

Análise de redundância - RDA

Gabriela Medeiros e Marília Melo Favalesso

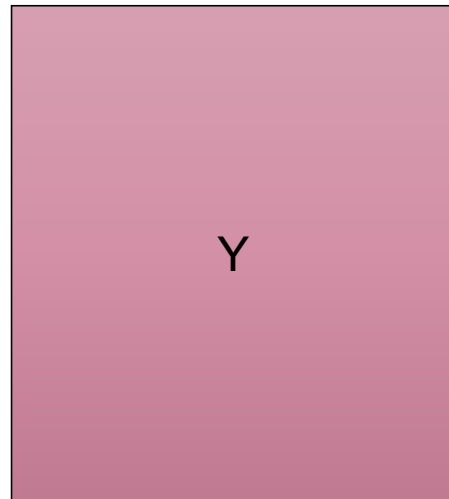
Mestrandas do Programa de Conservação e Manejo de Ambientes Naturais -
UNIOESTE, Cascavel.



- **Objetivo:** Estabelecer uma relação causal entre uma matriz de variáveis descritoras (Y) com uma matriz de variáveis preditoras (X); ordenação.
- Regressão múltipla multivariada \rightarrow PCA
- Criar uma ordenação com a relação de $Y \sim X$.
- É uma análise canônica (≥ 2 matrizes de dados)
- Rao (1965) / Wollenberg (1977)

RDA x outras análises

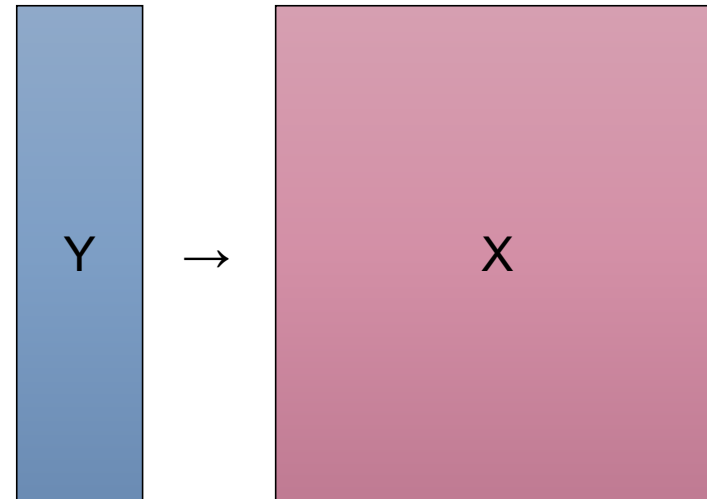
**Ordenação simples
da matriz Y**



Análise de Componentes Principais (PCA)

Análise de correspondência (CA)

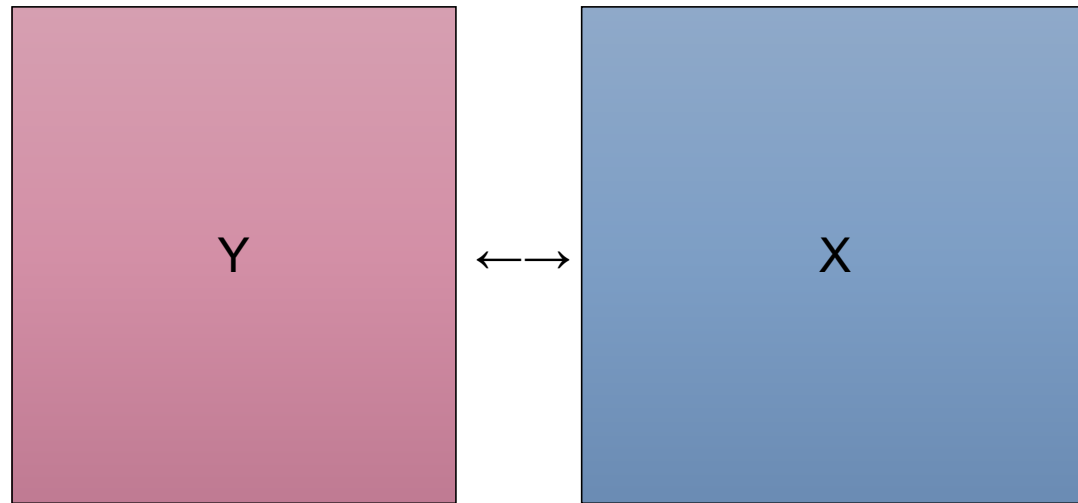
**Ordenação de Y (eixo único) sob
restrição de X**



Regressão múltipla

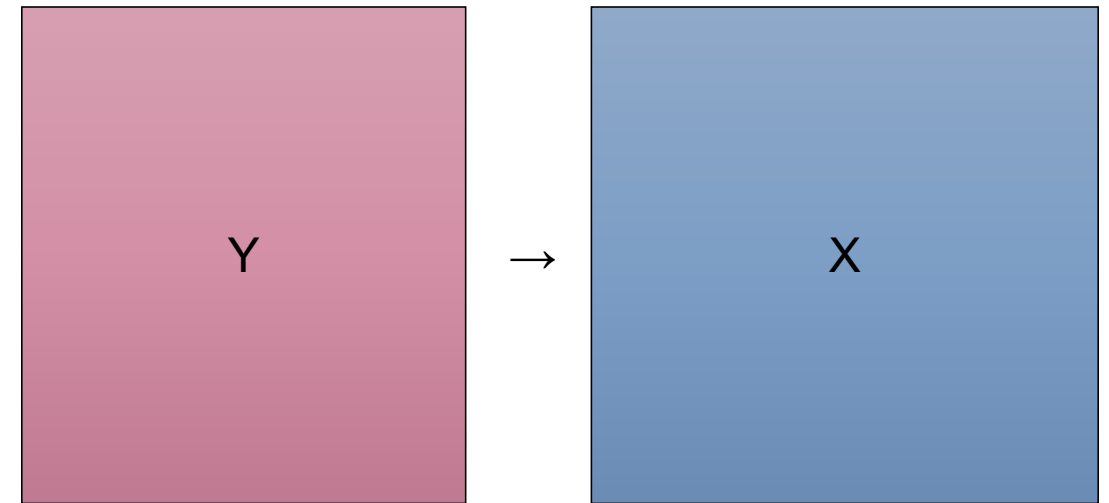
RDA x outras análises

Ordenação de Y sob restrição de X



Análise de correlação canônica (CCoA)

Ordenação de Y sob restrição de X

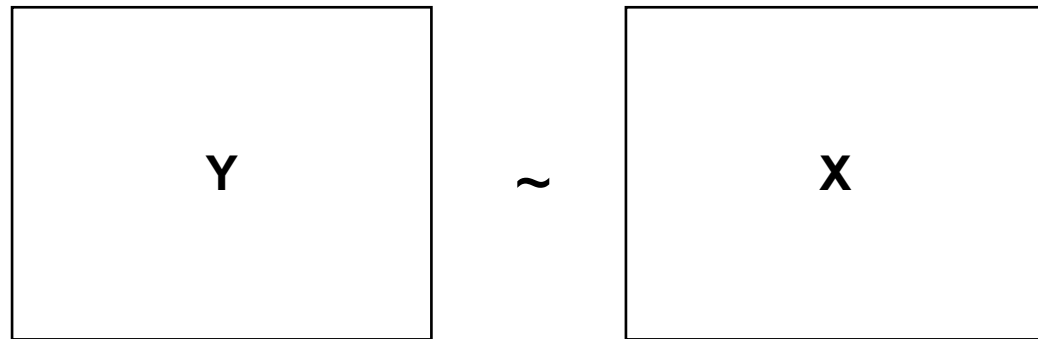


Análise de redundância (RDA)

Preparando a matriz Y e X

- Matriz de variável resposta Y deve possuir um tamanho $n \times p$, onde n é o tamanho de objetos e p o número de variáveis.
- A matriz de variáveis independentes deve possuir um tamanho $n \times m$ com $m \leq n$.
- Centralização das variáveis em suas médias ou padroniza-las caso as dimensões não sejam homogêneas (por exemplo, T°C, pH, etc)*

Regressar cada uma das variáveis resposta Y_j de \mathbf{Y} com as variáveis de \mathbf{X} e calcular os valores ajustados.



$$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon_j$$
$$\hat{Y} = X\hat{B}$$
$$\hat{B} = (X^T X)^{-1} X^T Y$$

The diagram shows a flow from the regression equation to the matrix equation $\hat{Y} = X\hat{B}$. An arrow points from the matrix equation to the vector of predicted values $\hat{Y} = [Y_j]$. Another arrow points from the matrix equation to the matrix equation $\hat{B} = (X^T X)^{-1} X^T Y$.

Fazer a matriz de variância/covariância correspondente a tabela de valores ajustados

$$S_{\hat{Y}' \hat{Y}} = [1/(n-1)] \hat{Y}^T \hat{Y}$$



Realizar uma análise de componentes principais com a finalidade de reduzir a dimensionalidade dos dados



Encontramos os autovalores e uma matriz de autovetores (matriz **U**)



Como na PCA, a **Matriz U** são os escores das variáveis da Matriz **Y**

RDA: Passo 4, 5 e 6

Gerar dois conjuntos de escores a partir da **matriz U**:

Escores dos locais no espaço de
variáveis resposta Y



$$F = YU$$

Escores ajustados dos locais no
espaço de variáveis preditoras X



$$Z = \hat{Y}U = XBU$$

Escores ajustados dos locais



$$C = BU$$

(Contribuição das
variáveis preditoras)



“Essa decomposição equivale a dizer que os valores em cada uma das colunas C são iguais aos coeficientes de regressão padronizados da matriz X” Gotelli & Ellison (2011)

Exemplo de RDA

A forma da concha do caramujo *Lirroraria angulifera* foi medida para amostras de 9 países

Matriz resposta Y: Medidas de conchas de caramujos

País	Proporcionalidade	Circularidade	Altura do epiral
Angola	1.36	0.76	1.69
Bahamas	1.51	0.76	1.86
Belize	1.42	0.76	1.85
Brasil	1.43	0.74	1.71
Flórida	1.45	0.74	1.86
Haiti	1.49	0.76	1.89
Libéria	1.36	0.75	1.69
Nicarágua	1.48	0.74	1.69
Serra Leoa	1.35	0.73	1.72

Matriz resposta X: Variáveis ambientais

País	Precipitação	Nº de meses secos	T°C média mensal	Altura média do dossel
Angola	363	9	26.4	30
Bahamas	1.181	2	25.1	3
Belize	1.5	2	29.5	8
Brasil	2.15	4	26.4	30
Flórida	1.004	1	25.3	10
Haiti	1.242	6	27.5	10
Libéria	3.874	3	27	30
Nicarágua	3.292	0	26	15
Serra Leoa	4.329	4	26.6	35

Pergunta: Como o formato das conchas covariam com os locais e as suas variáveis ambientais?

Exemplo de RDA

Matriz resposta Y

1.36	0.76	1.69
1.51	0.76	1.86
1.42	0.76	1.85
1.43	0.74	1.71
1.45	0.74	1.86
1.49	0.76	1.89
1.36	0.75	1.69
1.48	0.74	1.69
1.35	0.73	1.72

~

Matriz preditora X

363	9	26.4	30
1.181	2	25.1	3
1.5	2	29.5	8
2.15	4	26.4	30
1.004	1	25.3	10
1.242	6	27.5	10
3.874	3	27	30
3.292	0	26	15
4.329	4	26.6	35

$$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon_j$$

P.S: As variáveis foram transformadas

Matriz de valores ajustados

$$\hat{Y} = XB$$

$\hat{Y} =$

-0.05	0.01	-0.04
0.09	0	0.09
0	0.01	0.08
-0.05	-0.01	-0.08
0.05	0	0.03
0.04	0.02	0.09
-0.05	-0.01	-0.07
0.03	-0.01	0
-0.06	-0.01	-0.10

Matriz var/covar

$$S_{\hat{Y}'\hat{Y}}$$



Análise de componentes principais

$U =$

0.55	-0.81	-0.19
0.08	0.27	-0.96
0.83	0.52	0.21

Matriz autovetores

Exemplo de RDA

Escores dos locais

$$F = YU$$

$$F =$$

2.21	-0.03	-0.63
2.44	-0.06	-0.62
2.38	0	-0.61
2.26	-0.08	-0.62
2.4	-0.02	-0.59
2.45	-0.03	-0.61
2.21	-0.03	-0.62
2.28	-0.13	-0.63
2.23	-0.01	-0.59

Matriz dos escores ajustados dos locais

$$Z = \hat{Y}U = XB U$$

$$Z =$$

-0.06	0.02	-0.01
0.13	-0.03	0
0.07	0.04	0.01
-0.09	0	0
0.05	-0.02	0
0.1	0.02	0
-0.09	0	0
0.01	-0.03	0
-0.12	0	0

***CP1 explica 94% da variância total dos dados**

Contribuição da Matriz X



$$C = BU$$

0.01	0	0
0.03	0.01	0
0	0.02	0
-0.11	0	0

Correlação das variáveis ambientais X com a matriz F

	Componente 1	Componente 2	Componente 3
Precipitação	-0.55	-0.19	0.11
Nº de meses secos	-0.28	0.34	-0.18
T°C média mensal	0.02	0.49	0.11
Altura média do dossel	-0.91	0.04	-0.04

Exemplo de RDA

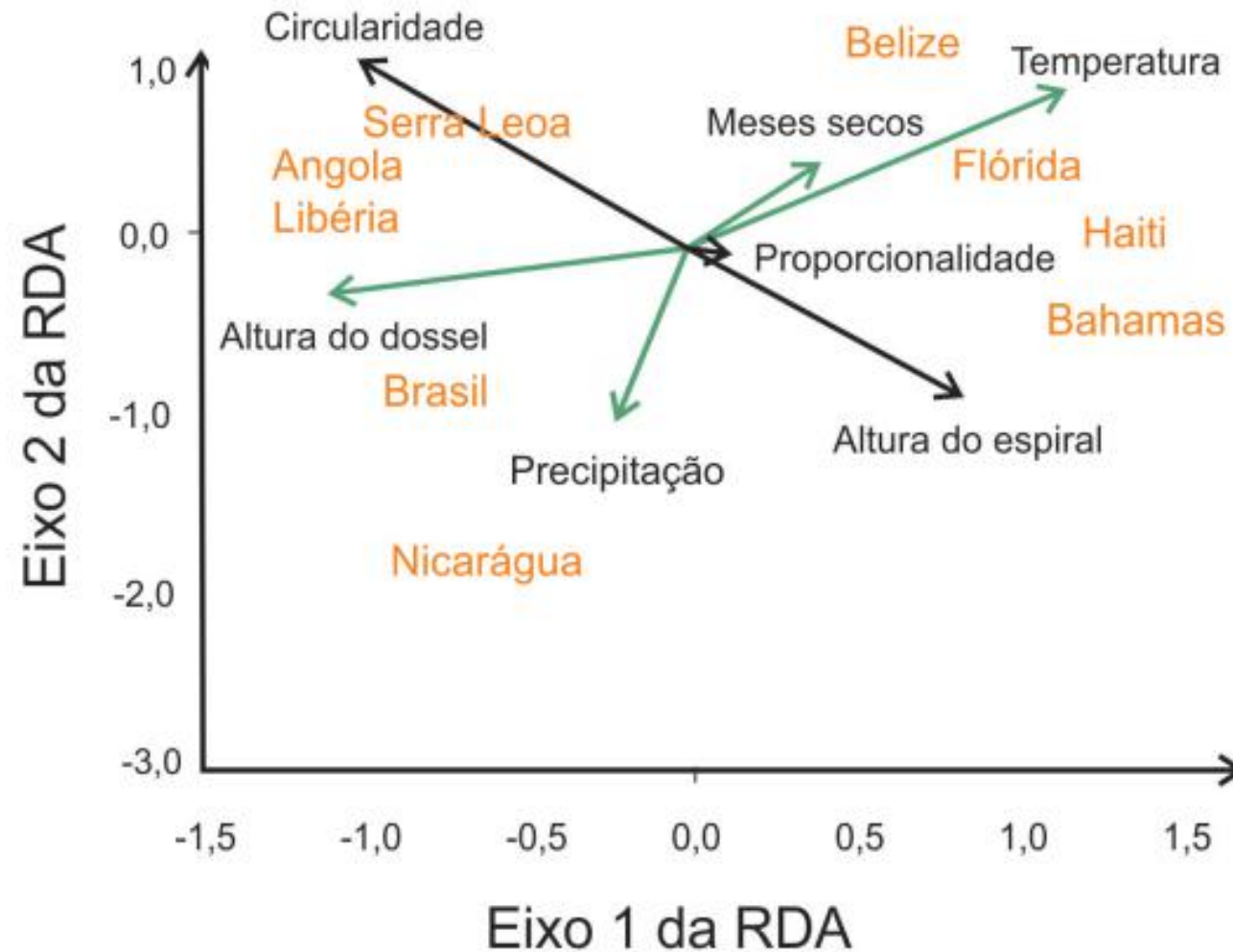


Figura 1 - Bi-plot dos dois primeiros eixos da RDA, da regressão dos dados de conchas de caramujos contra os dados ambientais, sendo todos eles medidos em 9 países.

use @ **R!**