

# UMA INTRODUÇÃO À ANÁLISE DE DADOS COMPOSICIONAIS

Ana Beatriz Tozzo Martins

Orientador: Prof. PhD Paulo Justiniano Ribeiro Junior.

3 de julho de 2008

# Motivação

- Ciências da Terra: dados expressos como frações ou porcentagens.
- Exemplos:
  - textura de solos;
  - composição química de uma rocha;
  - estruturas de dados resultantes de algoritmos de classificação.

Walvoort, D. J. J. e Gruijter, J.J. (2001).

# Qual o Tratamento Estatístico?

- Regressão Multivariada?
- Os componentes das composições somam 1?

# Dados Composicionais

- Butler e Glasbey (2008): Registram informação sobre frequências relativas associadas a diferentes componentes de um sistema.
- Aitchison (1986):
  - a. Vetores cujos elementos são proporções de algum todo.
  - b. **Composição:** Vetor  $\underline{X} = (X_1, X_2, \dots, X_D)'$  satisfazendo:
    - $X_1 \geq 0, \dots, X_D \geq 0$ ;
    - $X_1 + X_2 + \dots + X_D = 1$ .

# Características de Um Conjunto De Dados Composicionais

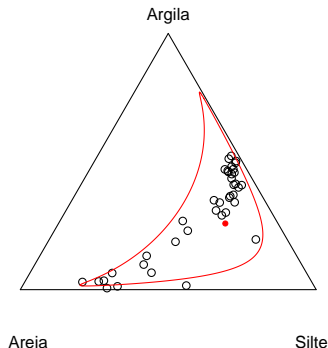
Reyment e Savazzi (1999) apud Labus (2005):

- representado em forma matricial;
- linha: unidade observacional;
- coluna: parte ou variável;
- elementos não negativos;
- cada linha soma 1 (100%);
- coeficientes de correlação mudam se uma das variáveis é excluída da matriz e as linhas somam 1 (100%). O mesmo ocorre se um novo componente é adicionado.

# Dados Lago Ártico: Aitchison (1986)

n	Areia	Silte	Argila
1	77.5	19.5	3
2	71.9	24.9	3.2
3	50.7	36.1	13.2
4	52.2	40.9	6.6
⋮	⋮	⋮	⋮
39	2	47.8	50.2

# Diagrama Ternário: Aitchison(1986)



**Figura:** Diagrama ternário para dados do Lago Ártico incluindo média geométrica e região 2-sigma.

# Construindo um Diagrama Ternário

- Triângulo equilátero de lado igual a 7,5 cm.

- $\underline{x} = (0, 17; 0, 33; 0, 50)$

- $$\begin{array}{lcl} 7,5 & \rightarrow & 100 \\ x_1 & \rightarrow & 17 \Rightarrow x_1 = 1,28\text{cm} \\ x_2 & \rightarrow & 33 \Rightarrow x_2 = 2,48\text{cm} \\ x_3 & \rightarrow & 50 \Rightarrow x_3 = 3,75\text{cm} \end{array}$$



# Construindo um Diagrama Ternário

- $(0, 17^{-1}; 0, 33^{-1}; 0, 50^{-1}) \Rightarrow (0, 54; 0, 28; 0, 18) \Rightarrow$

$$\begin{cases} x_1 &= 4,05 \\ x_2 &= 2,10 \\ x_3 &= 1,35 \end{cases}$$

- $\left. \begin{array}{l} \underline{x} = (0, 17; 0, 33; 0, 50) \\ \underline{y} = (0, 56; 0, 33; 0, 11) \end{array} \right\} \Rightarrow (0, 36; 0, 43; 0, 21) \Rightarrow$

$$\begin{cases} x_1 &= 2,70 \\ x_2 &= 3,23 \\ x_3 &= 1,60 \end{cases}$$

# Interpretando um Diagrama Ternário

- Para qualquer ponto no diagrama, as perpendiculares aos eixos satisfazem a restrição soma 1.
- Qualquer ponto numa linha paralela ao eixo oposto à um vértice tem a mesma proporção daquele componente correspondente a este vértice.

**Proporção:** distância relativa da linha ao eixo sobre a distância do vértice ao eixo.

Boogaart (2005)

# Interpretando um Diagrama Ternário

- Pontos em uma linha reta partindo de um dos vértices tem proporções relativas iguais dos outros dois componentes.

**Proporção:** proporções relativas representadas pelo ponto onde a linha cruza o eixo oposto.

Boogaart (2005)

- **Extensão da variabilidade da razão de 2 componentes:** extensão da intersecção das linhas retas partindo de cada um dos vértices e passando por todos os pontos com o eixo oposto.

Aitchison (1986)

# Dados Gener1Mod

n	X	Y	AMG	AM	AF	AMF	Silte	Argila
1	0	0	2.1	10.4	11.4	5.9	12.8	57
2	50	0	1.1	12.0	9.8	4.6	13.0	58
3	100	0	1.1	7.9	6.9	4.0	15.7	64
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
412	2800	300	2.5	5.3	31.2	13.5	12.4	35

# Dados Gener1Mod

$$\text{Areia} = \text{AMG} + \text{AM} + \text{AF} + \text{AMF}$$

$$\text{Silte} = 100 - \text{Areia} - \text{Argila}$$

n	X	Y	Areia	Silte	Argila	Soma	Altim.	Sup
1	0	0	29.8	13.2	57	100	641.37	1
2	50	0	27.5	14.5	58	100	640.84	1
3	100	0	19.9	16.1	64	100	640.00	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
412	2800	300	52.5	12.5	35	100	520.00	0

# Dados Basso

n	AGrossa	Silte	Argila	Soma
1	9	26	43	78
2	9	26	42	77
3	9	25	41	75
⋮	⋮	⋮	⋮	⋮
250	10	27	37	74

# Dados Basso

Formando composições

n	AGrossa	Silte	Argila	Soma
1	11.53846	33.33333	55.12821	100
2	11.68831	33.76623	54.54545	100
3	12.00000	33.33333	54.66667	100
⋮	⋮	⋮	⋮	⋮
250	13.513514	36.48649	50.00000	100

# Dados Basso

Diferença = Areia Fina

n	Areia	Silte	Argila	Soma
1	31	26	43	100
2	32	26	42	100
3	34	25	41	100
⋮	⋮	⋮	⋮	⋮
250	36	27	37	100



# Composição Regionalizada

Seja  $\Omega \subset \mathbb{R}^n$  domínio espacial,  $\mathbb{R}^n$  espaço real de dimensão  $n$ .

- Composição Regionalizada:**

$\underline{X}(\underline{u}) = [X_1(\underline{u}), X_2(\underline{u}), \dots, X_D(\underline{u})]'$ ,  $\underline{u} \in \Omega$ , satisfazendo:

- 1.  $X_1(\underline{u}) \geq 0, \dots, X_D(\underline{u}) \geq 0$ ;
- 2.  $X_1(\underline{u}) + X_2(\underline{u}) + \dots + X_D(\underline{u}) = 1$ .

- Espaço Amostral:**

$$\mathbb{S}^D = \{\underline{X}(\underline{u}) \in \mathbb{R}^D; X_i(\underline{u}) > 0, i = 1, \dots, D; \sum_{i=1}^D X_i(\underline{u}) = 1\}$$

# Base Regionalizada

- Vetor  $\underline{W}(\underline{u})$ ,  $\underline{u} \in \Omega \subset \mathbb{R}^n$  cujos componentes são todos medidos na mesma escala e positivos.

- **Espaço Amostral:**

$$\mathbb{R}_+^D = \{\underline{W}(\underline{u}) \in \mathbb{R}^D; W_i(\underline{u}) > 0, i = 1, \dots, D\}$$

- **Operador fechamento:** Base  $\Rightarrow$  Composição

$$\begin{aligned} \mathcal{C} : \quad \mathbb{R}_+^D &\longrightarrow \mathbb{S}^D \\ \underline{W}(\underline{u}) &\longrightarrow \mathcal{C}[\underline{W}(\underline{u})] = \frac{\underline{W}(\underline{u})}{\underline{j}'\underline{W}(\underline{u})}, \end{aligned}$$

$\underline{j}'$  vetor de  $1$ 's.

# Subcomposição Regionalizada

- Seja  $\underline{X}(\underline{u})$  composição regionalizada de  $D$  partes;
- $s \subset \{1, 2, \dots, D\}$  tal que  $\underline{X}_s(\underline{u})$  é subvetor cujos elementos são os componentes de  $\underline{X}(\underline{u})$  correspondentes às partes em  $s$ ;
- **Subcomposição Regionalizada:**

$$\mathcal{C}[\underline{X}_s(\underline{u})] = \frac{\underline{X}_s(\underline{u})}{\sum_{j \in s} \underline{X}_j(\underline{u})};$$

- Se  $\underline{s}(\underline{u}) = \mathcal{C}[\underline{X}_s(\underline{u})] \Rightarrow \frac{s_i(\underline{u})}{s_j(\underline{u})} = \frac{X_i(\underline{u})}{X_j(\underline{u})}, \quad i, j \in s.$

# Amalgamação

- $\underline{X}(\underline{u}) = [X_1(\underline{u}), X_2(\underline{u}), \dots, X_D(\underline{u})]'$  uma composição regionalizada de  $D$  partes;
- $\underline{X}(\underline{u})$  dividida em  $C$  ( $C \leq D$ ) subconjuntos mutuamente exclusivos;
- $\underline{X}_i(\underline{u})$ ,  $i = 1, \dots, C$ , vetor cujos componentes são os elementos do  $i$ -ésimo subconjunto.
- **Amalgamação Regionalizada:** composição regionalizada de  $C$  partes

$$\underline{A}(\underline{u}) = [A_1(\underline{u}), A_2(\underline{u}), \dots, A_C(\underline{u})]'$$

em que  $A_i(\underline{u}) = \underline{j}'\underline{X}_i(\underline{u})$ ,  $\underline{j}'$  vetor de  $1^s$ .

# Partição

**Partição Regionalizada:** Amalgamação juntamente com subcomposições.

- $P_C[\underline{X}(\underline{u})] = [\underline{s}_1(\underline{u}), \underline{s}_2(\underline{u}), \dots, \underline{s}_C(\underline{u}); \underline{A}(\underline{u})]$

em que  $\underline{s}_i(\underline{u}) = \mathcal{C}[X_i(\underline{u})]$ ,  $i = 1, \dots, C$ .

- Preserva informações contidas nos subvetores.

# Problema - Dados Composicionais

Efeito de **correlação espúria** (Pawlowsky e Olea, 2004):

- Covariâncias sujeitas à controles não estocásticos



interpretação errônea da estrutura de covariância espacial;

- Singularidade da matriz de covariância de uma composição.

# Transformação

- **Graf (2006):** Soma constante  $\Rightarrow$  correlação negativa entre os componentes.
- **Aitchison (1986):**
  - Propõe transformação que generaliza a transformação logística  $\ln \frac{X}{1-X}$  para um vetor composicional de 2 partes;
  - Magnitudes relativas ou razões  $\Rightarrow$  tratabilidade e interpretação estatística.
- **Transformações logísticas:** Aitchison (1982), Aitchison et al. (2000), Odeh et al. (2003).

# Transformação

**Transformação razão log-aditiva (ALR):**

$$\begin{aligned} \text{ALR: } \mathbb{S}^D &\longrightarrow \mathbb{R}^{D-1} \\ \underline{X}(\underline{u}) &\longrightarrow \text{ALR}[\underline{X}(\underline{u})] = \left( \ln \frac{X_1(\underline{u})}{X_D(\underline{u})}, \dots, \ln \frac{X_{D-1}(\underline{u})}{X_D(\underline{u})} \right)'. \end{aligned}$$

Pawlowsky et al. (1995), Pawlowsky e Olea (2004).



# Transformação

**Transformação logística generalizada aditiva (AGL):**

$$\begin{aligned} \text{AGL} : \quad \mathbb{R}^{D-1} &\longrightarrow \mathbb{S}^D \\ \text{ALR}[\underline{X}(\underline{u})] &\longrightarrow \text{AGL}\{\text{ALR}[\underline{X}(\underline{u})]\} = \underline{X}(\underline{u}) = \frac{\exp[\underline{Y}(\underline{u})]}{\underline{j}' \exp[\underline{Y}(\underline{u})]}, \end{aligned}$$

$$\underline{Y}(\underline{u}) = (X_1(\underline{u}), X_2(\underline{u}), \dots, 0)'$$

$$\text{Se } c \neq 0 \quad \Rightarrow \quad \underline{X}(\underline{u}) = c \cdot \text{AGL}\{\text{ALR}[\underline{X}(\underline{u})]\}.$$

# Transformação

## Transformação Razão Log Centrada (CLR):

$$\begin{aligned} \text{CLR} : \quad \mathbb{R}_+^D &\rightarrow \mathbb{R}^D \\ \underline{W}(\underline{u}) &\rightarrow \text{CLR}[\underline{W}(\underline{u})] = \ln \frac{\underline{W}(\underline{u})}{g[\underline{W}(\underline{u})]} \end{aligned}$$

ou

$$\begin{aligned} \text{CLR} : \quad \mathbb{S}^D &\longrightarrow \mathbb{R}^D \\ \underline{X}(\underline{u}) &\longrightarrow \text{CLR}[\underline{X}(\underline{u})] = \ln \frac{\underline{X}(\underline{u})}{g[\underline{X}(\underline{u})]} \end{aligned}$$

$g$ : média geométrica.

# Perturbação e Potência

Aitchison (1986), Aitchison e Egozcue(2005), Tolosana-Delgado et al.(2005):

Operações que definem uma estrutura de espaço vetorial de dimensão  $D - 1$  no simplex.

- 1. **Perturbação** ( $\oplus$ ) : operação comutativa;
- 2. **Potência** ( $\odot$ ) : produto externo.

# Perturbação

**Perturbação:** Multiplicação de composições componente a componente e divisão de cada componente pela soma de todos.

$$\begin{aligned}\underline{X} \oplus \underline{Y} &= (X_1, X_2, \dots, X_D) \oplus (Y_1, Y_2, \dots, Y_D) \\ &= \mathcal{C}(X_1 Y_1, X_2 Y_2, \dots, X_D Y_D).\end{aligned}$$

# Potência

- **Potência**: análogo à multiplicação por um escalar no espaço real.

$$\alpha \odot (X_1, X_2, \dots, X_D) = \mathcal{C}(X_1^\alpha, X_2^\alpha, \dots, X_D^\alpha).$$

- Vetor de **diferenças** composicionais:

$$\underline{X} \ominus \underline{Y} = \underline{X} \oplus (-1 \odot \underline{Y}).$$

# Geometria de Aitchison do Simplex

- **Produto interno:**

$$\langle \underline{X}, \underline{Y} \rangle = \sum_{i=1}^D \ln \frac{X_i}{g(\underline{X})} \ln \frac{Y_i}{g(\underline{Y})}.$$

- **Distância de Aitchison:**

$$d(\underline{X}, \underline{Y}) = \sqrt{\sum_{i=1}^D \left( \ln \frac{X_i}{g(\underline{X})} - \ln \frac{Y_i}{g(\underline{Y})} \right)^2}$$

# Estatísticas Descritivas

- Centro:**

$$\text{cen}(\underline{X}) = \frac{1}{g_s} [g(X_1) \quad g(X_2) \quad \dots \quad g(X_D)]'$$

- $g(X_i)$  - média geométrica do  $i$ -ésimo componente
- $g_s = g(X_1) + g(X_2) + \dots + g(X_D)$ .

- Medida de dispersão:**

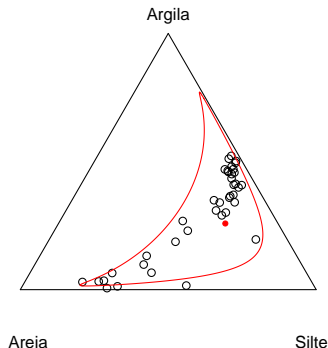
$$\text{Totvar}(\underline{X}) = \frac{1}{D} \sum_{i < j} \text{var} \left[ \ln \left( \frac{X_i}{X_j} \right) \right] = \text{tr}(\Sigma) - \frac{1}{D} \mathbf{1}_d' \Sigma \mathbf{1}_d$$

# Domínio de Confiança

$$D'_{1-\alpha}(\underline{X}) = \left\{ \underline{X} \in S^D / \left( \ln \frac{\underline{X}_{-D}}{\underline{X}_D} - \mu \right)' \Sigma^{-1} \left( \ln \frac{\underline{X}_{-D}}{\underline{X}_D} - \mu \right) \leq \chi^2_{d;1-\alpha} \right\}$$



# Diagrama Ternário: Aitchison(1986)



**Figura:** Diagrama ternário para dados do Lago Ártico incluindo média geométrica e região 2-sigma.

# Matriz Diagrama de Dispersão

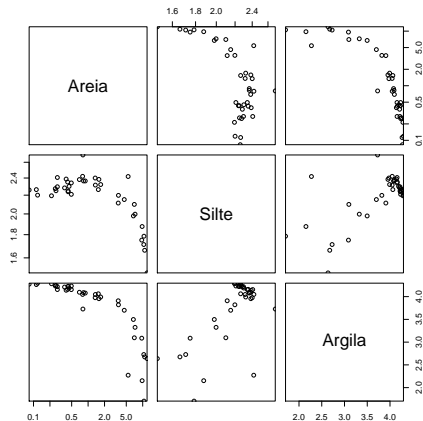
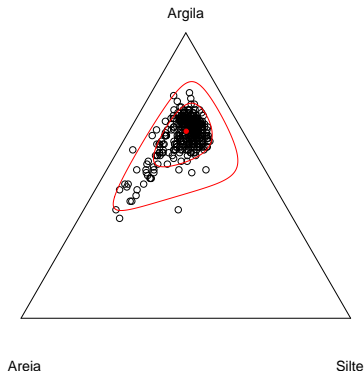


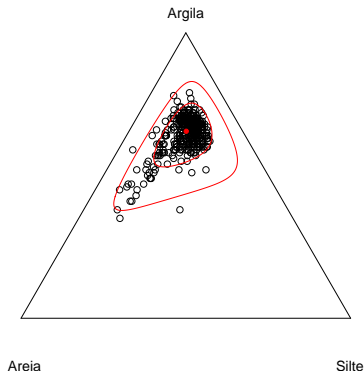
Figura: Areia vs Silte, Areia vs Argila e Silte vs Argila - Lago Ártico.

# Dados reais: Gener - Silte



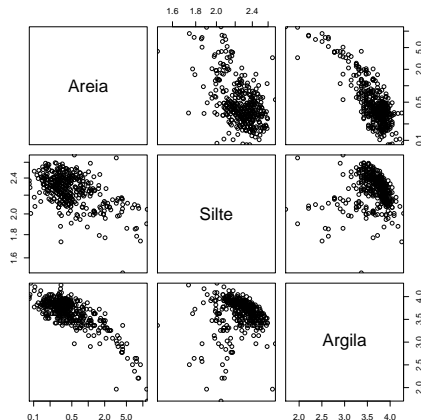
**Figura:** Diagrama ternário para dados de Gener com Silte obtido pelo complementar, incluindo média geométrica e região 2,4-sigma.

# Dados reais: Gener - Areia



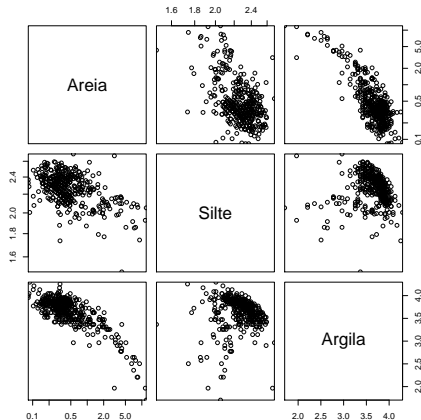
**Figura:** Diagrama ternário para dados de Gener com Areia obtida pelo complementar, incluindo média geométrica e região 2,4-sigma.

# Matriz Diagrama de Dispersão: Gener - Silte.



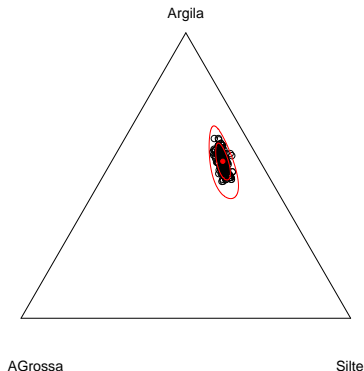
**Figura:** Areia vs Silte, Areia vs Argila e Silte vs Argila - Silte obtido pelo complementar.

# Matriz Diagrama de Dispersão: Gener - Areia.



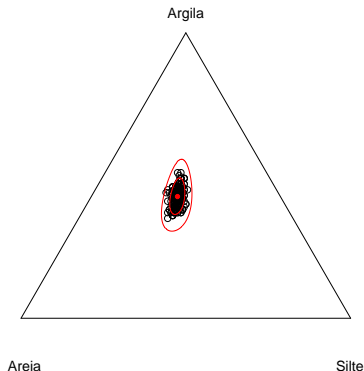
**Figura:** Areia vs Silte, Areia vs Argila e Silte vs Argila - Areia obtida pelo complementar.

# Dados reais: Bassoi - Composições.



**Figura:** Diagrama ternário para dados de Bassoi incluindo média geométrica e região 2,4-sigma-Formando composições.

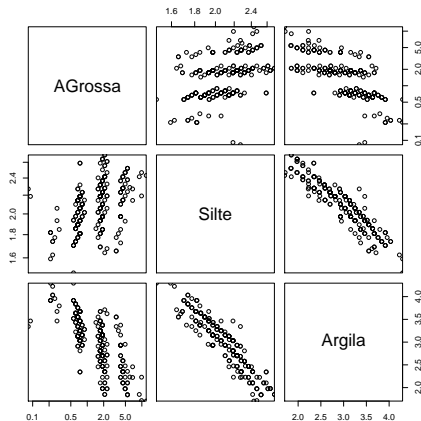
# Dados reais: Bassoi - Areia Fina.



**Figura:** Diagrama ternário para dados de Bassoi com Areia Fina como complementar, incluindo média geométrica e região 2,4-sigma.

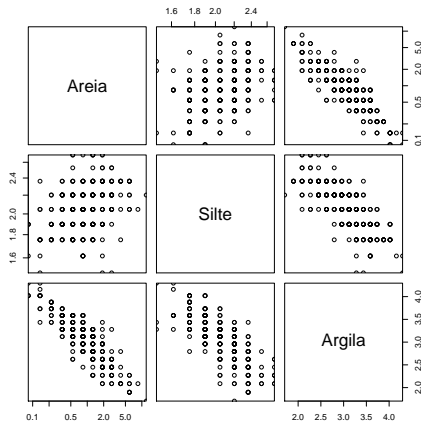


# Matriz Diagrama de Dispersão: Bassoi - Composições.



**Figura:** AGrossa vs Silte, AGrossa vs Argila e Silte vs AGrossa; Formando composições.

# Matriz Diagrama de Dispersão: Bassoí - Areia Fina.



**Figura:** Areia vs Silte, Areia vs Argila e Silte vs Areia - Areia Fina como complementar.

# Estacionariedade

Segundo Pawlowsky-Glahn e Olea (2004)

$\underline{X}(\underline{u})$  é **Estacionária de 2ª ordem** se  $f[\underline{X}(\underline{u})]$  satisfaz:

- $E[f(\underline{X}(\underline{u}))] = \mu$  existe e não depende de  $\underline{u}$ ;
- $\text{Cov}[f(\underline{X}(\underline{u}_2)), f(\underline{X}(\underline{u}_1))] = \Sigma(\underline{u}_2 - \underline{u}_1)$  existe e não depende de  $\underline{u}_1, \underline{u}_2$ , mas somente da diferença  $\underline{h} = \underline{u}_2 - \underline{u}_1$ .

# Função Intrínseca

$\underline{X}(\underline{u})$  é **função intrínseca** se  $f[\underline{X}(\underline{u})]$  satisfaz:

- $E[f(\underline{X}(\underline{u}))] = \mu$  existe e não depende de  $\underline{u}$ ;
- $\text{Cov}[f(\underline{X}(\underline{u}_2)) - f(\underline{X}(\underline{u}_1))] = \mathbf{\Gamma}(\underline{u}_2 - \underline{u}_1)$

existe e não depende de  $\underline{u}_1$ ,  $\underline{u}_2$ , mas somente da diferença  $\underline{h} = \underline{u}_2 - \underline{u}_1$ .

# Função Intrínseca

- **LR estacionária:**  $X(\underline{u})$  é log razão estacionária de 2ª ordem se o conjunto das log razão entre todos os pares é estacionário de 2ª ordem;
- **LR intrínseca:**  $X(\underline{u})$  é log razão intrínseca se o conjunto das log razão entre todos os pares é intrínseca;

# Estrutura de Covariância Espacial

**Estrutura de Covariância Espacial:** conjunto de funções  $D^4$

$$\sigma_{ij \cdot kl}(\underline{h}) = \text{Cov} \left[ \ln \frac{X_i(\underline{u})}{X_k(\underline{u})}, \ln \frac{X_j(\underline{u} + \underline{h})}{X_l(\underline{u} + \underline{h})} \right],$$

$i, j, k, l \in \{1, 2, \dots, D\}$ , para  $\underline{u}, \underline{u} + \underline{h} \in \Omega$ .

Pawlowsky e Olea (2004).

# Estrutura de Covariância Espacial

- Matriz de **LR autocovariâncias** (Matriz Variação)

$$\mathbf{T}(\underline{h})_{D \times D} = (\tau_{i \cdot j}(\underline{h})) = (\sigma_{ii \cdot jj}(\underline{h}))$$

- Matriz de **ALR covariâncias cruzadas**

$$\mathbf{\Sigma}(\underline{h})_{D-1 \times D-1} = (\sigma_{ij}(\underline{h})) = (\sigma_{ij \cdot DD}(\underline{h}))$$

- Matriz de **CLR covariâncias cruzadas**,  $\mathbf{\Xi}(\underline{h})$  com elementos

$$\xi_{ij}(\underline{h}) = \text{Cov} \left[ \ln \frac{X_i(\underline{u})}{g[\underline{X}(\underline{u})]}, \ln \frac{X_j(\underline{u} + \underline{h})}{g[\underline{X}(\underline{u} + \underline{h})]} \right],$$

# Estrutura de Covariância Espacial Intrínseca

**Estrutura de Cov. Espacial Intrínseca:** conjunto de funções  $D^4$

$$V_{ij \cdot kl}(\underline{h}) = \frac{1}{2} \text{Cov} \left[ \ln \frac{\underline{X}_i(\underline{u})}{\underline{X}_k(\underline{u})} - \ln \frac{\underline{X}_i(\underline{u} + \underline{h})}{\underline{X}_k(\underline{u} + \underline{h})}, \ln \frac{\underline{X}_j(\underline{u})}{\underline{X}_l(\underline{u})} - \ln \frac{\underline{X}_j(\underline{u} + \underline{h})}{\underline{X}_l(\underline{u} + \underline{h})} \right]$$

$i, j, k, l \in \{1, 2, \dots, D\}$  para  $\underline{u}, \underline{u} + \underline{h} \in \Omega$ .



# Estrutura de Covariância Espacial Intrínstica

- Matriz de **LR variogramas** (Matriz Variação Intrínstica)

$$\Gamma(\underline{h})_{D \times D} = (V_{i \cdot j}(\underline{h})) = (V_{ii \cdot jj}(\underline{h}))$$

- Matriz de **ALR variograma cruzado**

$$\Psi(\underline{h})_{D-1 \times D-1} = (\psi_{ij}(\underline{h})) = (V_{ii \cdot DD}(\underline{h}))$$

- Matriz de **CLR covariâncias cruzadas**,  $\Delta(\underline{h})$  com elementos

$$\delta_{ij}(\underline{h}) = \text{Cov} \left[ \ln \frac{X_i(\underline{u})}{g[\underline{X}(\underline{u})]}, \ln \frac{X_j(\underline{u} + \underline{h})}{g[\underline{X}(\underline{u} + \underline{h})]} \right]$$

# Bibliografia

- AITCHISON, J. The statistical analysis of compositional data. **Journal of the Royal Statistical Society, Series B**, v. 44, n.2, p. 139-177, 1982.
- AITCHISON, J. **The statistical analysis of compositional data**. New Jersey: The Blackburn Press, 1986.
- AITCHISON, J.; BARCELÓ-VIDAL, C.; MARTÍN-FERNANDÉZ, J. A.; PAWLOWSKY-GLAHN, V. Logratios analysis and compositional distance. **Mathematical Geology**, v. 32, n.3, p. 563-580, 2000.
- AITCHISON, J.; EGOZCUE, J. J. Compositional data analysis: where are we and where should we be heading?. **Mathematical Geology**, v. 37, n.7, p. 829-850, oct. 2005.
- BOOGAART, K.G.v.d. Using the R package "compositions", v. 0.9, 2005. Disponível em:  
<<http://www.stat.boogaart.de/compositions/UsingCompositions.pdf>>.  
Acesso em: 09/02/08.

# Bibliografia

- BUTLER, A.; GLASBEY, C. A latent Gaussian model for compositional data with zeros. **Journal of the Royal Statistical Society, Series C** (in press), 2008. Disponível em: <[http://www.bioss.ac.uk/staff/adam/documents/ButlerGlasbey\\_resubmit.pdf](http://www.bioss.ac.uk/staff/adam/documents/ButlerGlasbey_resubmit.pdf)>. Acesso em: 13/02/08.
- LABUS, M. Compositional data analysis as a tool for interpretation of rock porosity parameters. **Geological Quarterly**, v. 49, n.3, p. 347-354, 2005.
- PAWLOWSKY-GLAHN, V.; OLEA, R. A.; DAVIS, J. C. Estimation of regionalized compositions: a comparison of three methods. **Mathematical Geology**, v.27, n.1, p. 105-127. 1995.
- PAWLOWSKY-GLAHN, V.; OLEA, R. A. **Geostatistical analysis of compositional data**. New York: Oxford University Press, Inc., 2004.
- REYMENT, R. A. and SAVAZZI E. **Aspects of multivariate statistical analysis in geology**. Amsterdam: Elsevier, 1999.

# Bibliografia

- GRAF, M. Precision of compositional data in a stratified two-Stage cluster sample: comparison of the swiss earnings structure survey 2002 and 2004. **Survey Research Methods Section, ASA** , Session 415: Sample Survey Quality V, 3066-3072, 2006. Disponível em: <<http://www.amstat.org/sections/SRMS/proceedings/y2006/Files/JSM2006-000771.pdf>>. Acesso em: 18/05/08.
- ODEH, I. O. A.; TOOD, A. J.; TRIANTAFILIS, J. Spatial prediction of soil particle-size fractions as compositional data. **Soil Science**, v. 168, n.7, p. 501-515, July 2003.
- TOLOSANA-DELGADO, R.; OTERO, N.; PAWLOWSKY-GLAHN, V. Some Basic Concepts of Compositional Geometry. **Mathematical Geology**, v. 37, n.7, p. 563-580, oct 2005.
- WALVOORT, D. J. J.; GRUIJTER, J. J. Compositional kriging: A spatial interpolation method for compositional data. **Mathematical Geology**, v.33, n.8, p. 951-966, nov 2001.

**OBRIGADA PELA ATENÇÃO!**

# Agradecimentos

- PhD Paulo Justiniano Ribeiro Junior
- Laboratório de Estatística e Geoinformação - LEG