Monique Nouailhetas Simon

Mestre em Fisiologia (Ecofisiologia), doutoranda do Dep. de Genética e Biologia Evolutiva USP, Laboratório de Evolução de Mamíferos

exec

Propostas de Trabalho Final:

Plano A: função para testar **efeito de deriva genética** na diversificação morfológica dos crânios de um grupo de espécies, por comparação de matrizes de divergência morfológica dentro de uma espécie (W, de “within groups”) e entre espécies (B, de “between groups” - as tarefas seriam: 1. Ler arquivos csv (matrizes de covariância); 2. Retirar os autovalores das matrizes W e B, e guardá-los em um novo objeto; 3.Calcular uma regressão linear entre os conjuntos de autovalores de ambas as matrizes, e guardar em um outro objeto; 4. Identificar se o intervalo de confiança do coeficiente da regressão contém o valor 1 (de igualdade de autovalores entre as matrizes): se contiver, não se rejeita deriva; se não contiver, rejeita-se deriva.

Plano B: função para determinar a **similaridade entre duas matrizes de correlação** - tarefas: 1. Ler arquivos csv (matrizes); 2. Calcular a correlação de Pearson de cada matriz e guardar em um novo objeto; 3. Construir uma distribuição de correlações aleatórias por meio das matrizes originais com suas linhas e colunas permutadas (cenário nulo de DIFERENÇA entre as matrizes causadas pelas permutações) ; 4. Determinar se o valor de correlação empírico pertence ao intervalo de confiança dessa distribuição aleatória de correlações entre matrizes (se , as matrizes são similares).

Comentários das propostas (Leo)

As duas propostas parecem interessantes, úteis e factíveis. Só uma precaução: tente achar um balanço entre simplicidade e complexidade, ou seja, uma função que não seja trivial, mas que também não seja impossível de realizar no prazo. Ao longo do trabalho você deve perceber o tamanho do desafio e poderá adaptar a proposta de acordo com isto.

Projeto final repaginado

Continua sendo um **teste de efeito de deriva genética** para a diversificação morfológica de um grupo, porém após um melhor entendimento do teste, explico melhor seu sentido e corrijo aqui a descrição das etapas do teste. A idéia é a mesma de uma ANOVA, na qual se compara a variação morfológica dentro do grupo (por meio da matriz W, uma estimativa da matriz ancestral do grupo, que é ponderada pelo número de indivíduos de cada população) com a variação morfológica entre grupos (por meio do vetor B, relacionado com as médias dos caracteres morfológicos de cada taxon). Como esse teste foi pensado para evolução multivariada, com várias medidas para representar uma parte maior do organismo, para simplificar a comparação entre B e W, esta última é reduzida a seus autovalores e autovetores, e B é calculada pelas projeções das médias nos autovetores de W. Se a

diversificação do grupo foi por deriva, espera-se que a variação entre grupos seja proporcional à variação dentro do grupo, ou seja, a variação dos taxa terminais foi completamente dependente da variação inicial presente no ancestral do grupo. A proporcionalidade é avaliada com base na seguinte equação linear:
ln(Bi)=ln(t/Ne)+ βln(Wi),
onde **B** é a variância entre grupos, **W** é a variância dentro do grupo para o autovetor i, **t** é o tempo de divergência em gerações, e **Ne** é o tamanho efetivo da população (Akrmann e Cheverud, 2002). Se o intervalo de confiança de 95% de β incluir o valor 1.0 (proporcionalidade), a deriva não pode ser rejeitada.
Etapas do teste

1. A partir de um data frame contendo medidas lineares de partes dos organismos e o nome das taxa, criar matrizes de variância/covariância para cada taxon;
2. Calcular a matriz W utilizando as matrizes dos taxa e o número de indivíduos em cada taxon;
3. Retirar e guardar os autovalores e autovetores de W;
4. Projetar as médias dos caracteres morfológicos nos autovetores de W;
5. Calcular a variância entre as médias projetadas para cada autovetor de W (vetor B);
6. Construir uma regressão entre ln(B) e ln(autovalores de W);
7. Verificar se o intervalo de confiança de β contém o valor 1.0.

Help da função

deriva package:lem_packs R Documentation

Description:

```
  Testa se a deriva genética foi a causa da diversificação morfológica de um
grupo monofilético de taxa.
  A variação morfológica dentro do grupo (matriz W) é comparada com a
variação entre grupos
  (vetor B de médias).
```

Usage:

deriva=function(dados)

Arguments:

dados data frame. As colunas sao as medidas morfometricas, mais uma coluna com os nomes dos taxa. Essas colunas podem ter qualquer nome.

Details:

```
  A matriz W representa a matriz ancestral do grupo monofilético. Ela é
reduzida a seus componentes principais, e as médias dos taxa são projetadas
nos autovetores de W para se calcular a variancia entre grupos de cada
autovetor (vetor B). Por fim, é feita uma regressao entre log(B) e
log(autovalores de W) e o intervalo de confiança do coeficiente de
inclinacao da reta de regressao é avaliado.
```

Value:

```
   Lista:
```

```
"Coeficiente" : coeficientes da regressao de log(B) por log(Autovalores de
W)
```

```
"Intervalo de confianca 95%" : intervalos de confianca dos coeficientes da
regressao

"PCA de W": autovalores e autovetores da matriz W
```

```
"Vetor B": variancia das medias dos caracteres, por taxa, projetadas nos
autovetores de W
```

```
   Graficos:

 graficos diagnoticos da regressao log(B) em log(Autovalores de W)

 grafico de log(B) por log(Autovalores de W), com a reta de regressao
```

Warning:

```
 O objeto de entrada (dados) precisa estar configurado como descrito em
Arguments.
```

```
 As matrizes de variancia/covariancia sao criadas com a funcao cov e os
argumentos method="pearson" e use="complete.obs". Este ultimo argumento
exclui do calculo das matrizes pares de caracteres a serem correlacionados
que apresentam NAs.
```

Author(s):

```
Monique Nouailhetas Simon
Guilherme Garcia (palpites muito úteis)
```

References:

```
   Ackermann, R.G. e J.M. Cheverud (2002). American Journal of Physical
Anthropology, 117:260-271.
   Marroig, G. e J.M. Cheverud (2004). American Naturalist, 163:417-428.
```

See Also:

```
  cov(), eigen(), lm()
```

Examples:

```
  #Exemplo com o objeto iris:
```

```
   deriva(iris)
```

Código da função

```
deriva <- function(dados)
{
medidas <- dim(dados)[2]-1 #número de caracteres (medidas)
grupos <- unique(dados[,medidas+1]) #vetor com os nomes dos taxa
resultado <- list() #cria uma lista vazia
for(i in 1:length(grupos)) #laço para subdividir dados por taxa
{
  tabela <- subset(dados,dados[,medidas+1]==grupos[i],select=1:medidas)
  resultado[[i]] <- tabela
}
matrizes <- lapply(resultado,FUN=cov,method="pearson",use="complete.obs")
#cálculo das matrizes de V/CV por taxa
resultado.1 <- array(data=0, dim=(c(dim (matrizes[[1]]),length(matrizes) )))
for(i in 1:length(matrizes)) #laço para multiplicar o número de individuos
por taxa pela sua matriz de V/CV
{
  matrizes.1 <- length(resultado[[i]][,1])*matrizes[[i]]
  resultado.1[,,i] <- matrizes.1
}
W <- apply(X=resultado.1,MARGIN=c(1,2),FUN=sum)/length(dados[,1]) #Cálculo
da matriz ponderada pelo número de individuos
W.pc <- eigen(W, only.values=FALSE) #PCA em W
 names(W.pc) <- c("Autovalores", "Autovetores")
medias <- aggregate(dados[,1:medidas],FUN=mean,
by=list(Taxa=dados[,medidas+1]))
col.medias <- dim(medias)[2]
medias.Wpc <- as.matrix(medias[,2:col.medias])%*%W.pc$Autovetores #projecao
das medias dos taxa nos autovetores de W
B <- apply(medias.Wpc,2,var) #cálculo da variancia entre grupos
regressao <- lm(log(B)~log(W.pc$Autovalores))
par(mfrow=c(2,2))
plot(regressao) #graficos diagnosticos da regressao
par(mfrow=c(1,1))
plot(log(B)~log(W.pc$Autovalores), xlab="log(Autovalores de W)")
abline(regressao) #reta de regressao
coeficientes <- coef(regressao)
intervalo <- confint(regressao)
cat("\n\t", "Se o intervalo de confiança de log(W.pc$Autovalores) contiver o
valor 1.0, aceitar deriva")
 objeto <- list("Coeficiente"=coeficientes,"Intervalo de confianca
95%"=intervalo,"PCA de W"=W.pc,"Vetor B"=B)
return(objeto)
}
```

deriva.r

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

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2013:alunos:trabalho_final:monique.simon:start

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