

Pedro Milanesi Lopes



Mestrando em Conservação, Departamento de Ecologia do IB-USP. Projeto “Serviços Ecosistêmicos Hídricos em Paisagens Savânicas com uso de Silvicultura”, orientadora Rozely Santos.

Meu interesse se deve ao fato de ser um tema que envolve biologia, economia, interação com atores sociais e intervenção prática, além de ter como foco a água, condição básica para a existência da vida como a conhecemos.

[exec](#)

Trabalho Final

Arquivo da Função

[alloiso.r](#)

Código da Função

```
allo.iso = function(x, y, xdim=1, ydim=1)    #argumentos: recebem as
dimensões da variável (ex: área recebe 2)
{
  if(xdim==1 & ydim==2)                    #x é linear e y é quadrático
  {
    b.iso = 2                               #o coeficiente b, quando 2, representa
isometria.
  }
  if(xdim==1 & ydim==3)                    #x linear e y proporcional a volume
  {
    b.iso = 3                               #isometria: b é 3.
  }
  if(xdim==2 & ydim==1)                    #x bidimensional(ou quadrático) e y linear
  {
    b.iso = 0.5                             #isometria: b é 0.5
  }
  if(xdim==2 & ydim==3)                    #x bidimensional e y tridimensional
  {
    b.iso = 1.5                             #isometria: b é 1.5
  }
  if(xdim==3 & ydim==2)                    #x tridimensional e y quadrático
  {
    b.iso = 0.67                           #isometria: b é 0.67
  }
  if(xdim==3 & ydim==1)                    #x tridimensional e y linear
```

```
{
  b.iso = 0.33          #isometria: b é 0.33
}
if(xdim==ydim & xdim<=3) #quando x é da mesma natureza que y e menor ou
igual a 3.
{
  b.iso = 1           #isometria quando b é 1
}
else
{
  return(cat("xdim and ydim can only receive 1, 2, or 3 (see help)"))
#outros valores não entram.
}
lr = lm(log(y)~log(x)) #ajustando modelo linear em escala logarítmica
(erros lognormais)
a.lr = coef(lr)[1]     #obtendo o coeficiente a (intercepto)
b.lr = coef(lr)[2]     #obtendo coeficiente b (inclinação)
var.lr = var(log(y) - predict(lr)) #variância para cálculo do AICc
nlr = nls(y~a*x^b, start=list(a=a.lr, b=b.lr)) #ajustando o modelo de
potência (erros normais)
a.nlr = coef(nlr)[1]   #obtendo coeficiente a (constante de
proporcionalidade)
b.nlr = coef(nlr)[2]   #obtendo coeficiente b (coeficiente
alométrico)
var.nlr = var(log(y) - predict(nlr)) #variância para cálculo do AIC
L.lr = sum( 1 / (y*sqrt(2*pi*var.lr)) * exp( -(log(y)-
log(a.lr*x^b.lr))^2) / (2*var.lr))) #'likelihood' de "lr"
L.nlr = sum( 1 / (y*sqrt(2*pi*var.nlr)) * exp( -(y-(a.nlr*x^b.nlr))^2) /
(2*var.nlr))) #'likelihood' de "nlr"
k = 3 #graus de liberdade: 3 parâmetros(a,b e var)
AICc.lr = 2*k - 2*log(L.lr) + (2*k*(k + 1))/(length(x) - k - 1)
#Critério de Informação de Akaike para "lr"
AICc.nlr = 2*k - 2*log(L.nlr) + (2*k*(k + 1))/(length(x) - k - 1) #AICc
para "nlr"
delta.AICc = AICc.lr - AICc.nlr #Diferença entre os valores nos dará as
próximas opções:
if(delta.AICc < -2) #Se AICc.nlr for maior:
{
  ci.b = confint.lm(lr)[2,] #calculando intervalo de confiança para b
a partir de "lr"
  if(min(ci.b)<=b.iso & max(ci.b)>=b.iso) #se o valor "b.iso" definido
acima estiver dentro de "ci.b"
  {
    cat("\n", "\t", "AICc best model: Linear Log-Log Regression", "\n",
"\n", "Relation: Isometry", "\n",
"\n", "Confidence Intevals for b:", "\n", "2.5 %", "97.5 %", "\n",
round(ci.b, 3)) #temos isometria
  }
  if(min(ci.b>b.iso) | max(ci.b>b.iso)) #se o valor "b.iso" estiver
fora do "ci.b"
```

```

{
  cat("\n", "\t", "AICc best model: Linear Log-Log Regression", "\n",
"\n", "Relation: Allometry", "\n",
      "\n", "Confidence Intevals for b:", "\n", "2.5 %", "97.5 %", "\n",
round(ci.b, 3)) #temos alometria
} #para ambos os casos, o modelo linear foi considerado melhor pelo
método AICc.
}
if(delta.AICc > 2) #se AICc.lr for maior:
{
  library(MASS)
  ci.b = confint(nlr)[2,] #calculando intervalo de confiança para b em
"nlr"
  if(min(ci.b)<=b.iso & b.iso<=max(ci.b)) #se "b.iso" estiver dentro
de "ci.b"
  {
    cat("\n", "\t", "AICc best model: Non Linear Regression", "\n", "\n",
"Relation: Isometry", "\n",
        "\n", "Confidence Intevals for b:", "\n", "2.5 %", "97.5 %", "\n",
round(ci.b, 3)) #isometria
  }
  if(min(ci.b)>b.iso | max(ci.b)<b.iso) #se "b.iso" está fora do "ci.b"
  {
    cat("\n", "\t", "AICc best model: Non Linear Regression", "\n", "\n",
"Relation: Allometry", "\n",
        "\n", "Confidence Intevals for b:", "\n", "2.5 %", "97.5 %", "\n",
round(ci.b, 3)) #alometria
  } #em ambos os casos, o modelo escolhido foi "nlr" porque teve menor
valor AICc.
}
else
{
  cat("The models did not differ: AICc Averaging Models required") #não
foi possível completar o último passo
} #seria o caso de não haver diferença significativa (>=2) entre os
modelos e ser necessário calcular o
#peso dos modelos para um modelo intermediário.
X11() #abrindo janela
par(mfrow=c(1,2), bty="l", lwd=2, pch=16, mar=c(12,5,10,1))
#estabelecendo parâmetros para 2 gráficos
plot(y~x, xlab="Indicator Variable", ylab="Response Variable") #gráfico
com eixos normais mostrando ajuste do
#modelo não linear.
abline(lr, col=4, lwd=2) #linha de regressão linear
curve(a.nlr*x^b.nlr, col=2, lwd=2, add=T) #curva da lei de potência
lines(x=c(0.8, 1.0), y=c(6.3, 6.3), col=4) #legenda
lines(x=c(0.8, 1.0), y=c(6, 6), col=2) #legenda
text(x=0.7, y=6.3, "LR", cex=0.8) #legenda
text(x=0.7, y=6, "NLR", cex=0.8) #legenda
plot(log(y)~log(x), xlab="Log(Indicator Variable)", ylab="Log(Response
Variable)") #outro gráfico, mostrando o

```

```
#ajuste do modelo linear.  
abline(lr, col=4, lwd=2)    #linha de regressão linear  
curve(a.nlr*x^b.nlr, col=2, lwd=2, add=T)  #curva da lei de potência  
}
```

Página de Ajuda

allo.iso package:unknown R Documentation

Isometry or Allometry

Description:

```
~~ First compares non-linear models with linear models ~~  
~~ and then verifies isometric or allometric relationships ~~  
~~ considering the best model (method: AIC) ~~
```

Usage:

```
~~ allo.isso(x, y, xdim = 1, ydim = 1) ~~
```

Arguments:

```
~~ x and y: numeric vectors ~~  
~~ xdim: number of dimensions of indicator variable. Can be 1 when linear,  
2 when quadratic and 3 when cubic ~~  
~~ ydim: the same applies for response variable.
```

Details:

```
~~ The method used to choose the best model from which to conclude the  
case of Allometry/Isometry ~~  
~~ is the Akaike Information Criterion (see Xiao Xiao et. Al, 2011)
```

Value:

```
~Returns two graphics with adjusted models and a list which includes:
```

Chosen Model : The best model according to the AICc method.

Type of relationship : If the relationship is considered allometric or isometric.

CI of b: Confidence Intervals for B. To show that the isometric b is inside or outside the interval.

Warning:

```
....
```

Note:

Author(s):

~~Pedro M Lopes~~

References:

~Xiao X., White, E. P., Hooten, M. B., Durham, S. L. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. Ecology. Volume 92, Issue 10, Pages 1887–1894. October 2011.

~Snipes, M., Taylor, D.C. Model selection and Akaike Information Criteria: An example from wine ratings and prices. Science Direct~

See Also:**Examples:**

```
##---- Should be DIRECTLY executable !! ----  
##- - ==> Define data, use random,  
##- - or do help(data=index) for the standard data sets.
```

Oi Pedro! Sou o monitor encarregado de revisar suas propostas. Acho que a proposta B é interessante e dá para explorar idéias mais legais de implementação do que aquela na proposta A. Me deixa um pouco confuso como planeja implementar aquilo e gostaria que desse uma olhada nos meus comentários para entender melhor a sua idéia. Eu sugiro trocar a proposta A por uma C diferente, caso a proposta B apresente problemas mais para a frente.

Seguem meus comentários específicos logo embaixo da seção respectiva e minhas sugestões no cabeçalho de cada proposta. Pode-me encontrar pelo [email](#).

Abs, Gustavo A. Ballen

Proposta A

GAB: Sugiro trocar de idéia nesta proposta

National Sanitation Foundation Index

O índice oficial de qualidade de água norte-americano foi criado no começo dos anos 70 e se consolidou nos EUA nas décadas subsequentes, tendo sido também utilizado em outros países, como a Índia e o Brasil, onde foi utilizado como base para o desenvolvimento do IQA (Índice de Qualidade

de Água).

A importância dos índices de qualidade de água é a transposição de diversos parâmetros, como oxigênio dissolvido, temperatura, coliformes fecais, etc, medidos em unidades e escalas totalmente diferentes, em um único número que seja representativo da qualidade da água no que diz respeito aos parâmetros medidos. Assim, cria-se base para comparações.

A proposta é criar uma função que calcule o índice norte-americano para a qualidade da água, a partir dos parâmetros elegidos nesse índice, mais alguns opcionais, caso se queira customizar o índice.

ENTRADA

Objetos: vetores cujos valores devem seguir uma ordem determinada na página de ajuda. A ordem seria importante para que, na hora do cálculo, os valores entrassem na posição reservada para cada parâmetro. São nove básicos e dois adicionais para este índice.

GAB: Qual a estrutura que os objetos deverão ter? Quero dizer, serão vetores? data frames? listas? de tipo numerico? character? (é uma exageração, mas ajuda a ilustrar o ponto). Acho que vale a pena investir num nível de especificidade maior ao respeito.

Argumentos:

- Parâmetros suprimidos: caso não se utilize algum parâmetro, pode-se indicar qual deles estará faltando.
- Parâmetros opcionais: se o usuário quiser incluir parâmetros, atribui valores a argumentos com os nomes dos parâmetros (que irá procurar no 'help').
- Mudança de pesos: caso o usuário queira mudar o peso que algum parâmetro tem no índice final. Os pesos estão previstos nas fórmulas de cálculo do índice.

SAÍDA

Um número: o índice em questão, com a informação de que ele é o 'default' ou o customizado, e quais as mudanças feitas.

GAB: Entendo que o resultado final de um índice é precisamente resumir diferentes fontes de informação numa medida so. Embora útil em alguns casos, eu acho que neste não irá acrescentar muito à sua experiência em programação com uma implementação que poderia resultar pouco estimulante. E se talvez você considerase uma função que comparasse os índices entre diferentes conjuntos de dados? Eu acho que algo assim iria melhorar a experiência.

Olá Pedro, Concordo com o Gustavo que sua proposta é muito simples. Quais são os valores necessários para calcular o índice? Como é o objeto de entrada? O que a função retorna? Seria interessante poder calcular para uma ou mais localidades. A sua função poderia fazer gráficos comparando os valores do índice entre diferentes locais, por exemplo. Outra opção seria, caso os dados para o cálculo do índice estejam disponíveis on line, sua função fazer a busca dos dados necessários e calcular o índice. Desse modo, o usuário poderia dar a localidade e a função retornaria o índice. Enfim, apenas calcular o índice é uma proposta simples demais. Pense em como torná-la mais desafiadora!
— [Sara Mortara](#)

Proposta B

GAB: Gostei desta, vamos trocar uma idéia para avaliar sua executabilidade

Relações Isométricas e Alométricas

Muitas áreas da biologia utilizam como ferramenta de análise de dados a verificação da natureza da relação entre duas variáveis no que diz respeito à variação ou não da proporção entre elas, ou seja, relações alométricas ou isométricas. Normalmente, isso é aplicado a medidas de tamanho, volume e massa, porém há outras aplicações possíveis, como, por exemplo, mudanças na resposta de um organismo em relação a algum fator do ambiente.

GAB: Refere-se a fatores como variáveis categóricas ou a outras variáveis que poderiam ter uma relação com aquelas variáveis de interesse?

A ideia é criar uma função que estabeleça qual o tipo de relação entre dois vetores.

GAB: Inclusive se não existisse nenhuma? Nem todo par de variáveis numéricas vai apresentar algum tipo de relação.

ENTRADA

Objetos: 2 vetores ('x' e 'y') numéricos. Como a função tem a proposta de verificar se há relações alométricas ou isométricas, imagina-se que o usuário está lidando com dados que normalmente se relacionam, de alguma forma, em termos de mudanças direcionais em tamanho, peso, etc. Ou seja, já se espera alguma dessas respostas.

GAB: De que tipo? Imagino que numeric, mas isso deve sere explicitado na proposta.

Argumentos:

- Um para definir se os valores são unidimensionais (como tamanho ou algum valor pontual), bidimensionais (área, normalmente) ou tridimensionais (volume). Desta forma, pode-se corrigir a relação entre vetores com naturezas diferentes de valores. Como a área tem, automaticamente, uma relação alométrica com medidas unidimensionais (ou seja, é sempre o quadrado do valor linear), por exemplo, há a necessecidade de descontar isso na equação da regressão. Se uma medida linear dobra, naturalmente a área relacionada quadruplica e isso comunica uma mudança proporcional, portanto seria considerado alometria somente quando a área aumentasse em mais ou menos de quatro vezes. Assim, o expoente da equação teria que ser dividido por 2, descontando o efeito da bidimensionalidade.

GAB: A estrutura do(s) argumento(s) também deve-se explicitar na proposta. Além disso acho legal você explicar como planeja realizar aquela correção que menciona.

SAÍDA

- Mensagem informando qual a relação encontrada, podendo esta ser "isometria", "alometria linear positiva", "alometria linear negativa" ou "alometria curvilínea", esta última para casos em que o melhor modelo é aquele que considera mudanças não só na proporção entre as variáveis, mas na "taxa" de mudança dessa proporção, ou seja, um modelo polinomial. Os casos de alometria linear positiva são aqueles nos quais a regressão possui uma inclinação maior do que a regressão da isometria, enquanto para a negativa, a inclinação é menor. Pode ainda ser que nenhuma das relações seja apropriada, gerando uma resposta como "não há relação direcional", ou algo do tipo.

GAB: O que seria um melhor modelo? Como pretende testar ou decidir (quantificar?) aquilo? Concordo com que é possível, mas acho que isso também deve ser claro na proposta, até porque irá permitir você pensar na melhor maneira de implementar aquilo na função.

- Gráfico: eixos logarítmicos, dados, linha de isometria (1:1) e linha preditória do modelo que melhor

se ajustou aos dados.

GAB: Você acha que seria sempre necessário que os eixos fossem logarítmicos? e se a relação fosse isometria? Não iria isso atrapalhar a visualização da relação entre as variáveis? Da mesma maneira, se a relação não fosse exponencial, os eixos logarítmicos tenderiam a gerar uma visualização errada da relação. Concordo na utilidade de tais eixos, mas só em casos nos quais a relação é exponencial (e.g., $\log(a^x) = x \cdot \log(a)$).

• Teste anova para os modelos: isometria, alometria linear positiva, alometria linear negativa e alometria curvilínea. A ideia é que um modelo melhor resolva uma parte significativa dos resíduos e significativamente faça diferença em relação aos outros.

GAB: E se o modelo não seguisse as suposições de um anova? Como iria a função decidir aquilo? Planeja deixá-lo ao critério do usuário? Vale a pena dar uma olhada em formas alternativas de testar modelos tais como o critério de informação de Akaike (a.k.a. [AIC](#))

Pedro, não entendi bem esta proposta. O que você quer automatizar na função me parece algo que necessita de uma interpretação acurada do pesquisador para decidir o próximo passo. Me preocupa colocar isto numa caixa preta dentro da função, dado que as relações entre duas variáveis podem ter diferentes relações e depende de que tipo de processo a variável descreve. Por fim, para comparar três modelos o AIC seria uma melhor métrica do que anova. Mas eu ainda não estou segura de que a maneira que você propõe é a melhor para isto. Acho mais produtivo seguir com a primeira proposta, adicionando algum desafio a ela 😊 do que esta. — [Sara Mortara](#)

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

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2016:alunos:trabalho_final:pedro.milanesi.lopes:start

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