



# Anna Penna

Aluna de Graduação em Ciências Biológicas pela USP

Projeto de Iniciação Científica em andamento no Laboratório de Evolução de Mamíferos, LEM, no Dpto de Genética e Biologia Evolutiva da Universidade de São Paulo

*“Modularidade e Padrões de Integração em Linhagens Seleccionadas de Camundongos”*

Orientada pelo Prof<sup>o</sup> Gabriel Marroig

## Curso 2012

[exec](#)

### TRABALHO FINAL: CRIAR UMA FUNÇÃO

**PLANO A: Uma metodologia utilizada na comparação de matrizes V/CV, é a das Adagas Casualizadas (Random Skewers).**

Matrizes de Variância e Covariância (V/CV) são uma medida quantitativa importante da associação entre traços em um sistema. Dessa forma, tem um papel central em estudos de biologia evolutiva. Para realizar tais estudos, é necessário comparar matrizes.

Essa tecnica procura quantificar a diferença entre matrizes ao comparar os vetores de resposta evolutiva de cada matriz quando submetidas ao mesmo vetor de seleção. Ela é a aplicação da equação de resposta multivariada à seleção proposta por Lande, 1979:

$$\Delta z = G * \beta,$$

onde  $\Delta z$  é o vetor de resposta evolutiva (representa as diferenças entre as médias de cada carater em duas matrizes, uma antes e outra depois da seleção),  $G$  é a matriz de V/CV genética e  $\beta$  é o vetor de seleção.

A primeira parte dessa comparação consiste em gerar um grande número de vetores aleatórios de seleção ( $\beta$ ) e obter os respectivos vetores de resposta evolutiva ( $\Delta z$ ) para cada matriz, através de multiplicação simples. Essa primeira parte acaba por gerar uma distribuição aleatória de  $\Delta z$ , onde cada um dos vetores aleatórios gerados terá um par de vetores de resposta à seleção associado, um para cada matriz a ser comparada. Em seguida é feita a correlação desses diversos pares de  $\Delta z$ . A última parte consiste em tirar a média dessas comparações.

Passo a passo:

1. Criando os vetores aleatórios de seleção (com um numero de linhas igual ao da dimensão das matrizes).

- serão criados 10.000 vetores aleatórios que devem ser capazes de explorar uniformemente todas as direções do hiperespaço.
- Para tanto, esses vetores devem respeitar uma distribuição esférica simétrica. Para que isso ocorra, cada elemento desses vetores deve ser sorteado a partir de uma distribuição normal de média zero e variância arbitrária.
- Depois, todos os vetores gerados são normalizados, para que tenham tamanho igual a 1.

2. Multiplicar esses 10.000 vetores gerados por cada uma das matrizes, o que vai gerar 10.000 vetores de resposta à seleção para cada uma das matrizes. Cada vetor de seleção terá, então, um par de vetores  $\Delta z$  associado.

3. Comparar as respostas evolutivas de cada matriz (os pares de  $\Delta z$  gerados por  $\beta$ ) através de correlação de vetores, criando uma distribuição aleatória de correlações de  $\Delta z$ .

4. Extrair a média da correlação dos  $\Delta z$ , que será usada como a medida de similaridade de cada matriz.

---

## PLANO B:

## Comentários

## PRÓLOGO

### FALTOU O PLANO B!.

Oi Ana, sei que o “Clã das Adagas Voadoras” é ótimo e que sua proposta deve ser pertinente, mas todos os “coleguinhas” postaram o plano B. Temos algumas hipóteses sobre alunos que não o fazem :

- não houve tempo hábil para fazê-lo (desorganizado!)
- acredita que não precisa de um plano B
- não está interessado na disciplina
- acha que as regras não se aplicam a ele
- ... (tipo brigou com a namorada!)

Perceba que qq que seja a resposta (inclusive “...”) ela depõem contra vc. Minha impressão é outra: uma aluna interessada e motivada, ou seja, o tipo de aluno para o qual pensamos e fazemos esse curso. Um plano B não iria fazer mal algum, talvez até fizesse vc. avançar para planos futuros, afinal temos que “Dominar o Mundo”.

### Aquilo que interessa

Sabia que Random Skewers foi primeiro aplicado para a Ecologia por Manly? Ele é o “todo poderoso” da permutação e randomização! A aula final do curso é em homenagem a ele! Acho a proposta interessante e como vc. não tem alternativa, é ela que irá tocar! Tenho alguns pontos para pensar ou que talvez eu não tenha entendido direito (não é minha praia!):

- os pseudovalores para compor o  $\beta'$  (beta aleatório) não são amostrado de que distribuição uniforme de 0 a 1?
- Se a ideia é só ver a correlação (similaridade) entre as matrizes, porque tanta volta? Não seria mais facil comparar as matrizes diretamente? Qual a informação adicional que esse método fornece além de um teste de Mantel entre duas matrizes?

A implementação não parece difícil, vc. já produziu um esboço de pseudo-código, fica mais fácil ainda. Acho que uma boa alternativa é gerar o  $\beta'$  aleatório e operar as duas matrizes simultaneamente e calcular as correlações dos dois vetor  $\Delta z'(i)$  formados em um mesmo ciclo! Há jeitos mais eficientes (sem ciclos), mas vai dar mais trabalho... desse jeito fica mais intuitivo e, se as matrizes não forem muito grandes não vai demorar muito. Quando a função tiver rodando me avisa que vou te mostrar um jeito de incluir uma janela com uma barra de progressão da simulação...até com o tempo que já demorou processando.. bacaninha!

— [Alexandre Adalardo de Oliveira](#) 2012/04/03 19:21

## FUNÇÃO

```
Crazy.Skewers<- function(A, B, beta=1000)
{
  if # verificando as condições de entrada das matrizes
    (!(dim(A)[1] == dim(A)[2] & dim(B)[1] == dim(B)[2] &
dim(A)[1] == dim(B)[1]))
    { cat("Review your matrix dimensions, sweetheart!","\n")
      return(-1)
    }
  size<- dim(A) [1]
  #criando a matriz de betas aleatórios
  BETAmx<- matrix (data=rnorm(size * beta) , nrow= size,
ncol=beta, dimnames=list (c(1:size), c(paste("b", 1:beta))))
  normalz<- function ( x )
    {n.x<- x / sqrt(t(x) %*%(x))
      return(n.x)
    }
  N.BETAmx<- apply(BETAmx, 2, normalz)# normalizando os betas
gerados
  #calculando a correlacao entre os pares de Delta zes de cada
Beta aleatório (cada um referente a uma das matrizes)
  cor<- apply((apply (A %*% N.BETAmx, 2, normalz)) * (apply (B
%*% N.BETAmx, 2, normalz)), 2, sum )
  result<- c( mean(cor), sd(cor) )
  names(result) <- c("Media" , "SD")
  return(result)
}
```

[crazyskewers.r](#)

## HELP DA FUNÇÃO

Crazy.skewers {XingLing} package: Filme Chines B R  
Documentation

~~Random Skewers: comparando duas matrizes de variancia/covariancia~~

### Description:

~~ Compara duas matrizes atraves da correlacao entre seus respectivos vetores resposta (DeltaZes) a vetores aleatorios (beta) gerados\ a partir de uma distribuicao normal de media zero. ~~

### Usage:

~~Crazy.Skewers (A, B, beta=1000) ~~

### Arguments:

~~ A uma matriz de Var/Covar  
B uma matriz de Var/Covar  
beta numero desejado de vetores aleatorios gerados. default= 1000 ~~

### Details:

~~ A e B devem ser matrizes quadradas e simetricas, de mesma dimensoes. Se essas condicoes na forem verificadas, retorna uma mensagem de Erro para vc conferir as dimensoes das matrizes. ~~

### Value:

~~ De acordo com as dimensoes das matrizes A e B, a função cria um numero beta de vetores aleatórios apartir de uma distribuicao\ normal de media zero, cujos comprimentos sao iguais a dimensao das matrizes. As matrizes são, entao, multiplicadas por cada um dos vetores aleatórios gerados normalizados e em seguida) sao comparados\ atraves da medida da correlacao entre eles. A funçãofo retorna a media e o desvio padrao dos pares de correlacoes verificados.~~

### Author(s):

~~Anna Penna~~

### References:

~~ Cheverud, J & Marroig, G. (2007) 'Comparing covariance matrices: Random Skewers method compared to the Common Principal Components model', Genetics and Molecular Biology, vol. 30, pp. 461-469.

Marroig, G.; Melo D.; Porto, A.; Sebastiao, H.; Garcia, G. (2011) 'Selection Response Decomposition (SRD): A New Tool for Dissecting Differences and Similarities Between Matrices" , *Evol Biol*, vol. 38, pp. 225â–241

Examples:

```
# criando matrizes para entrar na funcao. Aqui voce pode escolher as
dimensoes de cada uma delas (1 e 2) em "nrow=" e "ncol="
row1<-
col1<-
row2<-
col2<-
M1<- matrix(data= 0, nrow=row1, ncol=col1, dimnames= list (c(paste("A",
1:col1)), c(paste("A", 1:row1)) ) )
M1[lower.tri(M1, diag= FALSE)]<- rnorm( (length(M1[lower.tri(M1,
diag=TRUE)]) - row1) , mean= 0, sd=1)
M2<- matrix(data= 0, nrow=row2, ncol=col2, dimnames= list (c(paste("B",
1:col2)), c(paste("B", 1:row2)) ) )
M2[lower.tri(M2, diag= FALSE)]<- rnorm( (length(M2[lower.tri(M2,
diag=TRUE)]) - row2), mean= 0, sd=1)
Mx1<- M1 + t(M1)
Mx2<- M2 + t(M2)
diag(Mx1)<- rnorm(row1, mean= 0, sd=1)
diag(Mx2)<- rnorm(row2, mean= 0, sd=1)
Crazy.Skewers( Mx1 ,Mx2 , beta= 10000)
```

[description\\_help\\_crazyskewers.r](#)

From:

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

Permanent link:

[http://ecor.ib.usp.br/doku.php?id=05\\_curso\\_antigo:alunos2012:alunos:trabalho\\_final:anapontopenna:start](http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:alunos2012:alunos:trabalho_final:anapontopenna:start)



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