programas_R/resultados/price_changeabs_dessa.txt",sep=";");

## dessasazonando o pre\c{c}o m{\`e}dio ##
midprice <- (dado$bid+dado$ask)/2;
dado.midprice <- cbind(dado.dummy,midprice);
remove(midprice);

```

```

agregado.midprice <- aggregate(dado.midprice$midprice,
                             list(dummy=dado.midprice$minute>30,hour=dado.midprice$hour,
                                   day=dado.midprice$day,month=dado.midprice$month,year=dado.midprice$year),mean);
aux.spline <- smooth.spline(agregado.midprice$x);
midprice.comp <- aux.spline$y;
remove(aux.spline);
agregado.midprice <- cbind(agregado.midprice, midprice.comp);
remove(midprice.comp);
dados.midprice<- merge(dado.midprice,agregado.midprice);
ind1 <- order(dado.midprice$year,dado.midprice$month,dado.midprice$day,
             dado.midprice$hour,dado.midprice$minute,dado.midprice$second);
dados.aux1 <- dados.midprice[ind1,];
remove(ind1);
midprice.dessa <- dados.aux1$midprice/dados.aux1$midprice.comp;
dados.midprice.dessa <- cbind(dados.aux1, midprice.dessa);
remove(midprice.dessa);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/acf_midprice_dessa.eps");
acf(dados.midprice.dessa$midprice.dessa,lag.max=200,main="");
dev.off();
#write.table(dados.midprice.dessa, file = "C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/midprice_dessa.txt",sep="");

## varia\c{\~a}o do midprice ##
change.midprice<-numeric(length(dados.midprice.dessa$midprice.dessa));
change.midprice[i]<-NA
for(i in 2:length(change.midprice)){
change.midprice[i]<-dados.midprice.dessa$midprice.dessa[i]-dados.midprice.dessa$midprice.dessa[i-1]
}
postscript(file="C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/acf_changemidprice_dessa.eps");
acf(change.midprice,na.action=na.omit,main="");
dev.off();

## dessasazonando o volume ##
agregado.vol <- aggregate(dados.dummy$volume,
                        list(dummy=dados.dummy$minute>30,hour=dados.dummy$hour,
                              day=dados.dummy$day,month=dados.dummy$month,year=dados.dummy$year),mean);
aux.spline <- smooth.spline(log(agregado.vol$x));
comp.vol <- aux.spline$y;
remove(aux.spline);
agregado.vol <- cbind(agregado.vol,comp.vol);
remove(comp.vol);
dados.aux4 <- merge(dados.dummy,agregado.vol);
ind2 <- order(dados.aux4$year,dados.aux4$month,dados.aux4$day,
             dados.aux4$hour,dados.aux4$minute,dados.aux4$second);
dados.aux4 <- dados.aux4[ind2,];
remove(ind2);
vol.dessa <- log(dados.aux4$volume)/dados.aux4$comp.vol;
dados.vol <- cbind(dados.aux4,vol.dessa);
ind1<-which(dados.vol$hour==9 & dados.vol$minute==30);
ind2<-which(dados.vol$hour==15 & dados.vol$minute>0);
dados.vol[ind1,] <- NA;
dados.vol[ind2,] <- NA;
dados.vol <- na.exclude(dados.vol);
remove(ind1,ind2);
postscript(file="C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/acf_log_vol.eps");
acf(dados.vol$vol.dessa,lag.max=200,main="");

```



```
dev.off();
#write.table(dado.vol, file = "C:/Documents and Settings/Savano/Meus documentos
            /savano/PUC/dissertacao/programas_R/resultados/vol_dessa.txt", sep="");
remove(dado.aux4, vol.dessa);
```

D.2

Modelo GACD(1,1)

```
rm(list=ls(all=TRUE));
#library(stepfun);
#library(tseries);
arq1<-read.table("C:/Documents and Settings/Savano/Meus
documentos/savano/PUC/dissertacao/programas_R/resultados/duracao_dessa.txt", sep="");
x <- arq1$dessa; rm(arq1);
# Definindo a f\c{c} de verossimilhan\c{c}a - restri\c{c}\{o}es dos par\{a}metros ref [1]
log.gama <- function(teta=c(w,a,b)){
psi <- numeric(length(x)); # inicializa\c{c}\{o} de psi
lambda <- numeric(length(x));
gama <- 1;
kapa <- 1;
psi[1] <- exp(teta[1])/(1-(exp(teta[3])/(1+exp(teta[2])+exp(teta[3])))); #valor inicial de psi[1]
for(i in 2:length(x)) {
psi[i] <- exp(teta[1])+(exp(teta[2])/(1+exp(teta[2])+exp(teta[3])))*x[i-1]+
            (exp(teta[3])/(1+exp(teta[2])+exp(teta[3])))*psi[i-1] #calculo de psi
}
for(i in 1:length(x)){
lambda[i] <- psi[i]*gamma(kapa)/gamma(kapa+1/gama)
}
res <- sum(log(gama)+(kapa*gama-1)*log(x/lambda)-log(lambda*gamma(kapa))-(x/lambda)^gama) #verossimilhan\c{c}a
return(res)
}
# maximizando a verossimilhan\c{c}a
valor.optim <- optim(c(1,1,1),log.gama,method ="BFGS",control=list(fnscale=-1),hessian=T);
#valores dos parametros
options(digits=5);
parametro<- numeric(3);
parametro[1] <- exp(valor.optim$par[1]);
parametro[2] <- exp(valor.optim$par[2])/(1+exp(valor.optim$par[2])+exp(valor.optim$par[3]));
parametro[3] <-exp(valor.optim$par[3])/(1+exp(valor.optim$par[2])+exp(valor.optim$par[3]));
print(parametro);
# Matriz de covariancia dos par\{a}metros - ref [1]
matrix.fisher<-(-valor.optim$hessian)^-1; # ref[2]
psi1<-valor.optim$par[1];
psi2<-valor.optim$par[2];
psi3<-valor.optim$par[3];
J<-matrix(data=NA,nrow=3,ncol=3);
J[1,1]<-exp(psi1);
J[1,2]<-0;
J[1,3]<-0;
J[2,1]<-0;
J[2,2]<- (exp(psi2)/(1+exp(psi2)+exp(psi3)))-(exp(psi2)/(1+exp(psi2)+exp(psi3)))^2;
J[2,3]<-((-1)*exp(psi2)*exp(psi3))/(1+exp(psi2)+exp(psi3))^2;
J[3,1]<-0;
J[3,2]<-((-1)*exp(psi2)*exp(psi3))/(1+exp(psi2)+exp(psi3))^2;
```

```
J[3,3]<-(exp(psi3)/(1+exp(psi2)+exp(psi3)))-(exp(psi3)/(1+exp(psi2)+exp(psi3)))^2;
cov.par<-J %*% matrix.fisher %*% t(J);
# Calculo dos erros
psi <- numeric(length(x));
erro <- numeric(length(x));
psi[1] <- valor.optim$par[1]/(1-valor.optim$par[3]);
erro[1] <- x[1]/psi[1];
for(i in 2:length(x)) {
psi[i] <- (parametro[1]+parametro[2]*x[i-1]+parametro[3]*psi[i-1]);
erro[i] <-x[i]/psi[i];
}
```

Bibliografia

- [1] DUFOUR, A.; ENGLE, R. F. **The ACD Model: Predictability of the Time Between Consecutive Trades.**
- [2] E., BERNDT; HALL, B.; HALL, R.; HAUSMAN, J. **Estimation and inference in nonlinear structural models.** Annals of Economic and Social Measurement, 1974.
- [3] ENGLE, R. F.; RUSSEL, J. R. **Autoregressive Conditional Duration: A new model for irregularly spaced transaction data.** Econometrica, 66(5):1127–1162, setembro 1988.
- [4] KOOPMAN, S.J. **Diagnostic checking and intra-daily effects in Time series models.** PhD thesis, London School of Economics and Political Science, 1992.
- [5] CROMWELL, J.B.; LABYS, W.C.; TERRAZA, M. **Univariate Tests for Time Series Models.** Número 99 em Quantitative Applications in the Social Sciences. SAGE, 1994.
- [6] VENABLES, V.N.; RIPLEY, B.D. **Modern applied statistics with S-plus.** Quantitative Applications in the Social Sciences. Springer-Verlag, 3 edition, 1996.
- [7] NOCEDAL, J.; WRIGHT, S. J. **Numerical Optimization.** Springer-Verlag, 1999.
- [8] ENGLE, R. F. **The Econometrics of High Frequency Data.** Econometrica, (68):1–22, 2000.
- [9] ZHANG, M. Y.; RUSSEL, R. J.; TSAY, R. S. **A nonlinear autoregressive conditional duration model with applications to financial transaction data.** Journal of Econometrics, (104):179–207, 2001.
- [10] ENGLE, R. F.; RUSSEL, J. R. **Analysis of High Frequency Data.** 2002.

- [11] BOWMAN, K. O.; SHENTON, L. R. **Problems with Maximum Likelihood Estimation and the 3 Parameter Gamma Distribution.** J. Statist. Comput. Simul., 2002.