

Diogo Destro Barcellos

✖ Sou aluno de doutorado pelo Instituto Oceanográfico da Universidade de São Paulo. Durante o mestrado desenvolvi estudo de crescimento da lula *Doryteuthis plei*, da região do litoral do Estado de São Paulo. Atualmente desenvolvo estudo relacionado com o Monitoramento Acústico Passivo de Cetáceos na costa do Estado de São Paulo, Brasil. <http://lattes.cnpq.br/3720211733498851>

Meus Exercícios

[Exercícios Diogo D. Barcellos](#)

Trabalho Final

Proposta A

Proponho o desenvolvimento de uma função que fará a leitura de um dataframe com apenas duas variáveis quantitativas, sendo elas x (dados de idade) e y (comprimento do animal). Será feito um teste lógico para testar o mesmo comprimento das variáveis (x, y), as quais devem ter mesma extensão. Se forem de comprimentos diferentes, uma mensagem será enviada para o usuário. Será feito também um argumento lógico de exclusão de NA. A relação entre as variáveis será testada através dos modelos matemáticos logarítmico, exponencial, potencial e linear. Será obtido o AIC de cada modelo. O menor valor de AIC será usado como indicativo de melhor modelo de ajuste de crescimento (em relação as variáveis testadas). O objeto de saída será um gráfico que demonstre a relação entre as variáveis x e y, com a curva do modelo que melhor se ajustou aos dados testados. Esta função R será muito útil para pessoas que estudam crescimento de diferentes espécies. Será possível testar de uma vez só diferentes modelos matemáticos (os quais serão descritos no help da função) nas variáveis x e y, e obter a melhor função matemática que se ajusta aos dados testados e estimativa da curva de crescimento, segundo o modelo.

Proposta B

Proponho executar a proposta A, adicionando os modelos matemáticos de crescimento: Gompertz, von Bertalanffy, Logístico e Schunute. Mas seria bem avançado!

Oi, Diogo,

Bom, vamos lá, em primeiro lugar queria me desculpar pela confusão, no final depois de conversar com os outros monitores percebi que eu errei e que na verdade a sua proposta dá sim para fazer como uma função. Acho que talvez eu tinha entendido errado o que você queria fazer, mas agora entendi.

Ainda assim, você acabou fazendo apenas uma proposta, mas tudo bem, vamos nela! Só algumas sugestões que eu pensei que podem ser interessantes (mas veja se seria realmente útil mesmo):

- Pode ser interessante colocar argumentos para o usuário da função escolher como quer o gráfico. Talvez uma opção seja um argumento que permite o

usuário escolher se quer plotar todos os modelos ou apenas aquele que melhor se ajusta ou todos que o usuário quiser. Outra opção legal seria incluir nos argumentos das função argumentos de gráfico, como cor, espessura da linha, símbolo de plotagem, local da legenda, etc.

- Haveria algum objeto de saída além do gráfico? Se sim, qual seria? Talvez os próprios coeficientes do modelo?

Bom, acho que é isso. São pequenas coisinhas apenas, mas já dá para ir mandar bala na função! De novo, desculpe pela confusão!

—[Diana Garcia](#)

Olá Diana, Obrigado pelas sugestões, vou tentar atender todas. Penso como objeto de saída: os gráficos e os coeficientes testados. Obrigado pela aprovação, conforme o desenvolvimento da função nos falamos! Diogo

Página de Ajuda

melhorajuste

package:nenhum

R Documentation

Análise exploratória de modelos matemáticos: linear, logaritmo, potencial e exponencial, na relação das variáveis x (dados de idade) e y (comprimento do animal)

Description:

Função que efetua análise exploratória de quatro modelos matemáticos (linear, logaritmo, potencial e exponencial) em duas variáveis, sendo elas x (dados de idade) e y (comprimento do animal). O conjunto de dados x e y são correlacionados e plotados os gráficos em uma única janela. Os quatro modelos supracitados são ajustados, fornecendo as funções matemáticas e seus respectivos parâmetros (valores de a e b), coeficiente de correlação (R^2) e o índice Akaike's information criterion (AIC). O menor valor de AIC é utilizado como critério de escolha do melhor ajuste (modelo) que descreva a correlação dos dados

Usage:

```
melhorajuste(x,y)
```

Arguments:

x variável contendo valores numéricos em série, com valores da idade do animal

y variável contendo valores numéricos em série, com valores do comprimento do animal

Details:

A função sugere o melhor ajuste dentre os modelos testados em dois vetores (x e y), usando como critério de escolha o menor valor de AIC do modelo correspondente

Value:

A função retorna a exibição de uma janela gráfica com a dispersão dos dados. Exibe a curva e ajuste (linha de tendência) de cada modelo testado e sua respectiva fórmula matemática, valor de R², AIC, e uma mensagem indicando o melhor modelo de ajuste aos dados testados

Warnings:

A função aceita valores numéricos armazenados nos objetos x e y. Os vetores x e y devem ter mesmo comprimento. Se houver NA's ou NAN's, estes serão omitidos, diminuindo o número amostral analisado

Note:

Versão beta da função

Author(s):

Diogo Destro Barcellos
contato: dbarcellos@usp.br

Reference:

Akaike, H. 1992. Information theory and an extension of the maximum Likelihood Principle. In: Kotz, S.; Johnson, N.L. Breakthroughs in statistics. Foundations and basics theory, Springer Series in statistics, Springer-Verlag, pgs 610-624

Examples:

#Exemplo de dados 1

```
dados1<-read.table("dados_de_idade_vs_comprimento.txt", header = T, dec=",")  
# objeto "dados1" recebe os valores que serão lidos na tabela do documento  
.txt  
x<- dados1$x # objeto x recebe os valores correspondentes  
y<- dados1$y # objeto y recebe os valores correspondentes  
melhorajuste(x,y) # função em ação
```

#Exemplo de dados 2

```
x<- c(4.84, 3.13, 2.79, 3.41, 4.90, 2.94, 3.34, 3.25, 2.36, 3.99, 3.77,  
2.42, 4.18, 2.59, 4.19, 3.10, 1.45, 2.60, 3.36, 5.04, 3.90, 4.74, 3.13,  
2.78, 3.56, 1.50, 2.64, 3.08, 1.57, 4.95, 2.61, 2.93, 1.39, 2.88, 3.27,  
1.57, 4.95, 2.61, 2.93, 1.39, 2.88, 3.27, 2.40, 2.76, 6.06, 2.44, 2.89,  
3.51, 0.88, 3.91, 3.28, 2.78, 3.27, 3.58, 2.60, 1.99, 1.81, 2.01, 2.62,  
3.14, 4.54, 4.09, 2.86, 2.49, 4.03, 4.41, 0.48, 4.10, 1.72, 1.76, 1.54,  
3.43, 4.19, 2.15, 1.74, 2.28, 3.80, 2.29, 4.74, 2.76, 3.00, 2.56, 4.40,  
3.70, 3.14, 3.12, 4.17, 4.99, 4.05, 2.43, 2.62, 2.17, 2.56, 3.20, 2.46,  
1.54, 3.52, 3.88, 3.55, 3.58)  
y<- c(14.58, 13.93, 13.27, 14.45, 14.51, 14.81, 14.86, 15.33, 16.55, 14.94,  
15.43, 14.02, 15.85, 14.73, 16.32, 15.19, 12.73, 14.64, 15.05, 16.09, 14.40,  
14.12, 16.01, 14.71, 14.07, 14.70, 15.68, 14.57, 15.99, 15.40, 14.41, 15.63,  
15.76, 16.36, 15.38, 14.16, 14.49, 14.66, 15.20, 14.91, 13.82, 12.88, 14.74,  
14.81, 14.08, 14.14, 13.04, 15.32, 15.51, 15.50, 13.95, 14.62, 16.05, 13.59,  
16.68, 13.49, 13.63, 13.56, 16.43, 13.37, 15.90, 14.42, 14.65, 16.04, 15.05,  
16.10, 14.08, 14.01, 15.28, 14.27, 15.34, 15.33, 13.16, 12.28, 14.59, 14.15,  
15.11, 13.50, 14.92, 14.03, 14.39, 15.37, 15.46, 14.60, 15.16, 15.72, 13.61,  
16.49, 15.48, 15.42, 15.52, 14.67, 15.03, 14.07, 13.79, 14.52, 16.07, 17.06,  
15.65, 15.16)  
melhorajuste(x,y)
```

#Exemplo de dados 3:

```
x<- c(3.98, 1.18, 1.63, 0.47, 2.81, 1.54, 1.44, 0.76, 1.42, 1.49, 1.01,  
2.04, 3.08, 1.38, 2.04, 3.36, 2.45, 1.04, 4.39, 2.07, 1.98, 2.37, 2.70,  
1.32, 2.97, 2.62, 1.79, 1.12, 0.45, 3.30, 1.86, 1.02, 1.23, 0.64, 3.20,  
1.40, 2.95, 1.70, 3.03, 2.10, 1.43, 0.41, 1.00, 2.26, 3.66, 3.23, 0.87,  
0.35, 1.38, 1.57, 3.54, 2.04, 3.40, 1.74, 0.68, 0.86, 1.10, 2.64, 1.52,  
3.86, 1.44, 2.16, 1.02, 3.02, 1.10, 2.60, 1.62, 3.61, 3.60, 3.27, 1.11,  
2.66, 1.05, 2.59, 0.01, 0.97, 1.81, 2.00, 2.36, 3.57, 1.64, 3.55, 3.73,  
0.50, 0.49, 1.61, 0.89, 3.08, 2.08, 3.19, 2.96, 2.05, 1.32, 4.03, 0.64,  
1.78, 0.04, 1.28, 3.00, 2.64)
```

```
y<- c(14.05, 16.80, 15.54, 15.05, 13.30, 14.14, 13.39, 15.78, 12.93, 16.59,
15.01, 15.97, 15.85, 14.72, 14.13, 14.37, 15.64, 14.77, 15.48, 16.47, 16.34,
14.87, 17.67, 17.00, 15.21, 16.03, 14.34, 15.98, 14.60, 16.00, 13.41, 15.22,
13.69, 14.44, 14.09, 15.10, 15.88, 15.28, 14.58, 14.75, 14.32, 13.94, 15.19,
15.30, 16.56, 16.57, 17.43, 15.57, 13.74, 14.65, 14.50, 15.65, 16.46, 14.51,
15.19, 14.37, 13.67, 14.09, 14.51, 13.84, 15.05, 15.54, 15.68, 14.38, 13.64,
14.91, 14.49, 16.81, 15.08, 13.50, 14.77, 14.66, 15.43, 13.68, 14.45, 15.57,
15.32, 14.29, 14.22, 15.62, 15.85, 13.81, 13.78, 16.42, 15.45, 15.90, 15.75,
15.71, 15.85, 15.33, 13.52, 15.36, 13.66, 14.04, 13.99, 12.93, 14.65, 14.06,
13.68, 15.59)
melhorajuste(x,y)
```

Código da Função

```
melhorajuste<- function (x,y, rmNA=TRUE, xname="x", yname="y") ## nome da
função: melhorajuste. Esta aceita duas variáveis ("x" e "y", e valores NA)##

{ if(length(x)!=length(y)) ##condição que verifica se as variáveis "x" e
"y" possuem mesmo comprimento.
{
  cat("ATENÇÃO: os vetores x e y possuem tamanhos diferentes") ## Envia
mensagem de alerta.
}
else
{

if (rmNA==T) ## se houver valores NA
{
x<- na.omit(x) ## os valores de NA na variável x, estes valores serão
omitidos
y<- na.omit(y) ## os valores de NA na variável x, estes valores serão
omitidos
namostral<- length (x) - length(x) ## verifica se após a exclusão de valores
NA, o valor do n amostral a ser trabalhado. Para isto, o objeto "namostral"
recebe este valor.
cat("Valores de NA excluídos", namostral) ## envia mensagem de aviso ref. a
valores de Na excluídos.
cat("n amostral", length(x)) ## envia mensagem de aviso do n amostral
trabalhado.
}

#Criando modelos matemáticos:
mod1 <- mod1<- lm(y ~ x) # modelo linear
mod2<- mod2<-lm(y ~ log(x)) # modelo logarítimo
mod3<- mod3<-nls(y~a*x^b, start = c(a=-3, b=2)) # modelo potencial ou com o
valor de start: start = c(a=.01, b=3))
mod4<- nls(y~a*x^exp(b), start = c(a=2, b=1)) # modelo exponencial ou com
valores de start: start = c(a=.1, b=.3))
```

```
#Estimativa do coeficiente de correlação:
RSS.p <- sum(residuals(mod1)^2) # estima valor de R2, coeficiente de
correlação para o mod1
TSS <- sum((y - mean(y))^2)
rmod1<- 1-(RSS.p/TSS)#resultado de R2 do mod1
rmod1<- round(rmod1, digits=2) # arredonda valor de r2 do mod1

RSS.p <- sum(residuals(mod2)^2) # estima valor de R2, coeficiente de
correlação para o mod2
TSS <- sum((y - mean(y))^2)
rmod2<- 1-(RSS.p/TSS) #resultado de R2 do mod2
rmod2<- round(rmod2, digits=2) #arredonda valor de r2 do mod2

RSS.p <- sum(residuals(mod3)^2) # estima valor de R2, coeficiente de
correlação para o mod3
TSS <- sum((y - mean(y))^2)
rmod3<- 1-(RSS.p/TSS) #resultado de R2 do mod3
rmod3<- round(rmod3, digits=2) #arredonda valor de r2 do mod3

RSS.p <- sum(residuals(mod4)^2) # estima valor de R2, coeficiente de
correlação para o mod4
TSS <- sum((y - mean(y))^2)
rmod4<- 1-(RSS.p/TSS) #resultado de R2 do mod4
rmod4<- round(rmod4, digits=2) #arredonda valor de r2 do mod4

#Estimando valor de AIC:
AICmod1<- AIC (mod1) # estima o valor de Akaike (AIC) do mod1
AICmod1<- round(AICmod1, digits=2) #arredonda valor de AIC mod1

AICmod2<- AIC (mod2) # estima o valor de Akaike (AIC) do mod2
AICmod2<- round(AICmod2, digits=2) #arredonda valor de AIC mod2

AICmod3<- AIC (mod3) # estima o valor de Akaike (AIC) do mod3
AICmod3<- round(AICmod3, digits=2) #arredonda valor de AIC mod3

AICmod4<- AIC (mod4) # estima o valor de Akaike (AIC) do mod4
AICmod4<- round(AICmod4, digits=2) #arredonda valor de AIC mod4

#obter melhor posicionamento para plotar no gráfico legendas internas nos
gráficos
pg2<-(max(x)/2)-(max(x)/4) # objeto pg2 baseado nos valores máximos do eixo
c
pgy<-max(y)/3 # objeto pgy baseado nos valores máximos de y
pgy4<-max(y)-pgy # posicionamento no eixo y

#obtenção de coeficientes mod1 (linear)
coefmod1<-as.numeric(coef(mod1)) #armazena coeficientes do mod1 no objeto
"coefmod1". Dados transformados em valores numéricos.
coeflmod1<-coefmod1[1] #indexa valor
valor2<-coeflmod1 #nomeia em um novo objeto
```

```
valor2<-round(valor2 , digits=2) # arredonda valor
coeflmod2<-coefmod1[2] #indexa valor
valor1<-coeflmod2 #nomeia em um novo objeto
valor1<-round(valor1 , digits=2) # arredonda valor

#obtenção de coeficientes mod2 (logaritmo)
coefmod2<-as.numeric(coef(mod2)) #armazena coeficientes do mod2 no objeto
"coefmod2". Dados transformados em valores numéricos.
coeflmod2<-coefmod2[1] #indexa valor
valor2mod2<-coeflmod2 #nomeia em um novo objeto
valor2mod2<-round(valor2mod2 , digits=2) # arredonda valor
coeflmod2<-coefmod2[2] #indexa valor
valor1mod2<-coeflmod2 #nomeia em um novo objeto
valor1mod2<-round(valor1mod2 , digits=2) # arredonda valor

#obtenção de coeficientes mod3 (potencial)
coefmod3<-as.numeric(coef(mod3)) #armazena coeficientes do mod3 no objeto
"coefmod3". Dados transformados em valores numéricos.
coeflmod3<-coefmod3[1] #indexa valor
valor2mod3<-coeflmod3 #nomeia em um novo objeto
valor2mod3<-round(valor2mod3 , digits=2) # arredonda valor
coeflmod3<-coefmod3[2] #indexa valor
valor1mod3<-coeflmod3 #nomeia em um novo objeto
valor1mod3<-round(valor1mod3 , digits=2) # arredonda valor

#obtenção de coeficientes mod4 (exponencial)
coefmod4<-as.numeric(coef(mod4)) #armazena coeficientes do mod4 no objeto
"coefmod4". Dados transformados em valores numéricos.
coeflmod4<-coefmod4[1] #indexa valor
valor2mod4<-coeflmod4 #nomeia em um novo objeto
valor2mod4<-round(valor2mod4 , digits=2) # arredonda valor
coeflmod4<-coefmod4[2] #indexa valor
valor1mod4<-coeflmod4 #nomeia em um novo objeto
valor1mod4<-round(valor1mod4 , digits=2) # arredonda valor

x11() # abre um dispositivo de tela para plotar os gráficos
par(mfrow=c(2,2)) # pareia lado a lado as figuras, para 4 gráficos

plot(y ~ x, xlim = c(0, max(x)), ylim = c(0, (max(y))), pch=16, xlab= "idade
(dias, meses, anos)", ylab = "comprimento (mm, metros)", main = "Modelo
linear", cex.lab = 1.2) # plota em gráfico de dispersão, os dados das
variáveis "x" em relação a "y"
text(pg2, pg4,paste("y=", valor1, "+", valor2), cex=.8) # legenda interna
do gráfico, fórmula do modelo testado
text(pg2, (pg4-(pgy/3.5)), paste("R2=", rmod1), cex=.8) # legenda interna
do gráfico, valores do coeficiente de correlação (r2).
text(pg2, (pg4-(pgy/1.7)),paste("AIC=", AICmod1), cex=.8) # legenda interna
do gráfico, Akaike Information Criterion (AIC).
abline(lm(y~x), col="black",lwd=3) # reta de ajuste do modelo linear

plot(y ~ x, xlim = c(0, max(x)), ylim = c(0, (max(y))), pch=16, xlab= "idade
```

```
(dias, meses, anos)", ylab = "comprimento (mm, metros)", main = "Modelo
logarítimo", cex.lab = 1.2) # plota em gráfico de dispersão, os dados das
variáveis "x" em relação a "y"
text(pg2, pg4, paste("y=", valor1mod2, "ln(x) +", valor2mod2), cex=.8) #
legenda interna do gráfico, fórmula do modelo testado
text(pg2, (pg4-(pg4/3.5)), paste("R2=", rmod2), cex=.8) # legenda interna do
gráfico, valores do coeficiente de correlação (r2).
text(pg2, (pg4-(pg4/1.7)), paste("AIC=", AICmod2), cex=.8) # legenda interna
do gráfico, Akaike Information Criterion (AIC).
xx <- seq(min(x), max(x), length = length(x))
lines(xx, predict(mod2, newdata = list(x = xx, y = y)), col="black", lwd=3)
#ajusta a a curva do modelo logarítimo

plot(y ~ x, xlim = c(0, max(x)), ylim = c(0, (max(y))), pch=16, xlab= "idade
(dias, meses, anos)", ylab = "comprimento (mm, metros)", main = "Modelo
potencial", cex.lab = 1.2) # plota em gráfico de dispersão, os dados das
variáveis "x" em relação a "y"
text(pg2, pg4, paste("y=", valor1mod3, "x^", valor2mod3), cex=.8) # legenda
interna do gráfico, fórmula do modelo testado
text(pg2, (pg4-(pg4/3.5)), paste("R2=", rmod3), cex=.8) # legenda interna do
gráfico, valores do coeficiente de correlação (r2).
text(pg2, (pg4-(pg4/1.7)), paste("AIC=", AICmod3), cex=.8) # legenda interna
do gráfico, Akaike Information Criterion (AIC).
lines(xx, predict(mod3, newdata = list(x = xx, y = y)), col="black", lwd=3)
#ajusta a a curva do modelo potencial

plot(y ~ x, xlim = c(0, max(x)), ylim = c(0, (max(y))), pch=16, xlab= "idade
(dias, meses, anos)", ylab = "comprimento (mm, metros)", main = "Modelo
exponencial", cex.lab = 1.2) # plota em gráfico de dispersão, os dados das
variáveis "x" em relação a "y"
text(pg2, pg4, paste("y=", valor1mod4, "e", valor2mod4, "x"), cex=.8) #
legenda interna do gráfico, fórmula do modelo testado
text(pg2, (pg4-(pg4/3.5)), paste("R2=", rmod4), cex=.8) # legenda interna do
gráfico, valores do coeficiente de correlação (r2).
text(pg2, (pg4-(pg4/1.7)), paste("AIC=", AICmod4), cex=.8) # legenda interna
do gráfico, Akaike Information Criterion (AIC).
lines(xx, predict(mod4, newdata = list(x = xx, y = y)), col="black", lwd=3)
#ajusta a a curva do modelo exponencial

z<-c(AICmod1, AICmod2, AICmod3, AICmod4) #objeto z recebe os valores de AIC
dos modelos
z<-min(z) #menor valor do objeto z
z

if(z==AICmod1)
{
  cat("O modelo linear melhor se ajustou aos dados testados. Valor AIC:")
## Envia mensagem de alerta.
}
```

```
if(z==AICmod2)
{
  cat("O modelo logarítimo melhor se ajustou aos dados testados. Valor AIC:")
## Envia mensagem de alerta.
}

if(z==AICmod3)
{
  cat("O modelo potencial melhor se ajustou aos dados testados. Valor AIC:")
## Envia mensagem de alerta.
}
if(z==AICmod4)
{
  cat("O modelo exponencial melhor se ajustou aos dados testados. Valor
AIC:") ## Envia mensagem de alerta.
}
resultado<-(z)
return(resultado)
}
}

## Testando o funcionamento da função:
melhorajuste(x,y) #Função em ação
```

Arquivos

dados de exemplo: idade vs. crescimento: [dados_de_idade_vs_comprimento.txt](#)

From:

<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2016:alunos:trabalho_final:dbarcellos:start



Last update: **2020/08/12 06:04**