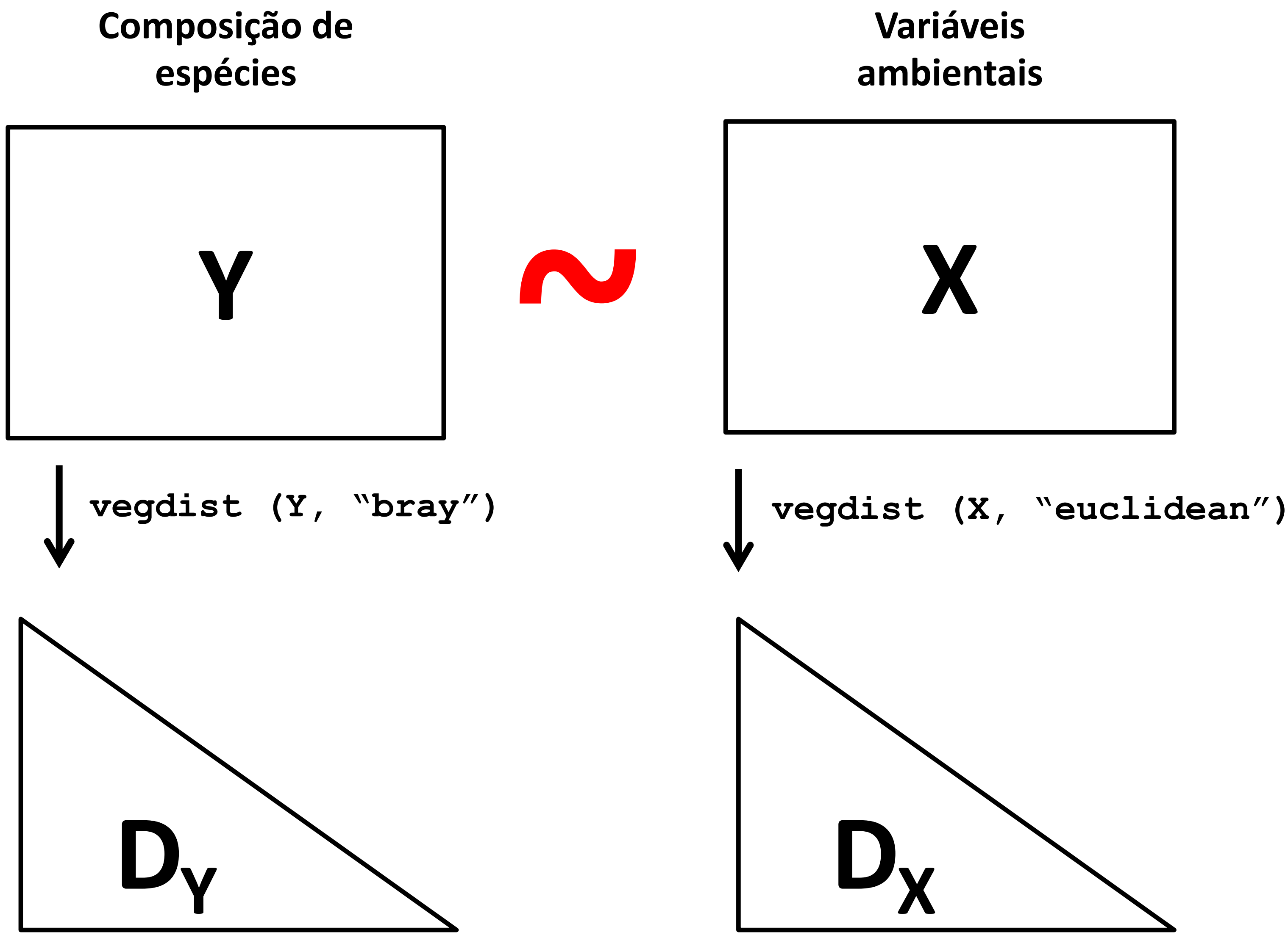

EU NÃO
AGUENTO
MAIS!

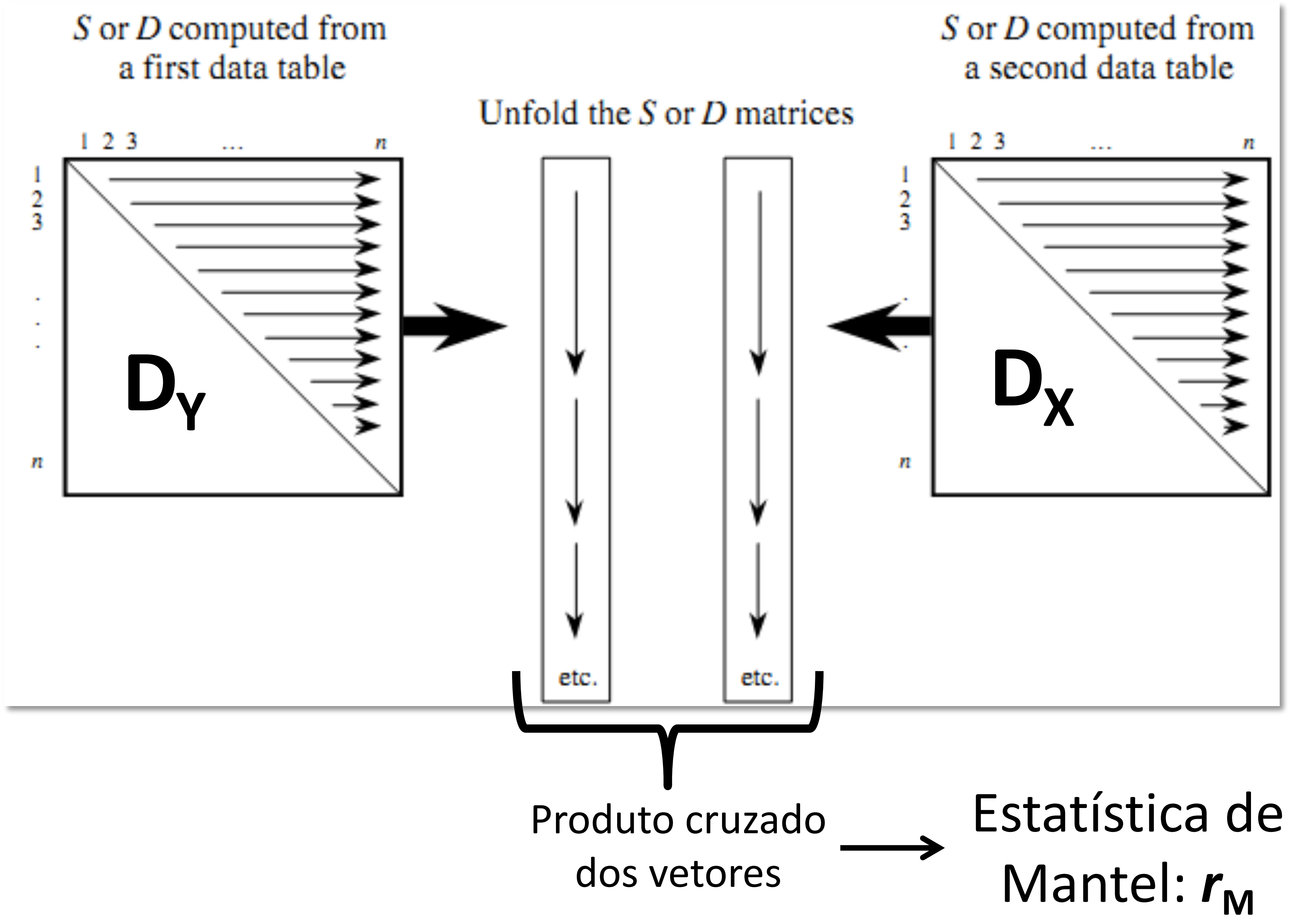
Aula 6: Métodos para testar relações entre variáveis e diferença entre grupos

p. 597-612 do Cap. 10 Legendre & Legendre

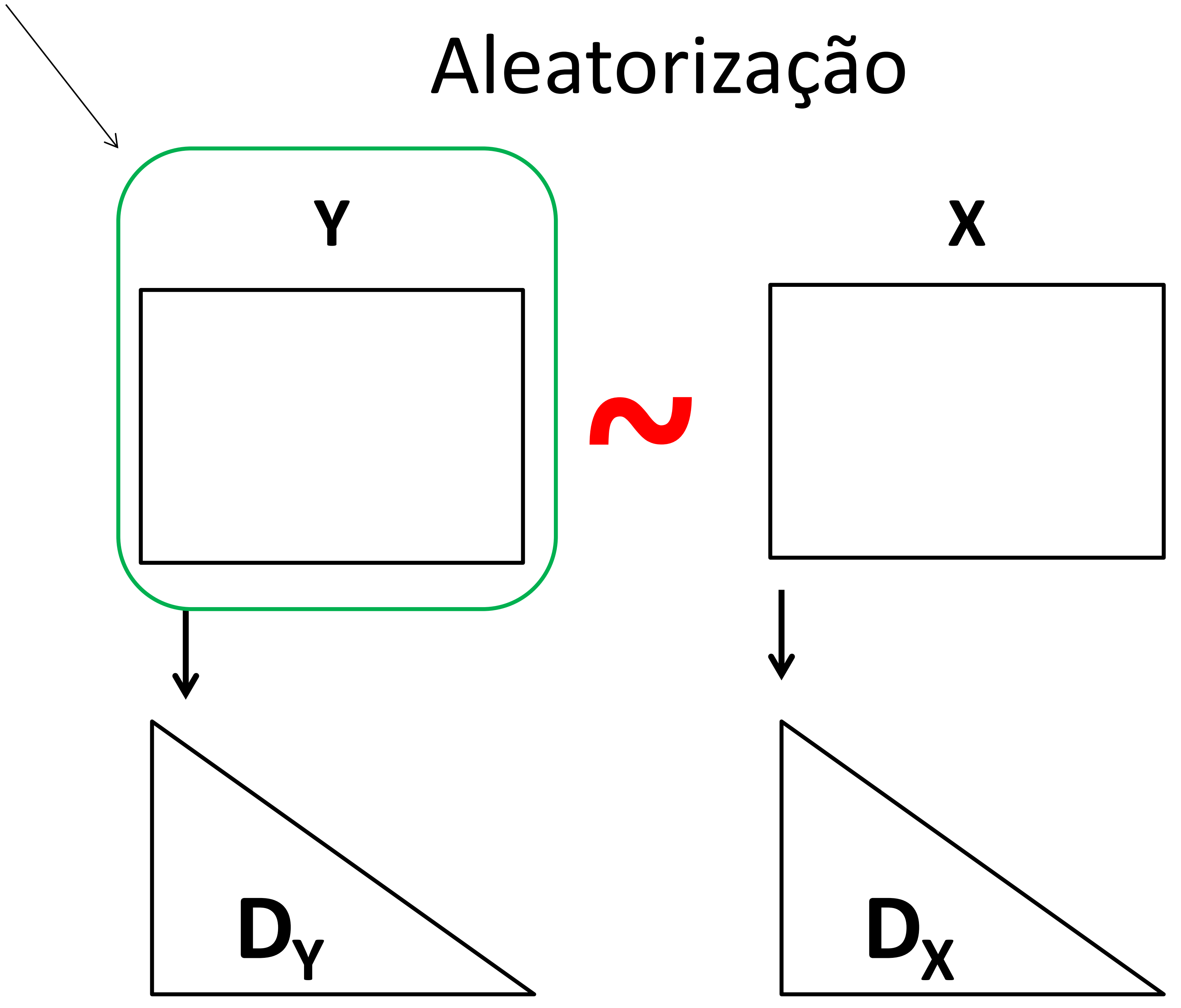
- Mantel
- Mantel parcial
- RMR
- Procrustes/Protest
- PERMANOVA
- betadisper+permdisp

Teste de Mantel




S or D computed from a first data table
S or D computed from a second data table
Unfold the S or D matrices
1 2 3 … n
1 2 3 . . . n
D_Y
D_X
etc.
etc.
Produto cruzado dos vetores
Estatística de Mantel: r_M

Aleatorização



Mantel parcial

- Relaciona dois conjuntos de matrizes de distância, controlando para uma terceira
- Lógica similar à da db-RDA, só que usando matrizes de distância
- A ordem das matrizes importa
- Útil para testar relação entre distância genética e distância ambiental, controlando para distância geográfica

Vantagens

Possibilidade de usar qualquer coeficiente de distâncias

Qualquer tipo de dado pode ser usado para construir a matriz de distância, inclusive mistura de dados

Permite usar mais de uma matriz (testes parciais)

Desvantagens

- Ter de transformar os dados brutos (matriz quadrada) em distância, com perda inerente de informação
- Interpretação é dada no espaço de distâncias
- Perda da informação das observações originais
- Baixo poder (Erro Tipo II) e alto Erro Tipo I

3. Two types of R^2 statistics and SS

The $R_M{}^2$ statistic of the Mantel test differs from the R^2 of correlation, regression, or canonical analysis. Demonstration:

R^2 of multiple regression and RDA : $$R^2 = \frac{SS(\widehat{\mathbf{Y}})}{SS(\mathbf{Y})}$$

Its denominator is $$SS(\mathbf{Y}) = \sum_{i=1}^{n} \sum_{j=1}^{p} (y_{ij} - \bar{y}_j)^2 = \frac{1}{n} \sum_{h=1}^{n-1} \sum_{i=h+1}^{n} D_{hi}^2 \quad *$$

The $R_M{}^2$ of the Mantel test is the square of the Mantel correlation (r_M).

Its denominator is SS(**D**), not SS(**Y**):

$$SS(\mathbf{D}) = \sum_{h=1}^{n-1} \sum_{i=h+1}^{n} (D_{hi} - \bar{D})^2 = \sum_{h=1}^{n-1} \sum_{i=h+1}^{n} D_{hi}^2 - \frac{\left(\sum\sum D_{hi}\right)^2}{n(n-1)/2}$$

number of distances

These two statistics are irreducible to one another.

Slide P. Legendre

4. Two types of SS – An ecological example

Consider a matrix of species presence-absence at four sites. There is variation among the sites (beta diversity).

Data **Y**

	Sp.1	Sp.2	Sp.3	Sp.4	Sp.5
Site 1	1	1	0	0	0
Site 2	1	0	1	0	0
Site 3	1	0	0	1	0
Site 4	1	0	0	0	1

There's variation in sp composition between sites

SS(**Y**) of the data = 3.000

SS(from the Jaccard D) =

$$6^*(0.667^2)/4 = 0.667$$

D = [1 – Jaccard similarity]

	Site 1	Site 2	Site 3	Site 4
Site 1	0	0.667	0.667	0.667
Site 2	0.667	0	0.667	0.667
Site 3	0.667	0.667	0	0.667
Site 4	0.667	0.667	0.667	0

The numbers are equal because they only differ in one species

SS(**D**) in the upper triangle = 0.000

Mantel would find there's no variation in the data

The variation of the distances in matrix **D** does not provide any information about the variation in the **Y** data.

Slide P. Legendre

5. Basic assumptions of the Mantel test: Do they hold in spatial analysis?

1. Mantel correlation: hypothesis of linearity (or at least of monotonic relationship) of the point dispersion in the **D–D** graph.

2. When D_1 is small, D_2 is small; when D_1 is large, D_2 is large. The distribution of values in the **D–D** graph is homoscedastic.

Are these assumptions verified in data on which ecologists and geneticists carry out spatial analysis?

⇒ Let us examine spatially autocorrelated (SA) simulated data.

⇒ Avantage of simulated data: we know for sure the range of the spatial autocorrelation.

Slide P. Legendre

Mantel serve para análises espaciais?

J. Statist. Comput. Simul., 2000, Vol. 67, pp. 37 – 73

COMPARISON OF PERMUTATION METHODS FOR THE PARTIAL CORRELATION AND PARTIAL MANTEL TESTS

PIERRE LEGENDRE*

CONCEPTS & SYNTHESIS
EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

Ecological Monographs, 75(4), 2005, pp. 435–450

ANALYZING BETA DIVERSITY: PARTITIONING THE SPATIAL VARIATION OF COMMUNITY COMPOSITION DATA

Pierre Legendre,[1] Daniel Borcard, and Pedro R. Peres-Neto[2]

MOLECULAR ECOLOGY RESOURCES
Molecular Ecology Resources (2010) **10**, 831–844 doi: 10.1111/j.1755-0998.2010.02866.x

METHODOLOGICAL ADVANCES - INFERENCE OF SPATIAL STRUCTURE

Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data

PIERRE LEGENDRE* and MARIE-JOSÉE FORTIN†

Methods in Ecology and Evolution
British Ecological Society
Methods in Ecology and Evolution 2015, **6,** 1239–1247 doi: 10.1111/2041-210X.12425

Should the Mantel test be used in spatial analysis?

Pierre Legendre[1]*, Marie-Josée Fortin[2] and Daniel Borcard[1]

- Like the previous results, these new simulations show that the power of the Mantel test is always lower than that of canonical analysis (RDA).

- Our results also show that spatial variation is, at best, weakly captured by direct regression of a response distance (or dissimilarity) matrix on a geographic distance matrix.

- None of the transformations of the distances that we tried increased the performance of the Mantel test.

- On the RDA side, representation of the spatial relationships among sites by Moran's eigenvector maps (MEM) produces a much more powerful test than the use of simple geographic coordinates or in the form of a polynomial.

Slide P. Legendre

9. Conclusion

Researchers like the Mantel test because it is simple to use. One simply has to type: mantel(D1, D2) to obtain a p-value.

1. The Mantel test is not appropriate to test for the presence of spatial structures in raw survey data, for several reasons.

1.1. The null hypothesis (H_0) of the test of correlation between vectors of raw data is different from H_0 of a test involving dissimilarities.

1.2. The statistics used in these two tests are different and cannot be reduced to one another.

1.3. The Mantel correlation assumes that, in a **D–D** plot, the relationship among points is linear and homoscedastic. That is not the case for spatially-structured data, except when the range of the spatial correlation extends to the whole surface under study or beyond.

Slide P. Legendre

9. Conclusion

2. If one still applies the Mantel test to spatially-structured data, its power is always lower than that of RDA using Moran's eigenvector maps (MEM) to represent spatial relationships among sites.

Our simulations showed that the Mantel R^2 is always much lower than the R^2 produced by MEM analysis. It should not be interpreted as the proportion of the variance of **Y** explained by **X**.

=> Analysis by spatial eigenvectors (MEM, AEM) produces output richer than a p-value, which is the only output of a Mantel test. In particular, the ordination triplots display the relationships among sites, species, and MEM eigenvectors. They also allow one to produce maps of the fitted values at different spatial scales.

3. The Mantel test is inappropriate to test the correlation between raw data vectors or matrices, irrespective of the fact that these vectors are spatially-structured or not, because the Mantel statistic is inappropriate to answer that type of question; the test would have low power. See simulations in Legendre et Fortin (2010).

Slide P. Legendre

9. Conclusion

4. Our main conclusion is that the Mantel test should only be used to answer questions that, in the application field, clearly and solely concern *relationships between distances*. Such questions are rarely found in ecology and genetics.

It should not be used to answer questions derived from questions that originally concerned the raw data, from which dissimilarities have been computed.

5. Who wants to use a test that has low power (low capacity to detect an effect), instead of a test that has high power?

Received: 20 December 2017 | Accepted: 7 December 2018

DOI: 10.1111/2041-210X.13141

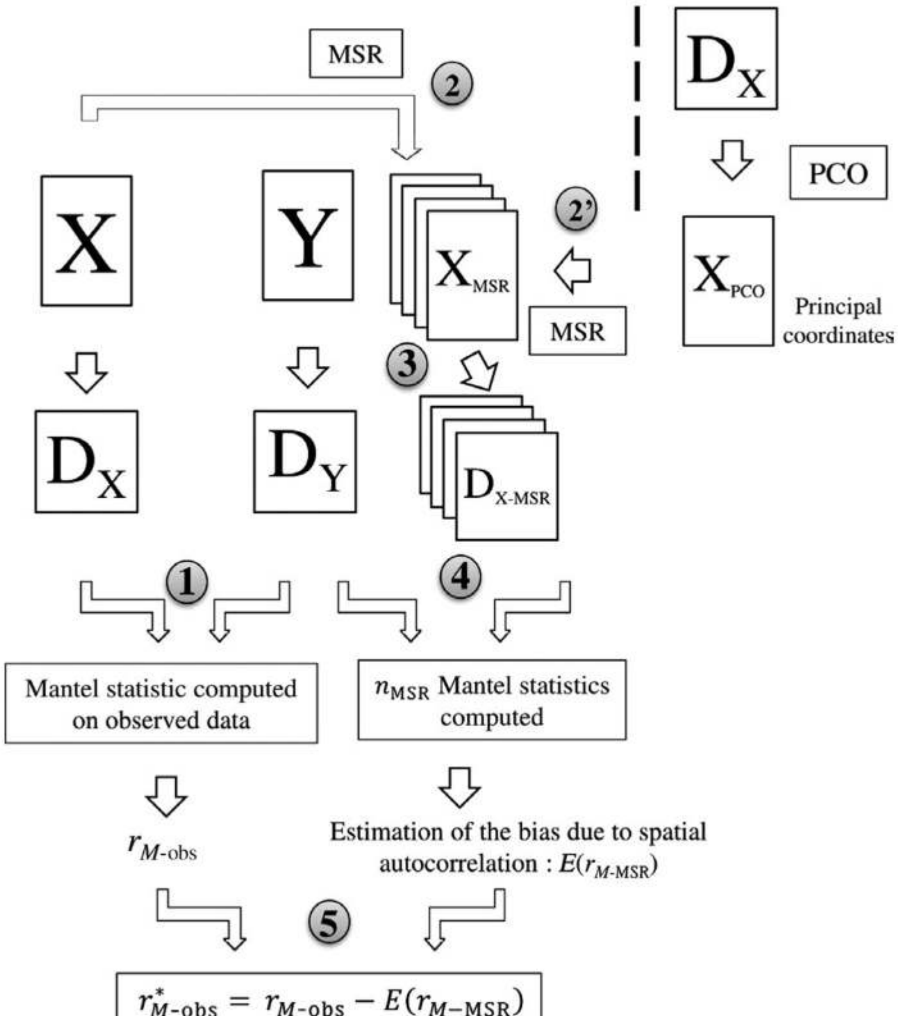
RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Testing the Mantel statistic with a spatially-constrained permutation procedure

Julie Crabot[1*] | Sylvie Clappe[2*] | Stéphane Dray[2] | Thibault Datry[1]

Methods Ecol Evol. 2019;10:532–540.



Usa um novo procedimento para gerar modelos nulos espacialmente restritos

Regressão Múltipla de Matrizes de distância

O que é e como funciona?

- Muito similar ao Mantel, mas ao invés de calcular correlação de Pearson, calcula regressão
- Também permite várias matrizes preditoras
- Portanto requer variável preditora e resposta
- Fornece R^2

Lichstein 2007

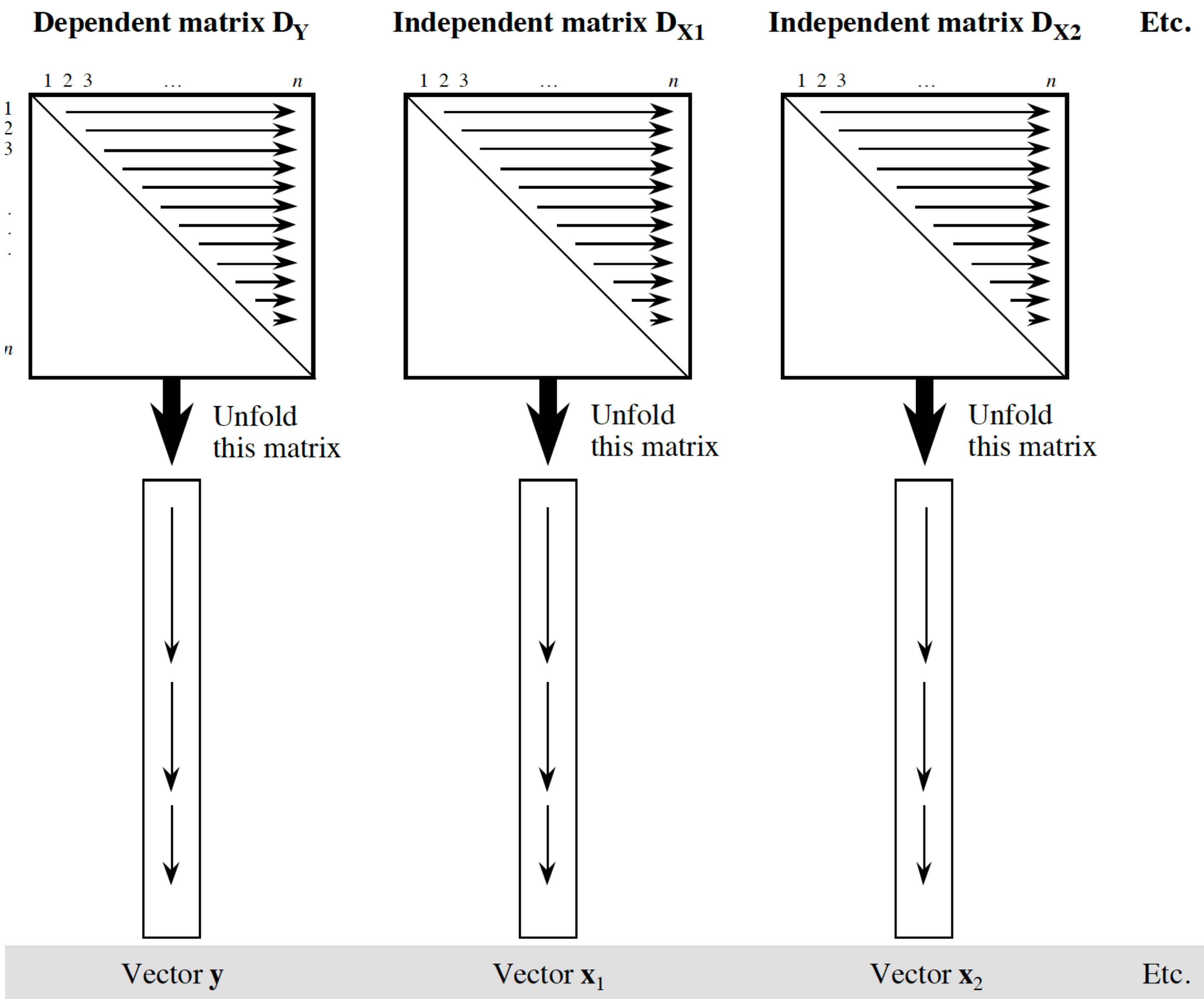

Dependent matrix D_Y
Independent matrix D_{X1}
Independent matrix D_{X2}
Etc.
1 2 3 … n
1 2 3 … n
Unfold this matrix
Vector **y**
Vector $\mathbf{x}_1$
Vector $\mathbf{x}_2$
Etc.

Fig. 10.22 L&L

Generalized Dissimilarity Modelling (GDM)

Syst. Biol. 51(2):331–363, 2002

Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning: Where to from Here?

SIMON FERRIER

Diversity and Distributions, (Diversity Distrib.) (2007) **13**, 252–264

BIODIVERSITY RESEARCH

Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment

Simon Ferrier[1]*, Glenn Manion[1], Jane Elith[2] and Karen Richardson[3]

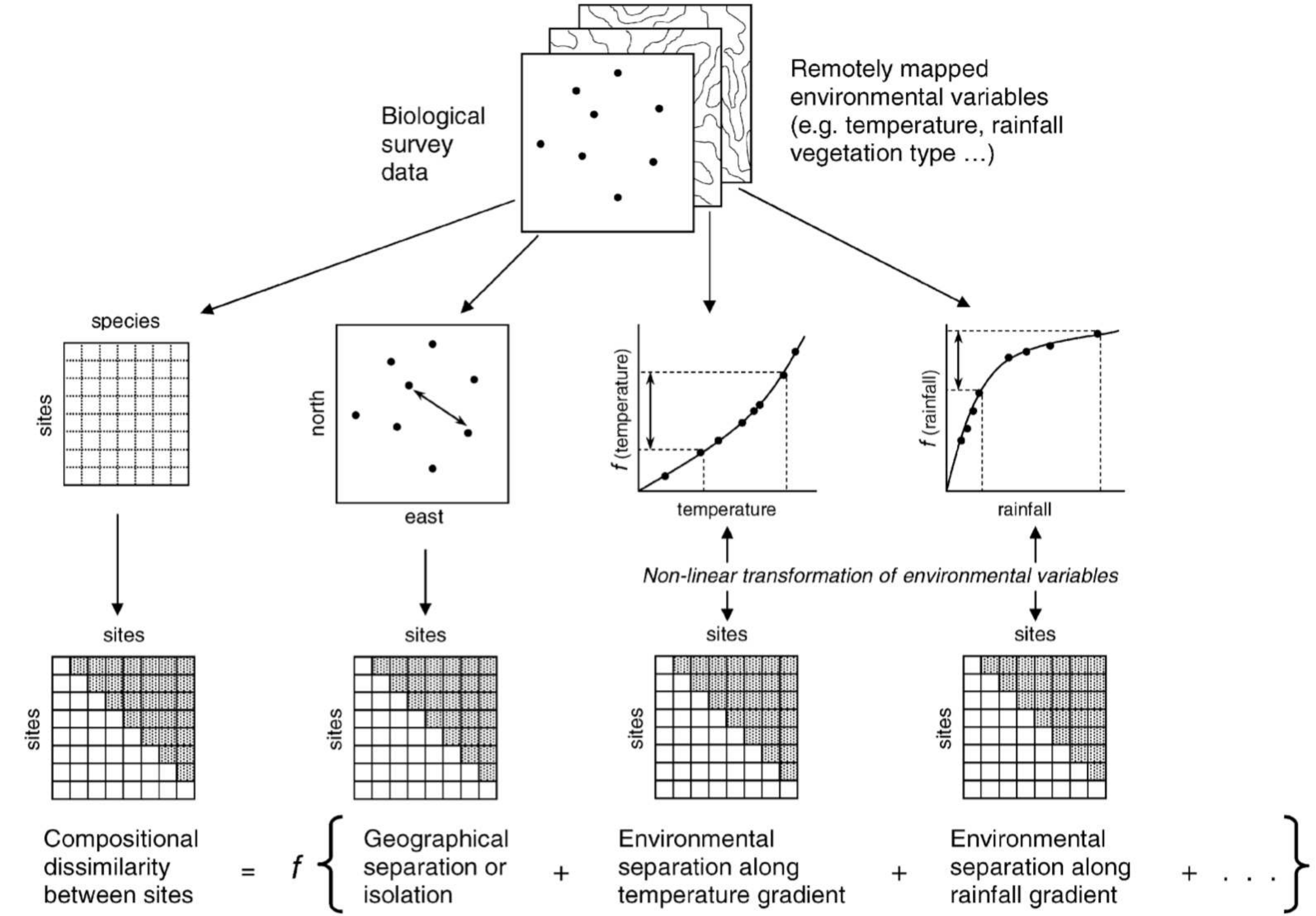


Figure 8. A diagrammatic illustration of the GDM approach to modelling compositional dissimilarity

Generalized Dissimilarity Modelling

- Extensão da MRM
- Acomoda relações não-lineares
- Bom desempenho em modelar matrizes de distância com saturação
- Aplicações em modelagem de comunidades, mapeamento

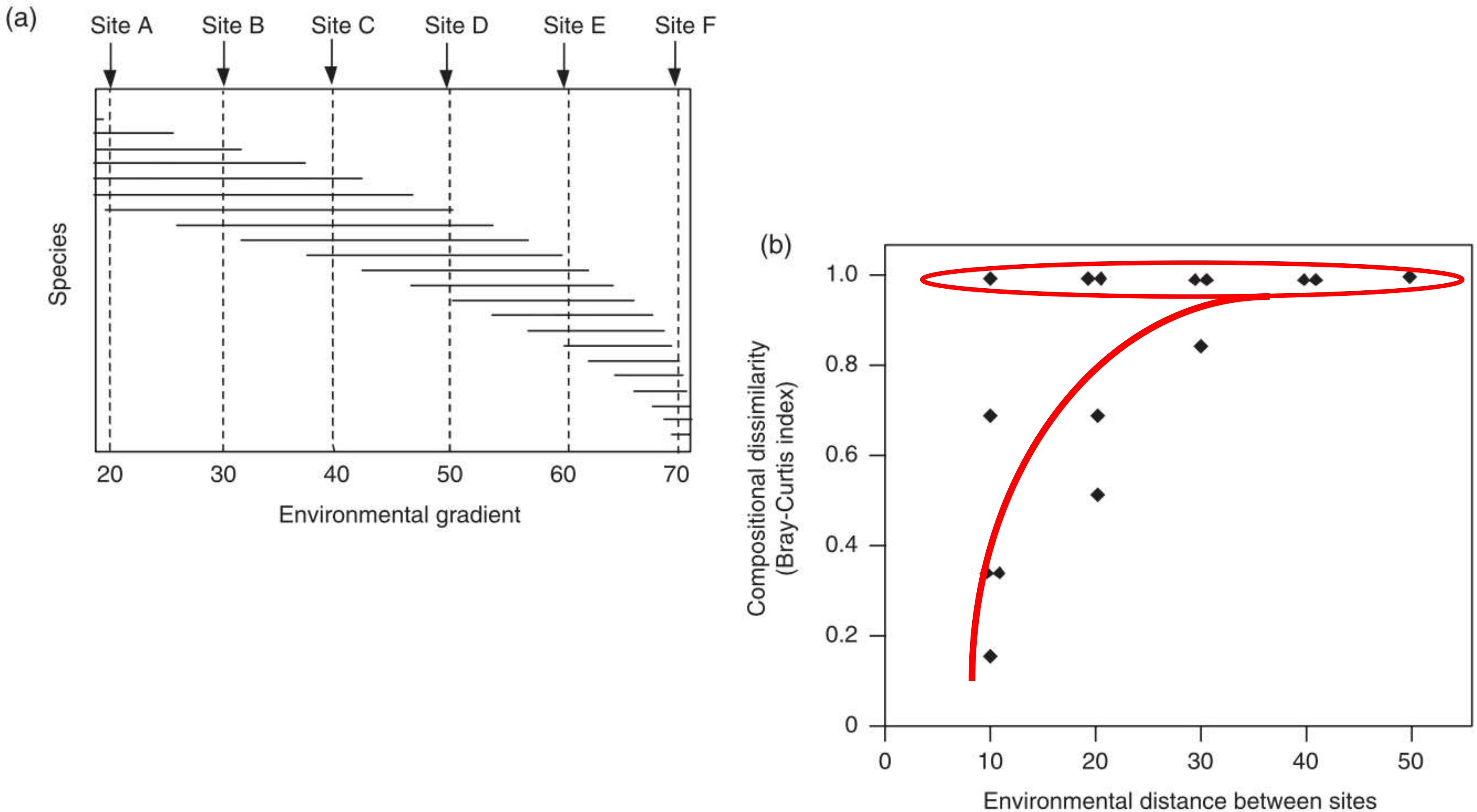


Figure 1 Hypothetical example illustrating the problems discussed in the text relating to the application of linear matrix regression to large-scaled ecological data sets. (a) Species' distributions (horizontal lines) in relation to a hypothetical environmental gradient, with six survey sites positioned along this gradient. (b) Plot of compositional dissimilarity vs. environmental distance for all possible pairwise combinations of the six survey sites.

Ferrier et al. 2007

Análise PROcrustes

In Greek mythology, **Procrustes was a son of Poseidon** and a **rogue**. He invited travellers to spend the night with him, then tied them down to an iron bed and either cut off their limbs if they were taller than the bed, or stretched the victims if they were too short, till they fitted in.

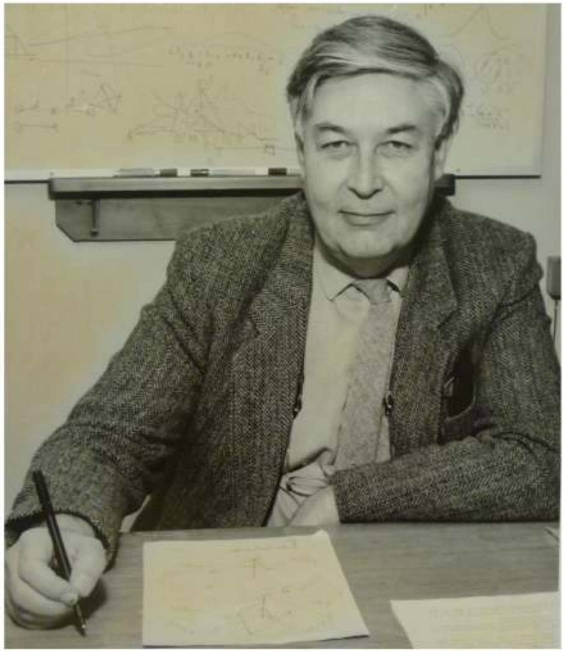
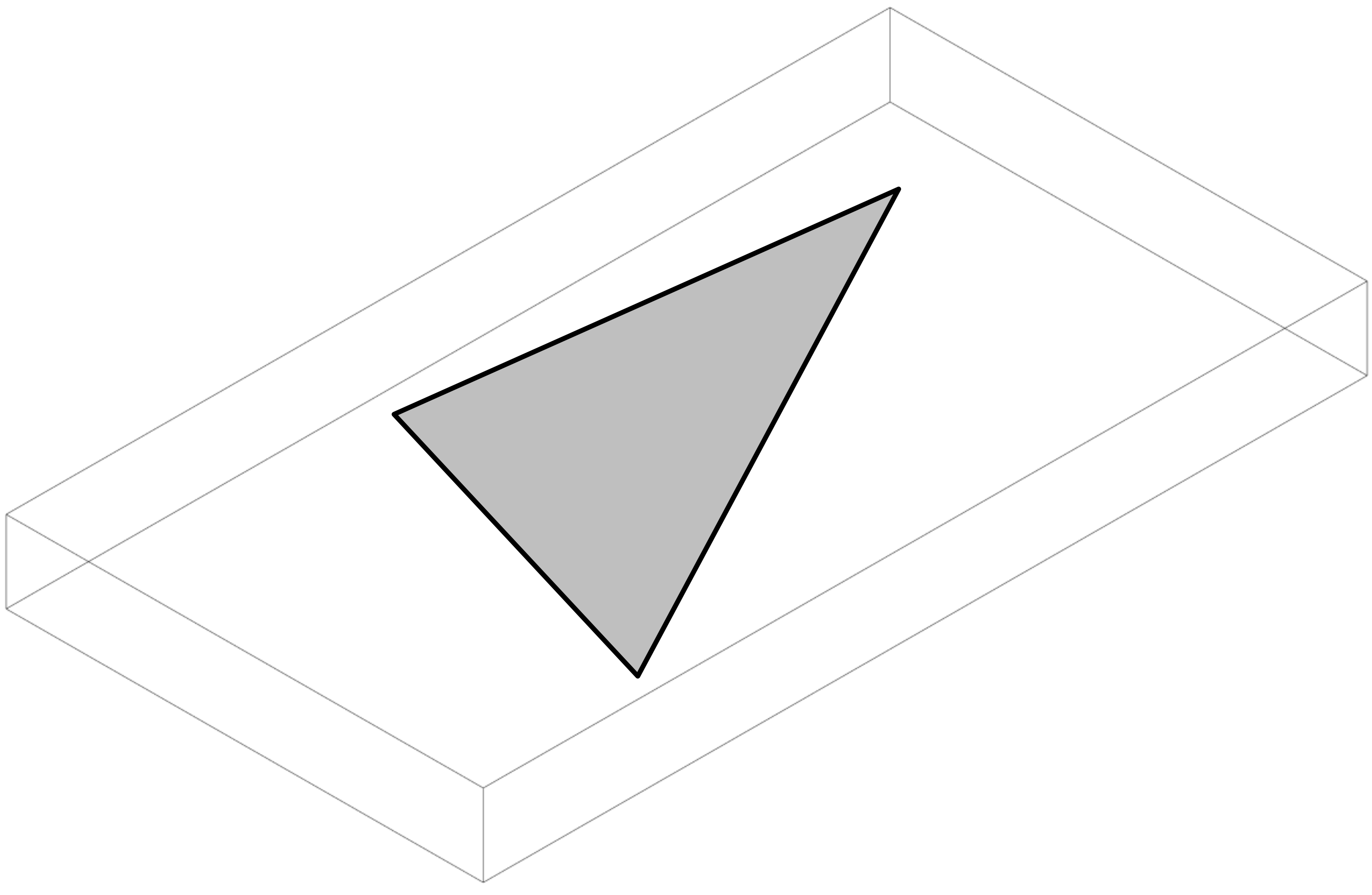
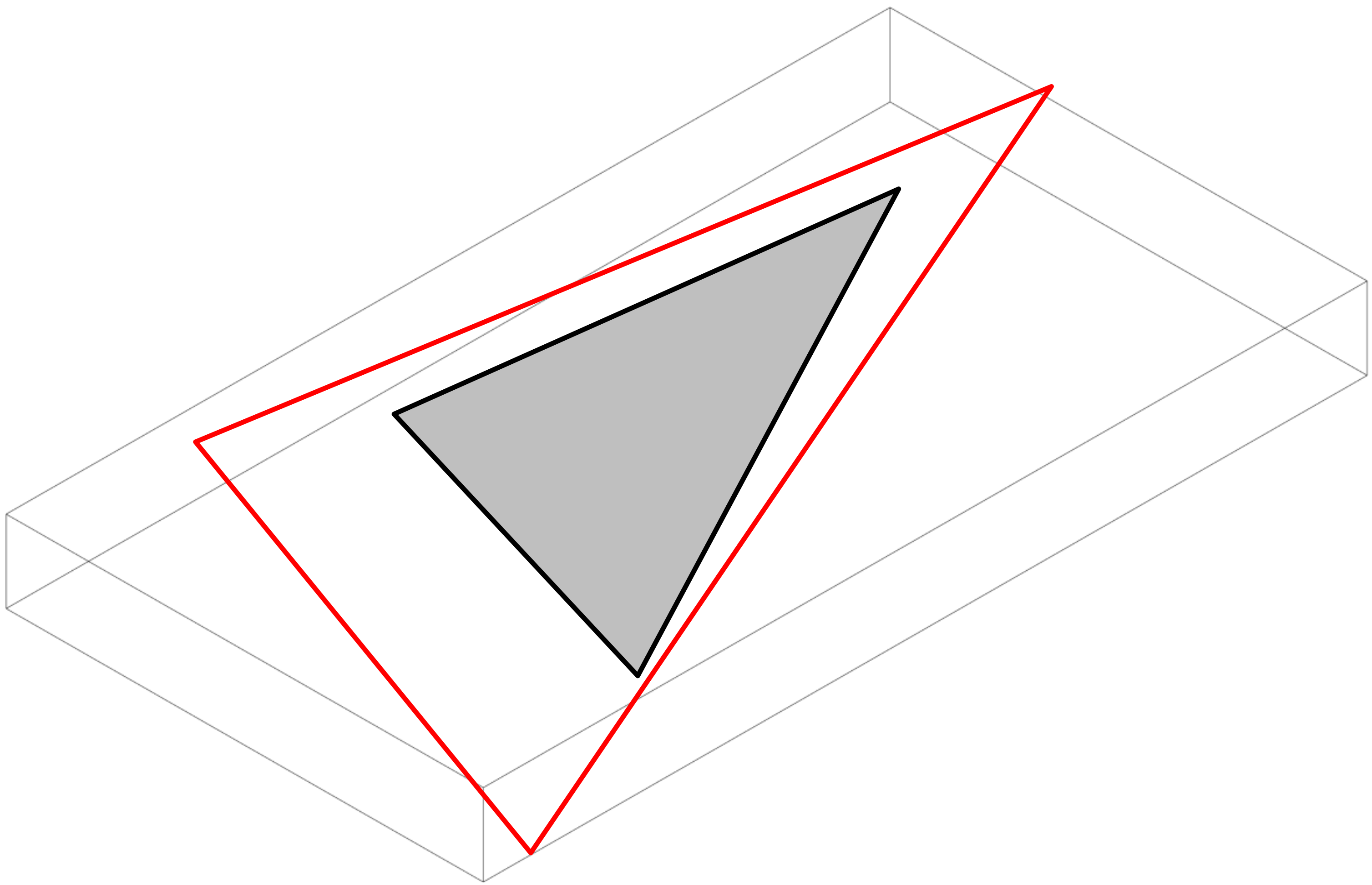


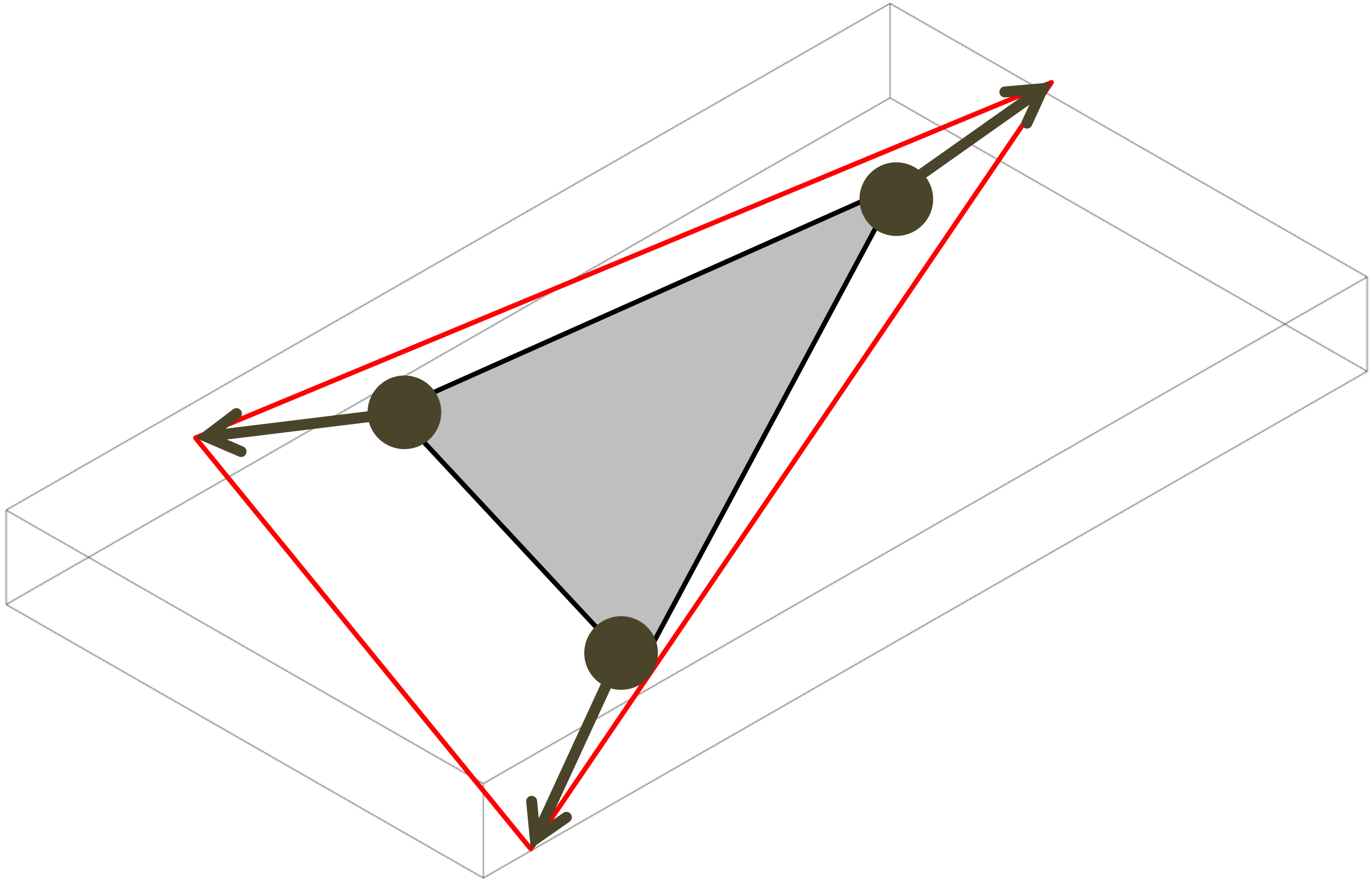
Figure 1. *John Gower, circa 1985.*

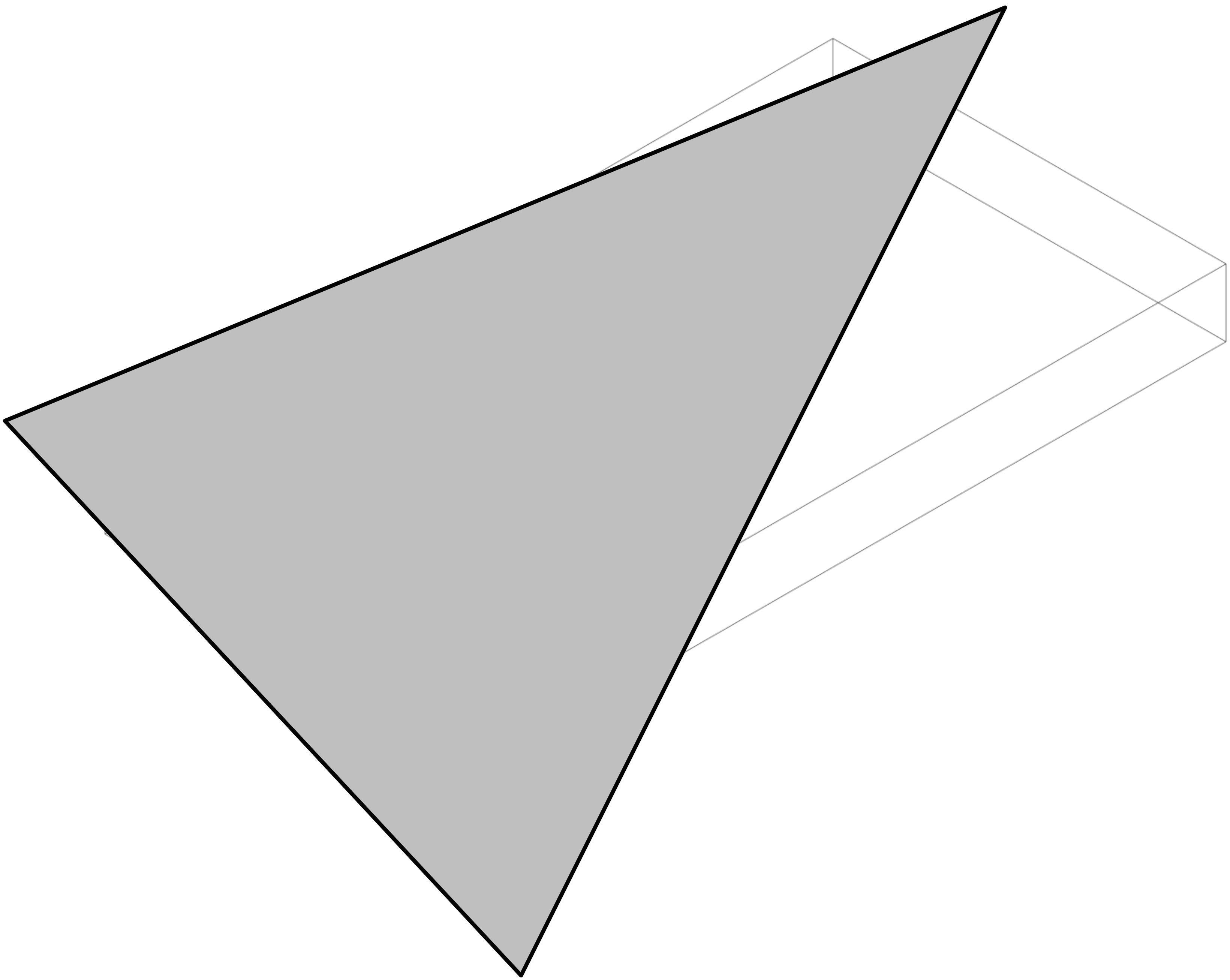
Tradução de procrusto (em português) = "O esticador"

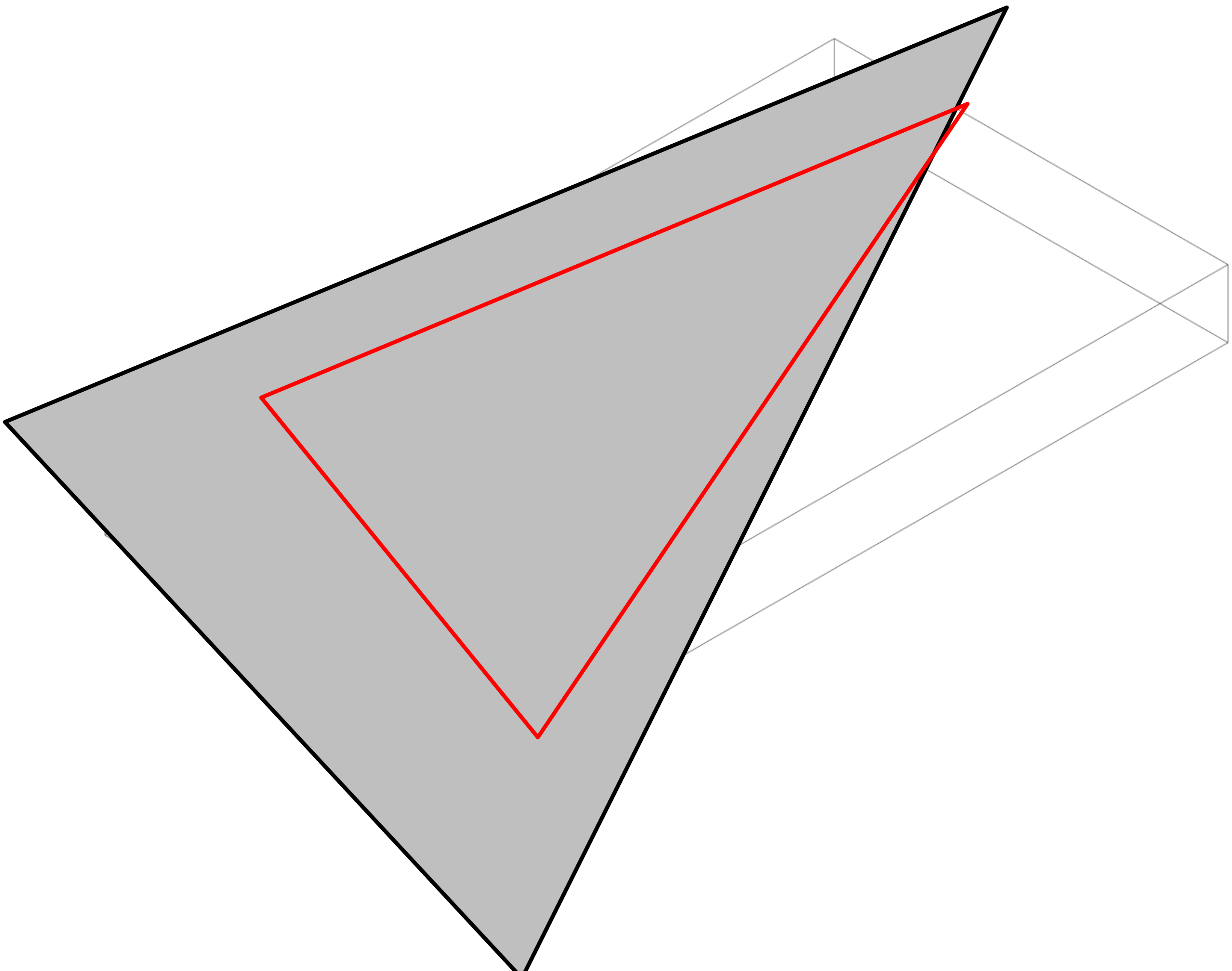
Procrustes e Teseu

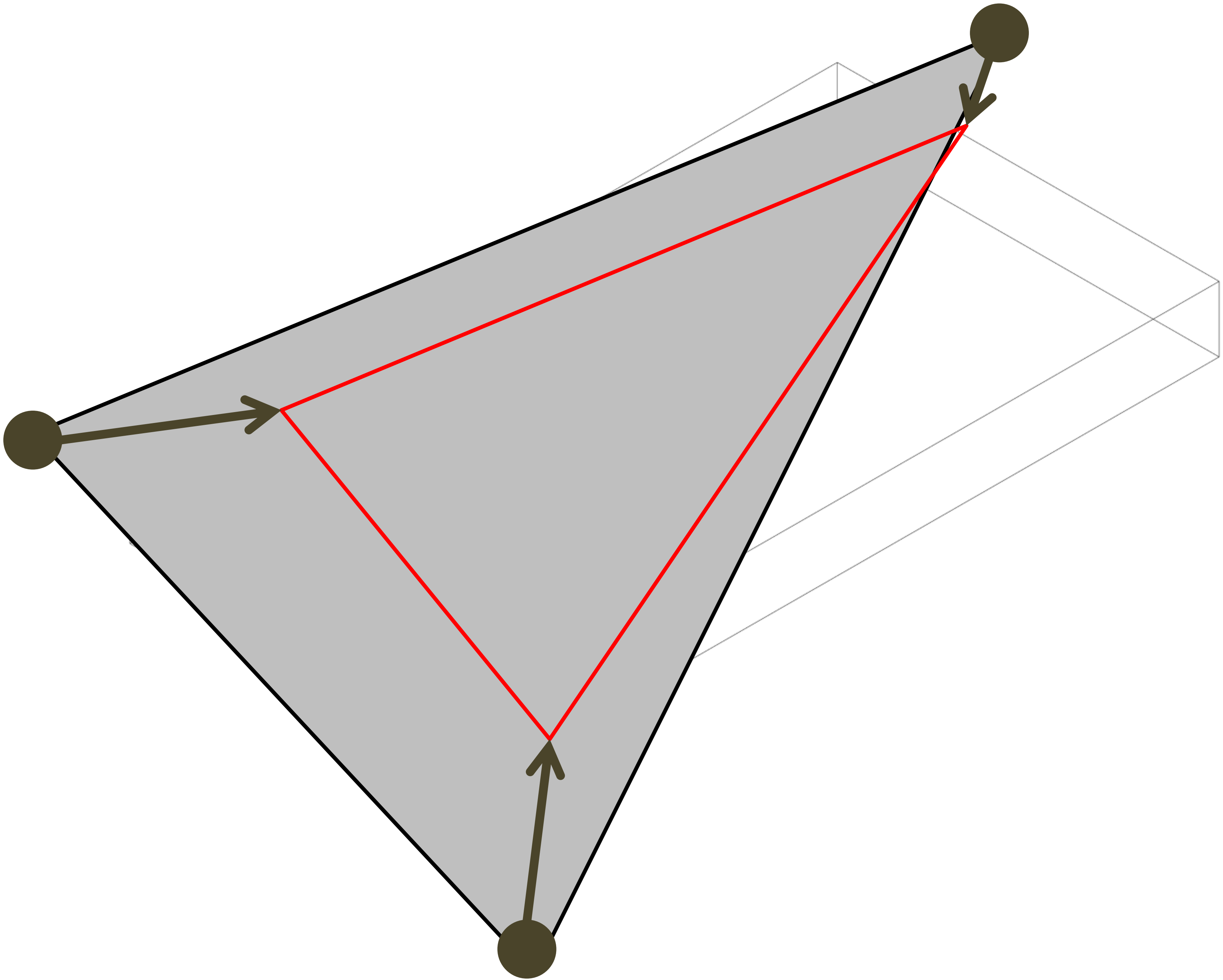






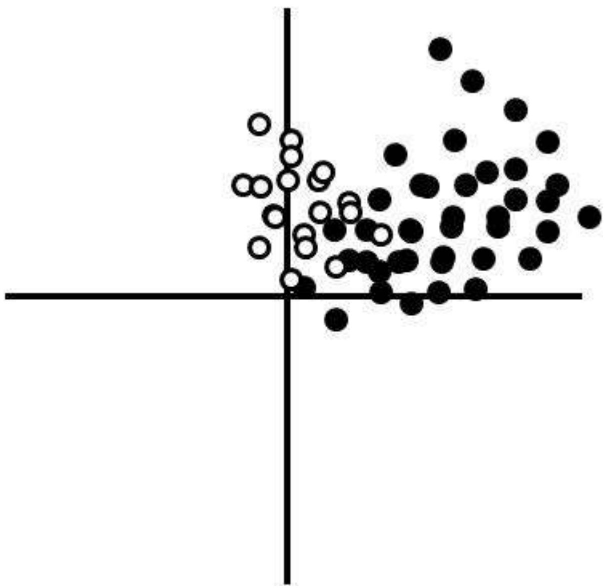




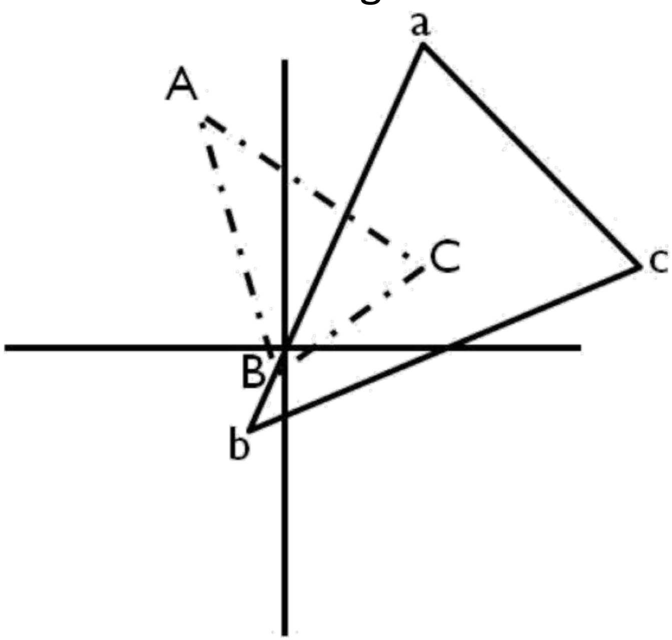


Análise PROcrustes

Dados originais



Dados originais



Utiliza um algoritmo que minimiza a soma de quadrados da distância entre dois pontos correspondentes em duas matrizes

If $\mathbf{Y}$ is a matrix of Cartesian coordinates:

$$\mathbf{Z} = \left[tr \left[\left(\mathbf{Y} - \bar{\mathbf{Y}} \right) \left(\mathbf{Y} - \bar{\mathbf{Y}} \right)^{\mathbf{T}} \right] \right]^{-1/2} \left(\mathbf{Y} - \bar{\mathbf{Y}} \right) \mathbf{H}$$

Procrustes Residuals

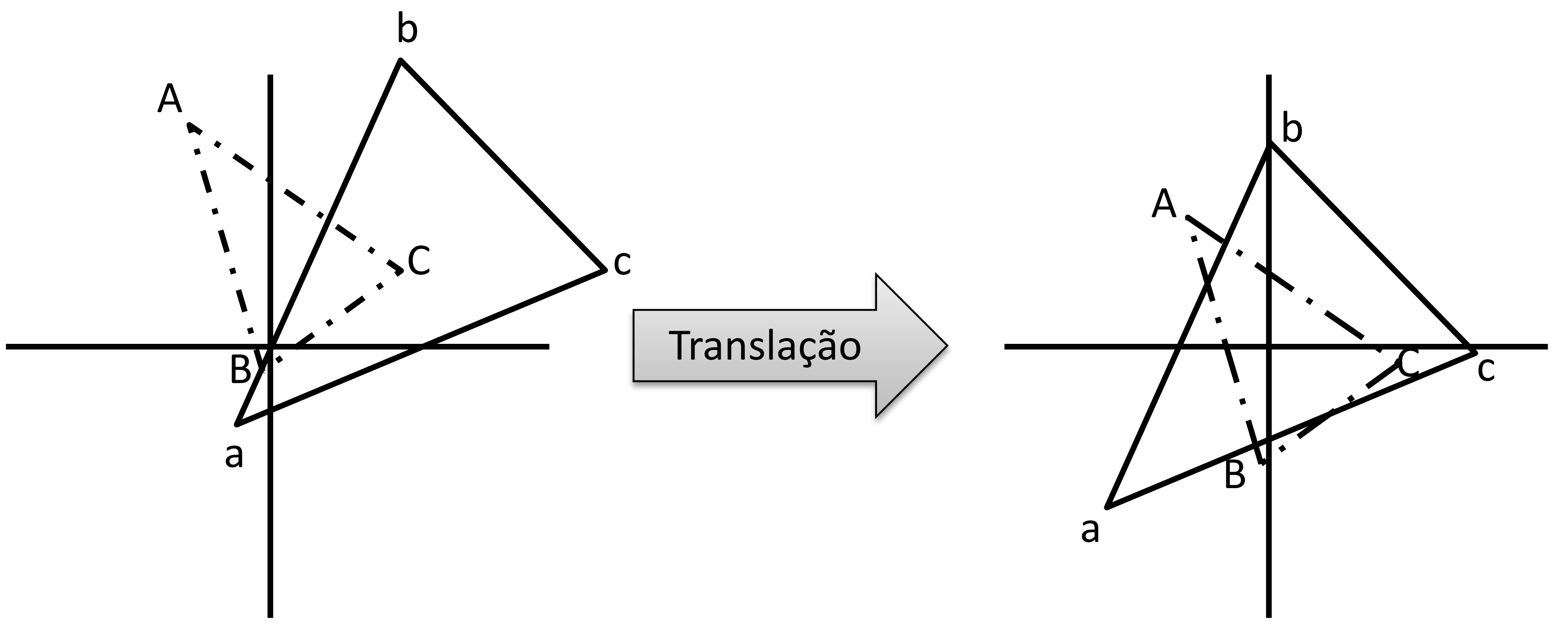
Inverse of Centroid Size Rescaling

Centering

Rotation

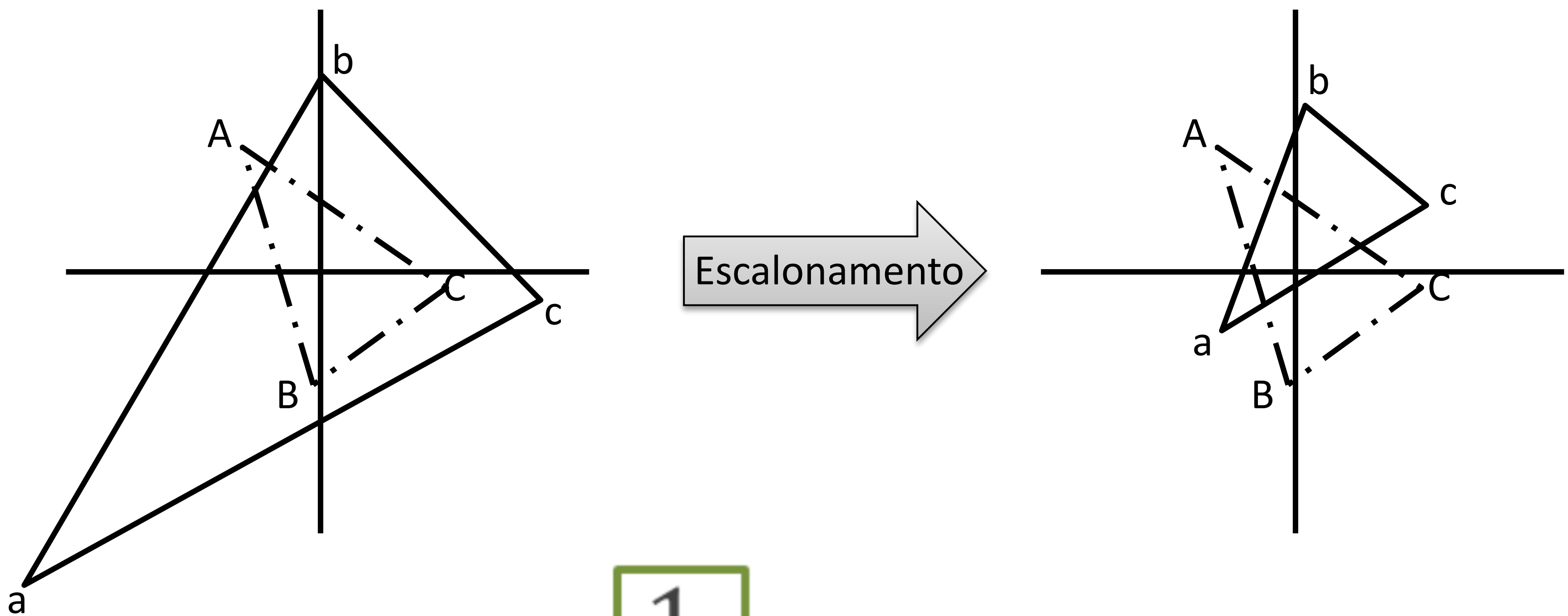
Slide M. Collyer

PROcrustes



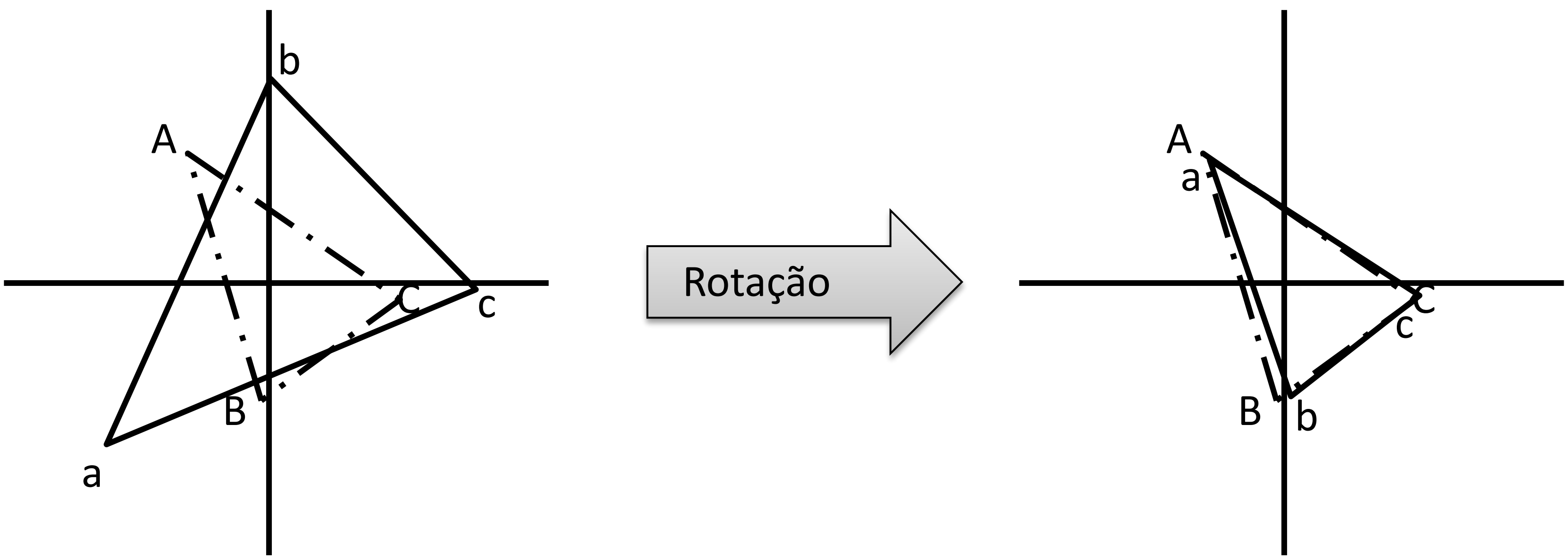
$$\mathbf{Z} = \frac{1}{cs}(\mathbf{Y} - \bar{\mathbf{Y}})\mathbf{H}$$

PROcrustes



$$\mathbf{Z} = \frac{1}{cs}(\mathbf{Y} - \bar{\mathbf{Y}})\mathbf{H}$$

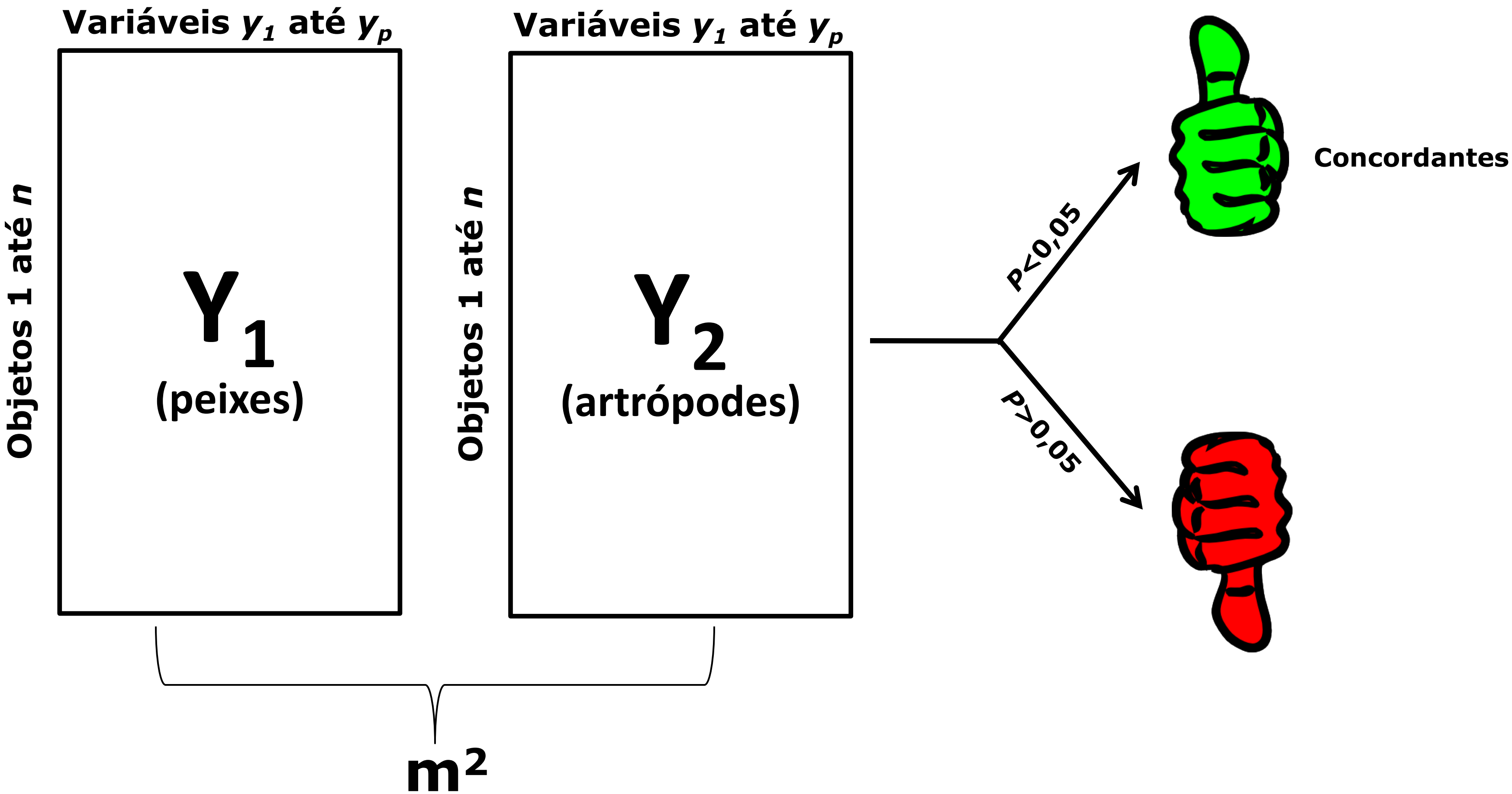
PROcrustes



$$\mathbf{Z} = \frac{1}{\mathrm{cs}}(\mathbf{Y} - \bar{\mathbf{Y}})\boxed{\mathbf{H}}$$

Análise PROcrustes

Mede o **grau de concordância** entre duas matrizes



Estatística do método

$$m_{12}^2 = \text{Trace}\left(\mathbf{Y}_1\mathbf{Y}_1'\right) - \frac{\left(\text{Trace}\,\mathbf{W}\right)^2}{\text{Trace}\left(\mathbf{Y}_2\mathbf{Y}_2'\right)}$$

ÉCOSCIENCE 2 (3) : 297-303 (1995)

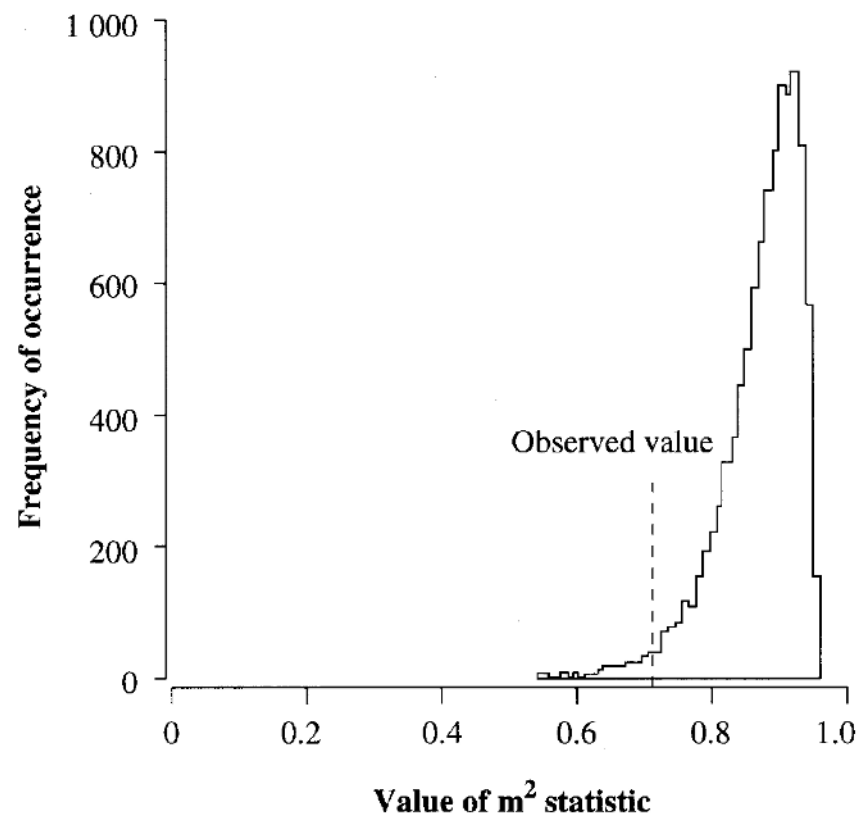
Teste de hipótese

PROTEST: A PROcrustean Randomization TEST of community environment concordance[1]

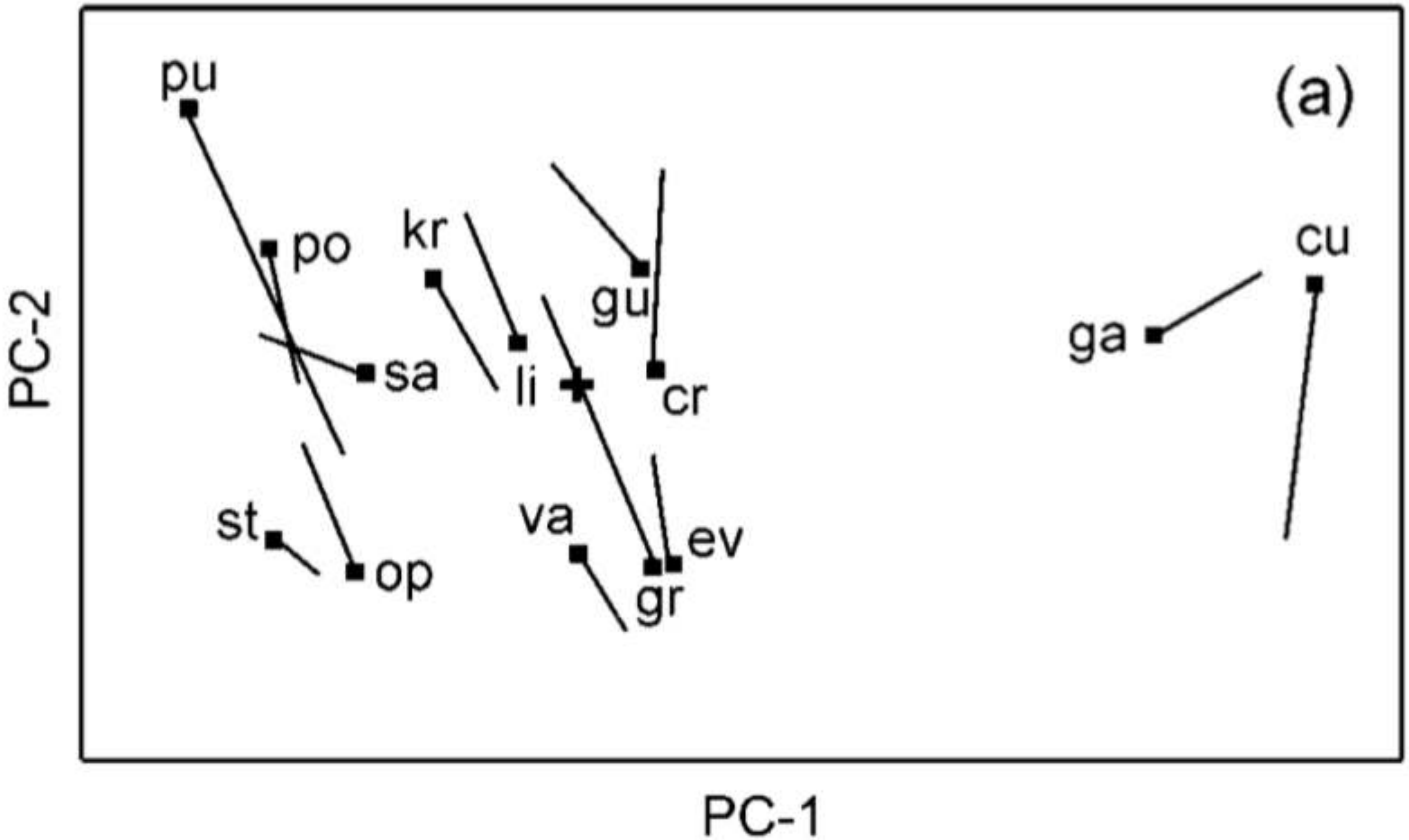
Donald A. JACKSON, Department of Zoology, University of Toronto, Toronto, Ontario M5S 1A1, Canada, e-mail: jackson@zoo.utoronto.ca

- PROTEST
  - Teste de hipótese para m^2
  - Baseado em procedimento de aleatorização

  de Monte Carlo



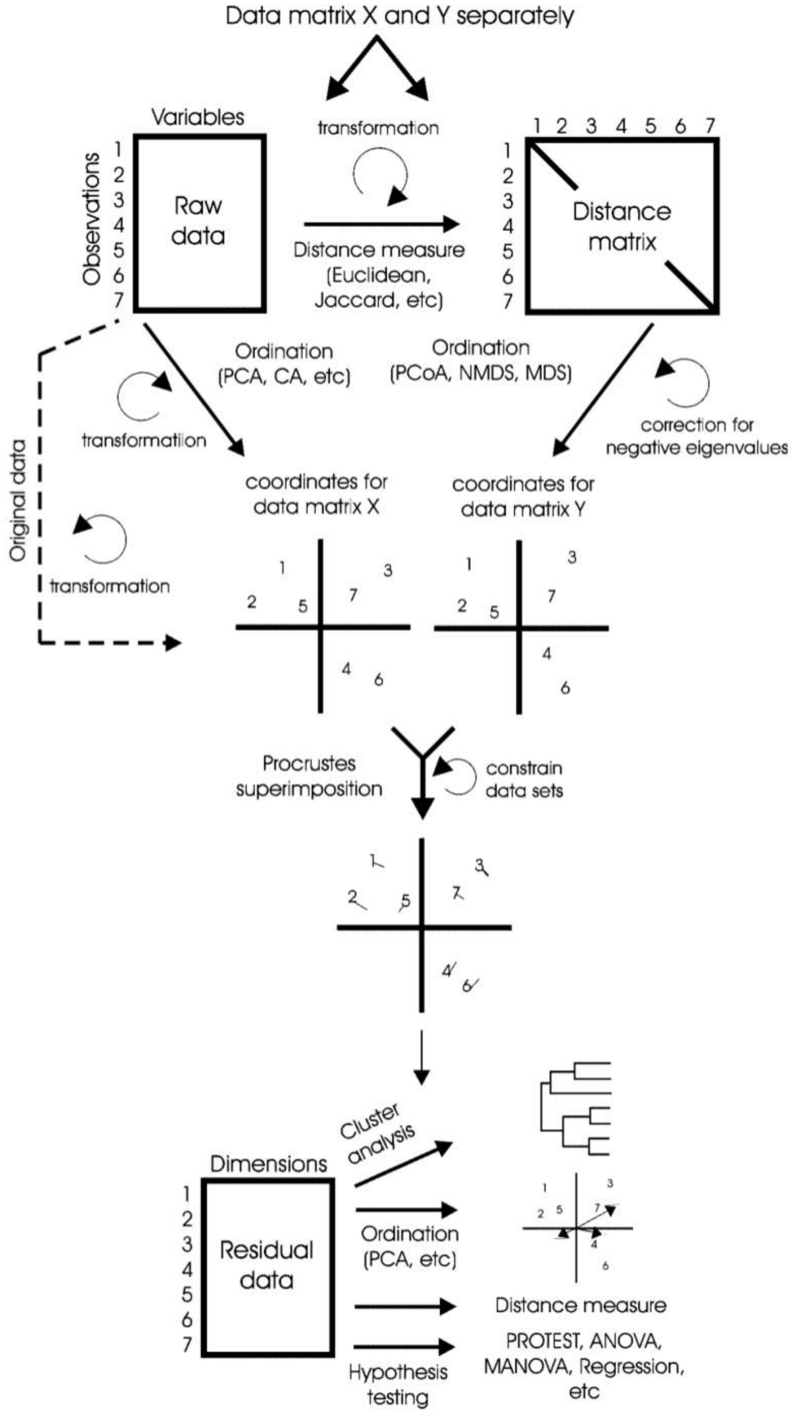
Exemplo de resultado gráfico



Vantagens do Procrustes/PROTEST

- Mais poderoso (baixo Erro Tipo I) do que Mantel para detectar padrões
- Não perde informações sobre observações individuais
  - Mantel reduz dados à distância, portanto perde informação sobre os objetos

Peres-Neto & Jackson 2001



Testar concordância entre conjunto de dados

Permite comparar ordenações

"Seed" agrupamentos

Peres-Neto & Jackson 2001

PERMANOVA

A test of the general multivariate hypothesis of differences in the composition and/or relative abundances of organisms of different species (variables) in samples from different groups or treatments.

Anderson (2001) Austral Ecology

PERMANOVA

A test of the general multivariate hypothesis of differences in the composition and/or relative abundances of organisms of different species (variables) in samples from different groups or treatments.

Aplicável para múltiplas variáveis resposta (não independentes) medidas em 1 ou mais fatores

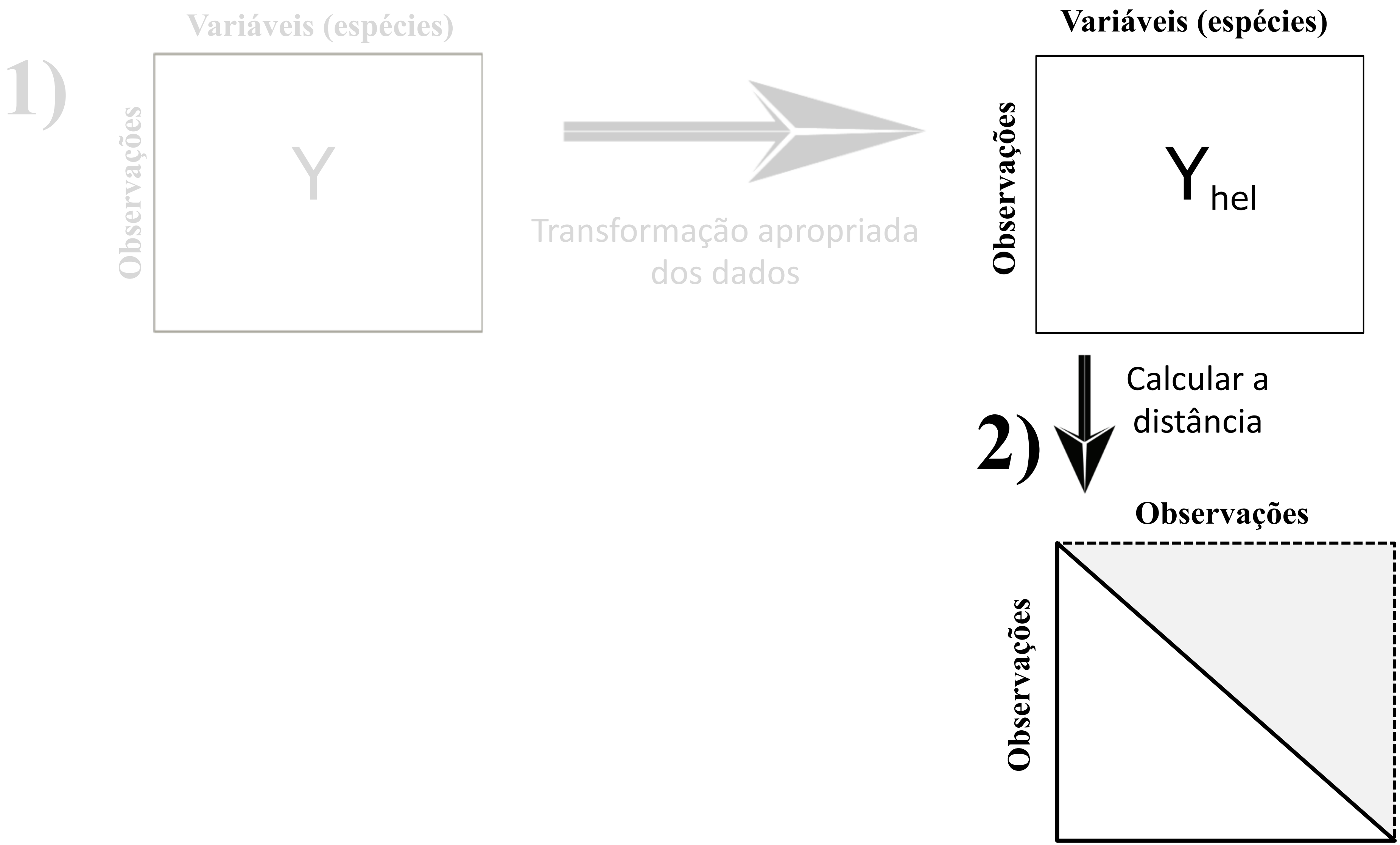
Anderson (2001) Austral Ecology

PERMANOVA

1)

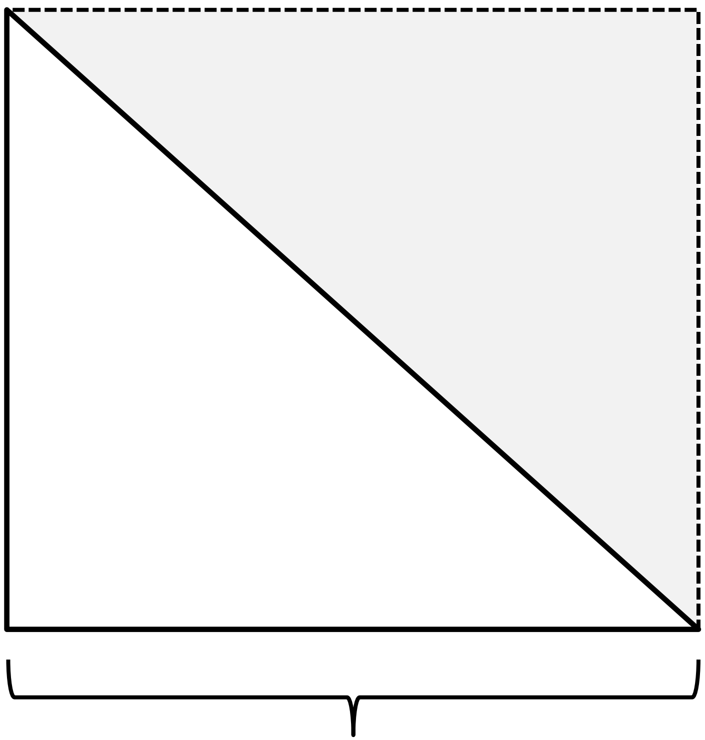

Variáveis (espécies)
Observações
Y
Transformação apropriada dos dados
Variáveis (espécies)
Observações
Y_{hel}

PERMANOVA


1)
Variáveis (espécies)
Observações
Y
Transformação apropriada dos dados
Variáveis (espécies)
Observações
Y hel
Calcular a distância
2)
Observações
Observações

PERMANOVA (simples – 1 fator)

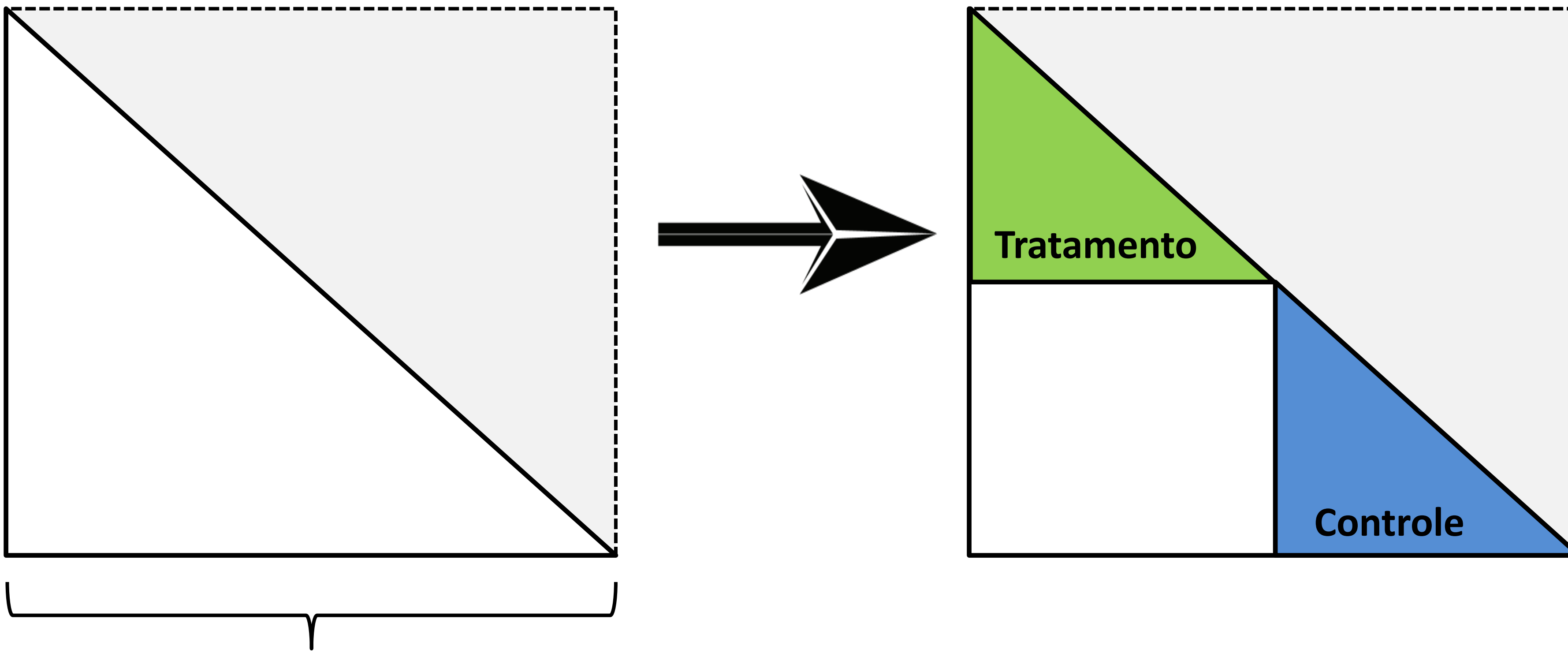
3)



$$SS_T = \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} {d_{ij}}^2\right] \Big/ N$$

PERMANOVA (simples – 1 fator)

3)

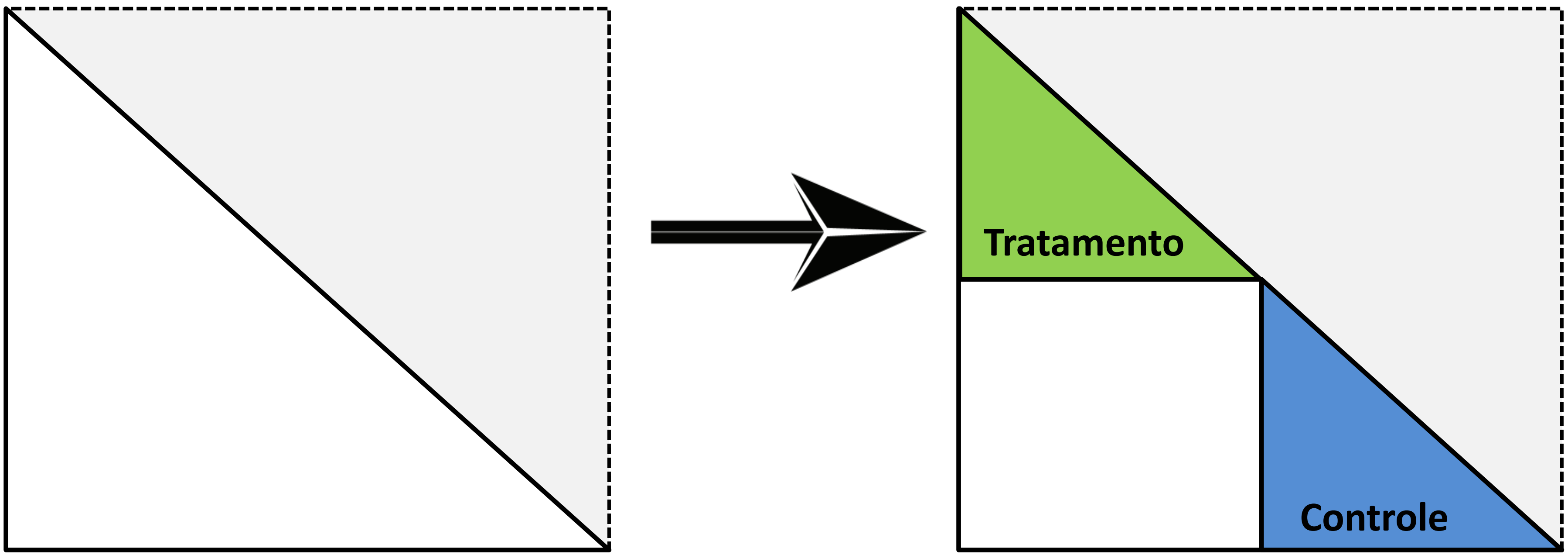


$$SS_T = \left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} {d_{ij}}^2\right]\Big/N$$

$$SS_w = \left[\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \epsilon_{ij} * {d_{ij}}^2\right]\Big/n$$

PERMANOVA (simples – 1 fator)

3)

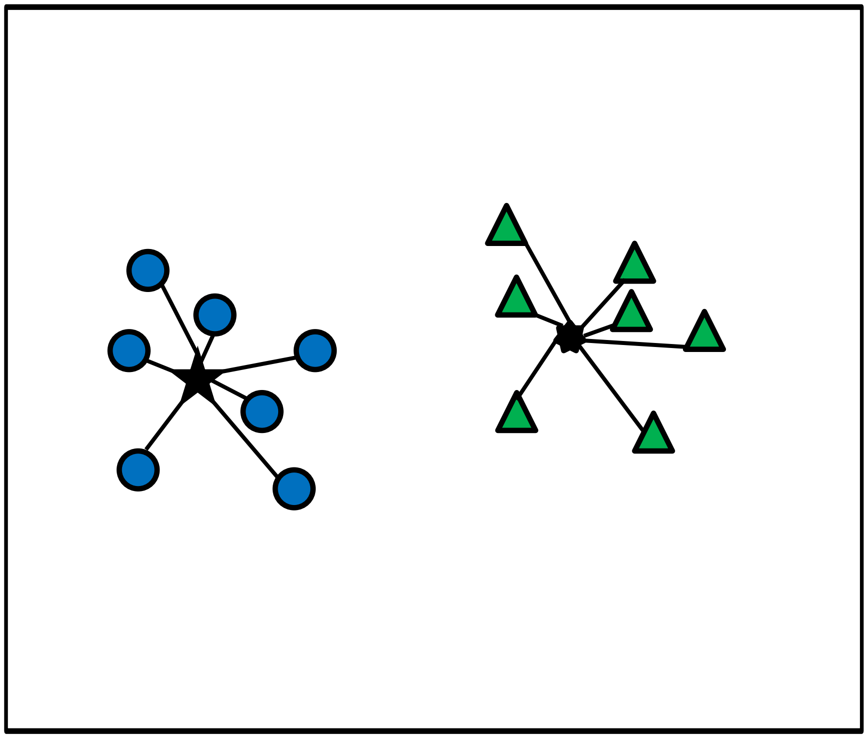


$$SS_A = SS_T - SSW$$

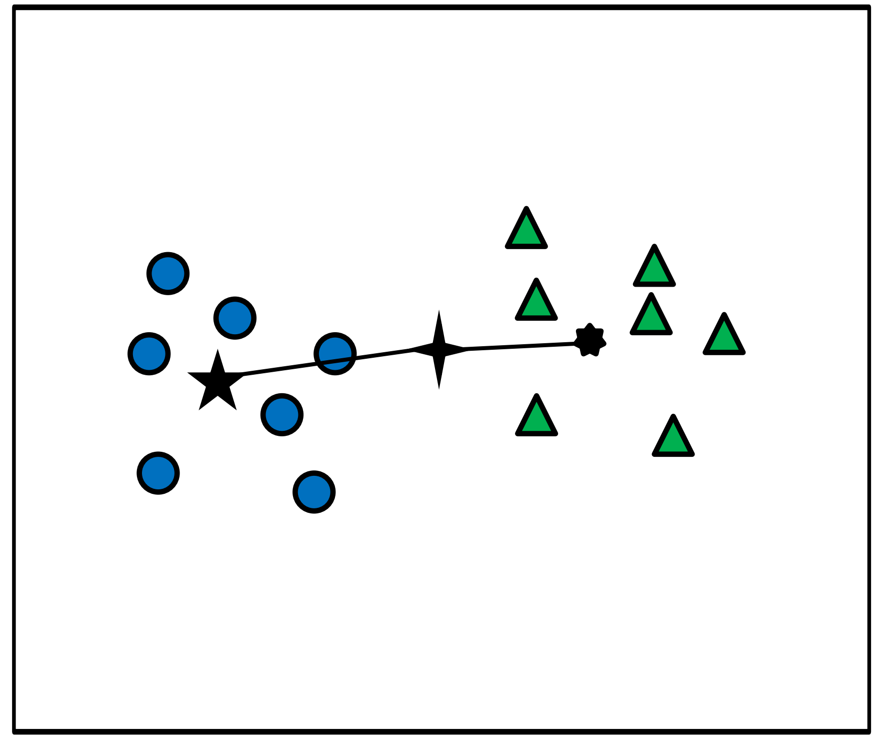
$$F = \frac{SS_A/(a-1)}{SS_W/(N-a)}$$

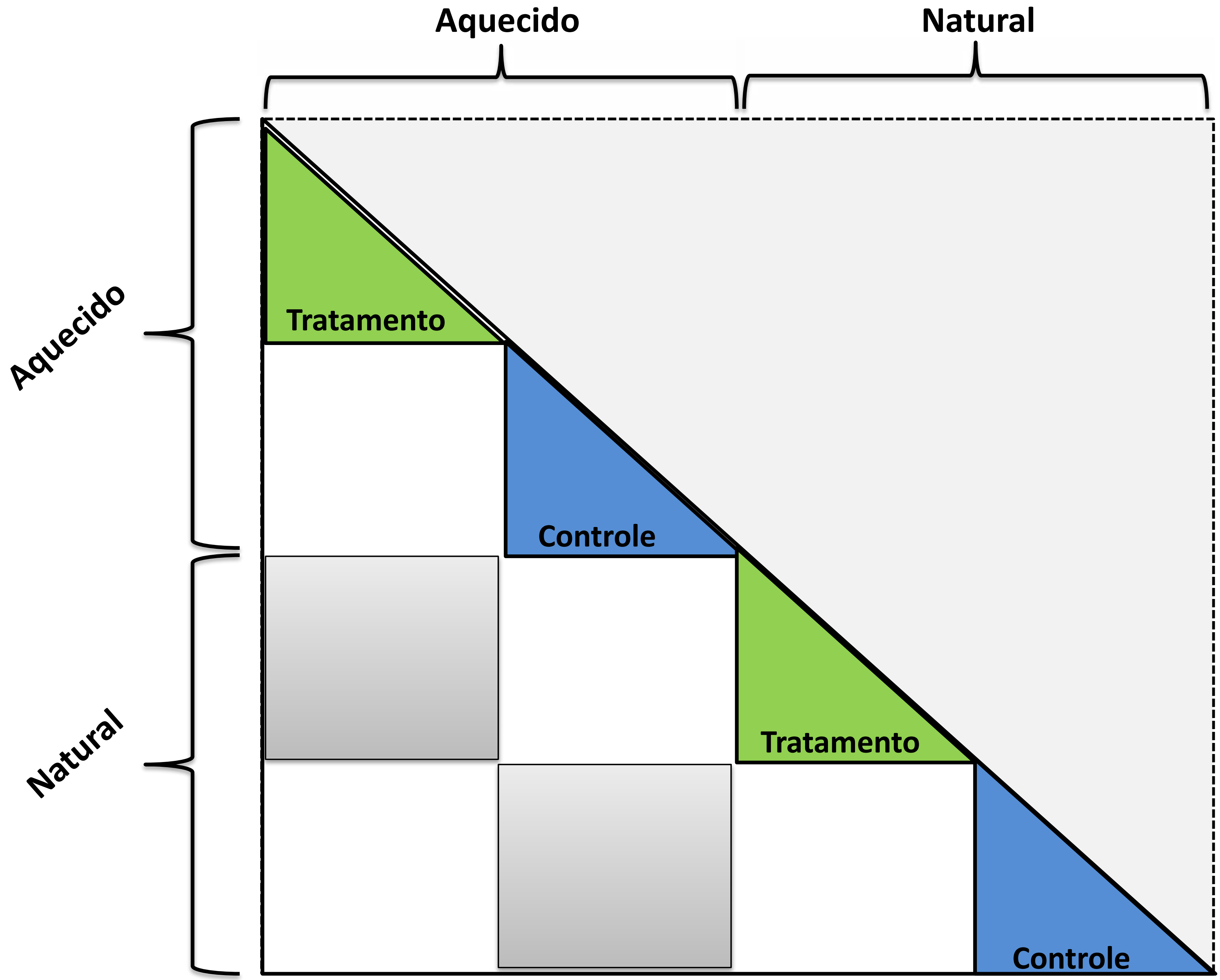
a = nº de grupos

Distância das amostras ao centróide do grupo



Distância dos centróides do grupo ao centróide geral

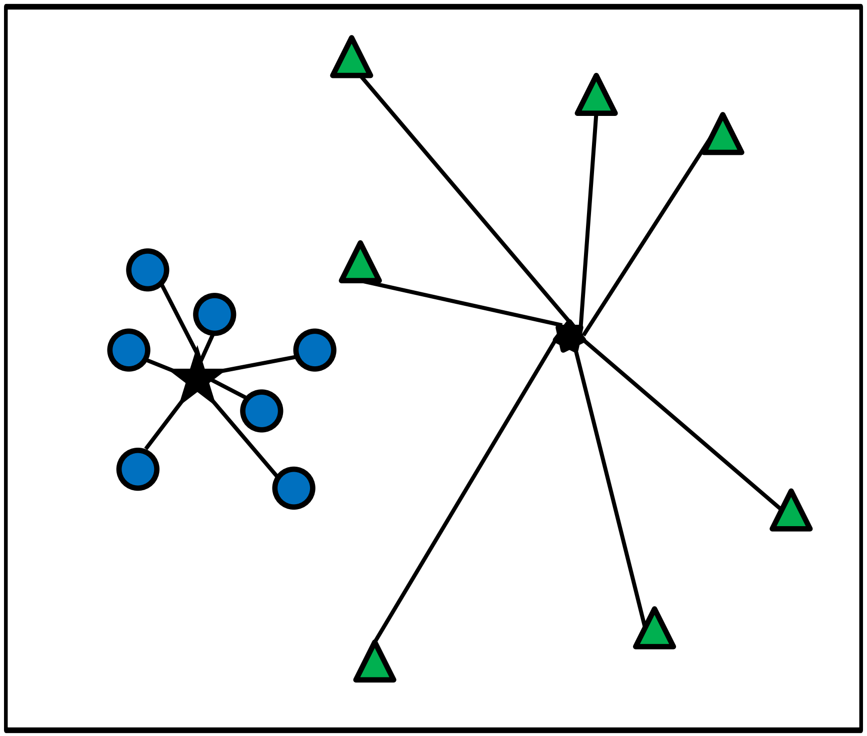




Aquecido
Natural
Aquecido
Natural
Tratamento
Controle
Tratamento
Controle

Cuidado! Heterogeneidade nas dispersões

- PIOR (ou proibitivo?) PARA ANOSIM E MANTEL



Anderson & Walsh 2013, Ecol Monogr 83

Tipos de perguntas

Mudança na posição: “Houve uma mudança fundamental na estrutura da comunidade?”

Mudança na dispersão: “a estrutura da comunidade se tornou mais ou menos variável?”

Ecological Monographs, 83(4), 2013, pp. 557–574

PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing?

Marti J. Anderson[1,3] and Daniel C. I. Walsh[2]

- Para desenhos balanceados, PERMANOVA é mais robusta que ANOSIM e Mantel
- Análises testam diferentes hipóteses nulas
    - ANOSIM e Mantel testam se *amostras no mesmo grupo não são mais agrupadas do que amostras de diferentes grupos*
    - PERMANOVA testa se *existe diferença nos centroides entre grupos*
- Leva em conta a posição do grupo no espaço multivariadao e não só a dispersão
- Com MANTEL e ANOSIM não é possível distinguir se a diferença é devido à posição, dispersão, ou formato dos pontos do grupo

— Diogo, preciso confessar, não gostei nada da disciplina, você é um péssimo professor.

—Tudo bem jovem, aqui tem o link pra o curso com o Legendre, com material e vídeos

http://adn.biol.umontreal.ca/~numericalecology/Trieste16/


That's all Folks!

Obrigado pela paciência
e perseverança!

Dúvidas?

Querem compartilhar o desespero?

diogo.provete@ufms.br
diogoprovete.weebly.com