

## ANÁLISE DOS DADOS COMPOSICIONAIS DE Bassoil

Considerando a diferença como Areia Fina

Carregando o pacote *compositions*

```
> require(compositions)
```

### PREPARANDO OS DADOS PARA ANÁLISE

Lendo os dados

```
> dados <- read.table("DadosBassoil.txt",head=TRUE)
> # outra maneira de ler os dados seria
> # dados <- read.csv('DadosBassoil.csv',head=TRUE)
> options(width=80)
> head(dados)
```

	Linha	Coluna	Cota	AGrossa	Silte	Argila	pHAgua	pHKCl	Ca	Mg	K	Al	H
1	1	1	578.295	9	26	43	5.8	4.9	3.6	0.8	0.50	0	3.1
2	1	2	578.460	9	26	42	5.9	4.9	3.4	0.8	0.44	0	3.2
3	1	3	578.491	9	25	41	5.9	4.9	3.7	0.9	0.59	0	2.5
4	1	4	578.699	11	28	40	5.8	4.9	3.7	0.8	0.52	0	3.5
5	1	5	578.749	9	27	41	5.8	5.0	4.2	0.9	0.56	0	3.4
6	1	6	578.726	8	27	43	6.1	5.1	4.1	0.9	0.56	0	3.1

	C	N	CTC	S	V	M	NC	CEC	CN
1	1.2	0.12	8.00	4.90	61.250	0	1.050	4.90	10
2	1.1	0.12	7.84	4.64	59.184	0	1.272	4.64	9
3	1.2	0.13	7.69	5.19	67.490	0	0.289	5.19	9
4	1.3	0.12	8.52	5.02	58.920	0	1.416	5.02	10
5	1.4	0.13	9.06	5.66	62.472	0	1.023	5.66	10
6	1.2	0.13	8.66	5.56	64.203	0	0.753	5.56	9

Criando a coluna Areia Fina (AFina) que contém o complementar das colunas AGrossa, Silte e Argila

```
> dadosC <-with(dados,matrix(c(Linha,Coluna,Cota,AGrossa,Silte,Argila,
+                               AFina=(100-AGrossa-Silte-Argila),pHAgua,pHKCl,
+                               Ca,Mg,K,Al,H,C,N,CTC,S,V,M,NC,CEC,CN),nrow=250))
```

Transformando dadosC em data.frame

```
> dadosC <- as.data.frame(dadosC)
> names(dadosC) <- c('Linha','Coluna','Cota','AGrossa','Silte','Argila',
+                    'AFina','pHAgua','pHKCl','Ca','Mg','K','Al','H','C',
+                    'N','CTC','S','V','M','NC','CEC','CN')
> head(dadosC)
```

	Linha	Coluna	Cota	AGrossa	Silte	Argila	AFina	pHAgua	pHKCl	Ca	Mg	K	Al
1	1	1	578.295	9	26	43	22	5.8	4.9	3.6	0.8	0.50	0
2	1	2	578.460	9	26	42	23	5.9	4.9	3.4	0.8	0.44	0
3	1	3	578.491	9	25	41	25	5.9	4.9	3.7	0.9	0.59	0
4	1	4	578.699	11	28	40	21	5.8	4.9	3.7	0.8	0.52	0
5	1	5	578.749	9	27	41	23	5.8	5.0	4.2	0.9	0.56	0
6	1	6	578.726	8	27	43	22	6.1	5.1	4.1	0.9	0.56	0

	H	C	N	CTC	S	V	M	NC	CEC	CN
1	3.1	1.2	0.12	8.00	4.90	61.250	0	1.050	4.90	10
2	3.2	1.1	0.12	7.84	4.64	59.184	0	1.272	4.64	9
3	2.5	1.2	0.13	7.69	5.19	67.490	0	0.289	5.19	9
4	3.5	1.3	0.12	8.52	5.02	58.920	0	1.416	5.02	10
5	3.4	1.4	0.13	9.06	5.66	62.472	0	1.023	5.66	10
6	3.1	1.2	0.13	8.66	5.56	64.203	0	0.753	5.56	9

Calculando Areia que é a soma de AGrossa e AFina

```
> dadosC <- with(dadosC, matrix(c(Linha, Coluna, Cota, Areia=(AGrossa+AFina),
+                               Silte, Argila, pHAgua, pHKCl, Ca, Mg, K, Al, H, C,
+                               N, CTC, S, V, M, NC, CEC, CN), nrow=250))
```

Transformando dadosC em data.frame

```
> dadosC <- as.data.frame(dadosC)
> names(dadosC) <- c('Linha','Coluna','Cota','Areia','Silte','Argila',
+                    'pHAgua','pHKCl','Ca','Mg','K','Al','H','C','N',
+                    'CTC','S','V','M','NC','CEC','CN')
> options(width=80)
> head(dadosC)
```

	Linha	Coluna	Cota	Areia	Silte	Argila	pHAgua	pHKCl	Ca	Mg	K	Al	H	C
1	1	1	578.295	31	26	43	5.8	4.9	3.6	0.8	0.50	0	3.1	1.2
2	1	2	578.460	32	26	42	5.9	4.9	3.4	0.8	0.44	0	3.2	1.1

	N	CTC	S	V	M	NC	CEC	CN
1	0.12	8.00	4.90	61.250	0	1.050	4.90	10
2	0.12	7.84	4.64	59.184	0	1.272	4.64	9
3	0.13	7.69	5.19	67.490	0	0.289	5.19	9
4	0.12	8.52	5.02	58.920	0	1.416	5.02	10
5	0.13	9.06	5.66	62.472	0	1.023	5.66	10
6	0.13	8.66	5.56	64.203	0	0.753	5.56	9

Confirmando que Areia, Silte e Argila formam uma composição

```
> dadosC <-with(dadosC,matrix(c(Linha,Coluna,Cota,Areia,Silte,Argila,
+                               Soma=(Areia+Silte+Argila),pHAgua,pHKCl,Ca,Mg,K,Al,H,C,
+                               N,CTC,S,V,M,NC,CEC,CN),nrow=250))
```

Novamente transformando dadosC em data.frame

```
> dados <- as.data.frame(dadosC)
> names(dados) <- c("Linha", "Coluna", "Cota", "Areia", "Silte",
+ "Argila", "Soma", "pHAgua", "pHKCl", "Ca", "Mg", "K", "Al",
+ "H", "C", "N", "CTC", "S", "V", "M", "NC", "CEC", "CN")
> options(width = 80)
> head(dados)
```

	Linha	Coluna	Cota	Areia	Silte	Argila	Soma	pHAgua	pHKCl	Ca	Mg	K	Al	H
1	1	1	578.295	31	26	43	100	5.8	4.9	3.6	0.8	0.50	0	3.1
2	1	2	578.460	32	26	42	100	5.9	4.9	3.4	0.8	0.44	0	3.2
3	1	3	578.491	34	25	41	100	5.9	4.9	3.7	0.9	0.59	0	2.5
4	1	4	578.699	32	28	40	100	5.8	4.9	3.7	0.8	0.52	0	3.5
5	1	5	578.749	32	27	41	100	5.8	5.0	4.2	0.9	0.56	0	3.4
6	1	6	578.726	30	27	43	100	6.1	5.1	4.1	0.9	0.56	0	3.1

	C	N	CTC	S	V	M	NC	CEC	CN
1	1.2	0.12	8.00	4.90	61.250	0	1.050	4.90	10
2	1.1	0.12	7.84	4.64	59.184	0	1.272	4.64	9
3	1.2	0.13	7.69	5.19	67.490	0	0.289	5.19	9
4	1.3	0.12	8.52	5.02	58.920	0	1.416	5.02	10

```
5 1.4 0.13 9.06 5.66 62.472 0 1.023 5.66 10
6 1.2 0.13 8.66 5.56 64.203 0 0.753 5.56 9
```

Criando o arquivo com os dados composicionais

```
> dados <- with(dados, matrix(c(Areia, Silte, Argila), nrow = 250))
> head(dados)
```

```
      [,1] [,2] [,3]
[1,]    31    26    43
[2,]    32    26    42
[3,]    34    25    41
[4,]    32    28    40
[5,]    32    27    41
[6,]    30    27    43
```

## INICIANDO ANÁLISE COMPOSICIONAL

Na classe `acomp` utilizada a seguir as quantidades individuais são parte de um todo e os dados são analisados na geometria relativa.

```
> comp <- acomp(dados)
> names(comp) <- c("Areia", "Silte", "Argila")
> head(comp)
```

```
[1] 0.31 0.32 0.34 0.32 0.32 0.30
```

```
> par(mar = c(5, 0.5, 0, 0.5), mgp = c(2, 0.8, 0))
> plot(comp)
```

Pela Figura 1 pode-se observar que as proporções dos componentes nas composições são aproximadamente iguais, com proporções menores para o componente Silte. Ainda, a razão Areia/Silte apresenta maior variabilidade que a razão Argila/Silte.

Vendo os dados como quantidades

```
> quant <- aplus(dados)
> names(quant) <- c("Areia", "Silte", "Argila")
> head(quant)
```

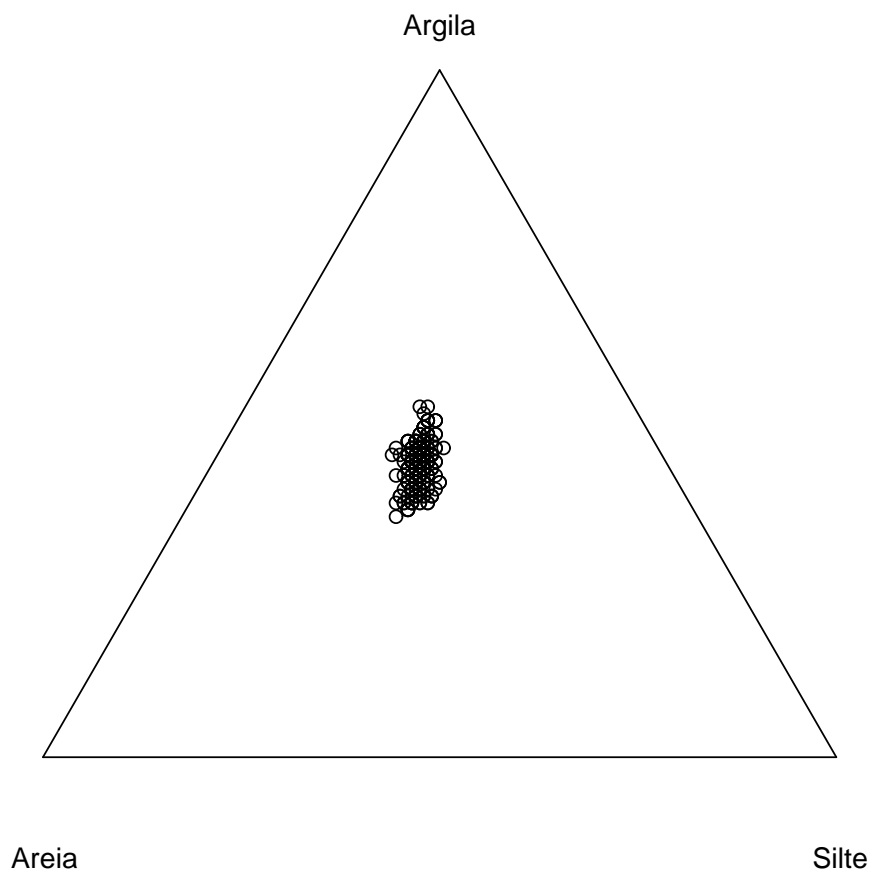


Figura 1: Diagrama Ternário para Areia, Silte e Argila.

```
[1] 31 32 34 32 32 30
```

Construindo uma matriz de diagramas de dispersão

```
> plot(quant)
```

Incluindo a média geométrica e a região 2 e 4-sigma no diagrama ternário

```
> plot(comp)
> plot(mean(comp), add = T, pch = 20, col = "red")
> ellipses(mean(comp), var(comp), col = "red", r = 2)
> ellipses(mean(comp), var(comp), col = "red", r = 4)
```

O centro da distribuição é dado por

```
> mean(comp)
```

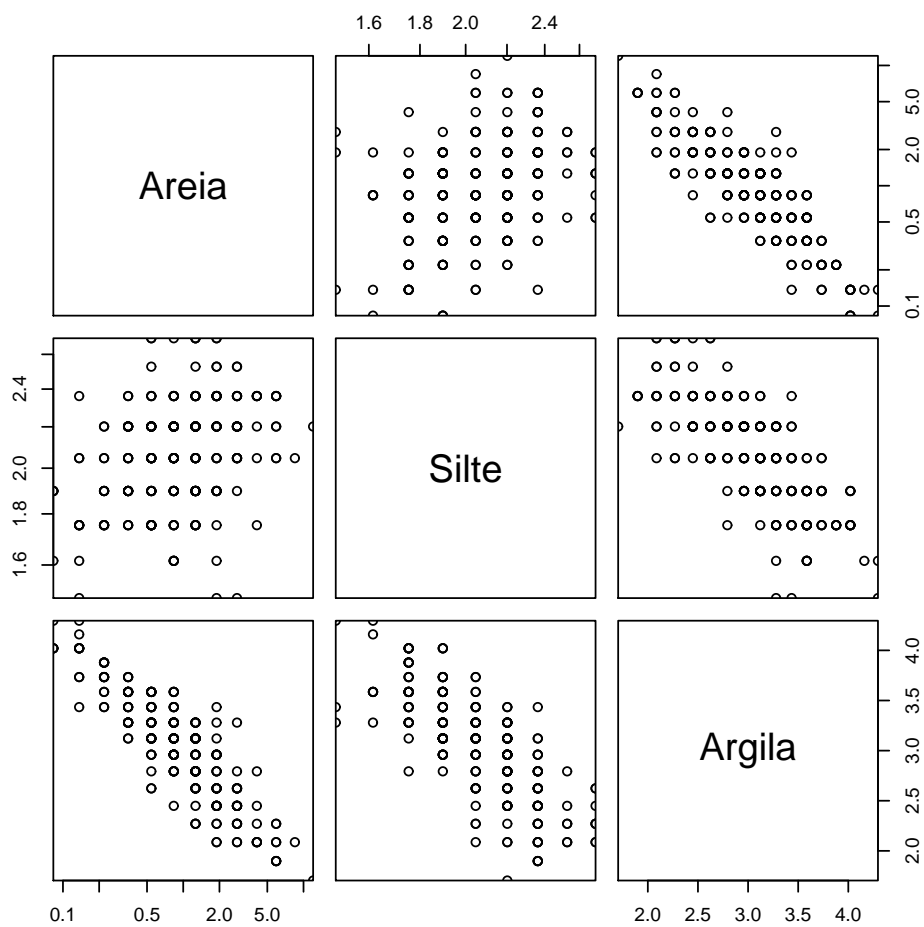


Figura 2: Areia  $\times$  Silte, Areia  $\times$  Argila e Silte  $\times$  Argila.

```

Areia    Silte    Argila
0.3121436 0.2614635 0.4263928
attr(,"class")
[1] "acomp"

```

A variância na estrutura CLR do espaço euclidiano é

```

> var(comp)

      Areia      Silte      Argila
Areia  0.0040091771  0.0003951434 -0.004404321
Silte  0.0003951434  0.0028150763 -0.003210220
Argila -0.0044043205 -0.0032102198  0.007614540

```

Obtém-se a variância métrica que nada mais é do que o traço da matriz CLR, fazendo-se

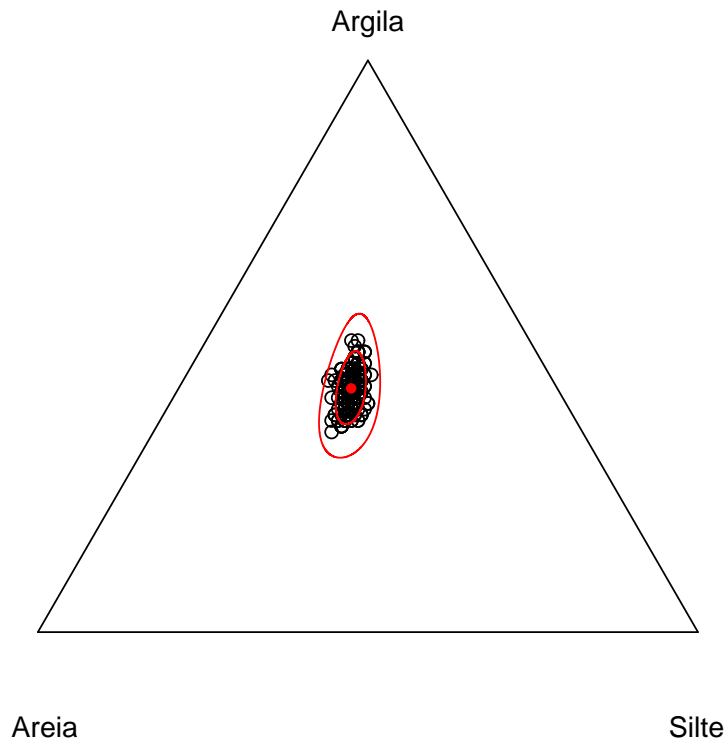


Figura 3: Diagrama ternário com região de confiança.

```
> mvar(comp)
```

```
[1] 0.01443879
```

O desvio padrão métrico -  $\sqrt{\text{mvar}/(D-1)}$  é dado por

```
> msd(comp)
```

```
[1] 0.08496703
```

Já o desvio padrão clássico é obtido através de

```
> sd(comp)
```

Areia	Silte	Argila
0.02244436	0.01646030	0.03205620

Matriz variação -  $\text{var}[\log(x_i/x_j)]$ :

```
> variation(comp)
```

	Areia	Silte	Argila
Areia	0.000000000	0.006033967	0.02043236
Silte	0.006033967	0.000000000	0.01685006
Argila	0.020432358	0.016850056	0.000000000

Por fim, a variância total, a partir da variância na estrutura CLR é obtida. Para isto cria-se um vetor com 3 elementos iguais a 1 e o transposto deste

```
> x <- c(1, 1, 1)
```

```
> x
```

```
[1] 1 1 1
```

```
> xt <- t(x)
```

```
> xt
```

	[,1]	[,2]	[,3]
[1,]	1	1	1

e, então:

```
> options(width = 80)
```

```
> varTot <- mvar(comp) - (1/3) * (xt %*% var(comp) %*% x)
```

```
> varTot
```

	[,1]
[1,]	0.01443879