Análise de componentes principais

Prof. Caio Azevedo

Introdução

- Objetivo: construir variáveis não correlacionadas que retenham a maior parte da estrutura de variabilidade e correlação, a partir da variáveis originais, através de transformações lineares.
- Benefícios:
  - Pode-se usar metodologias de análise univariada.
  - Pode-se trabalhar com um menor número de variáveis.
  - Pode-se obter detalhes do comportamento dos dados os quais são difíceis de serem deduzidos a partir das variáveis originais.

Estrutura estatística

- $\boldsymbol{X} = (X_1, ..., X_p)' \sim D_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, que $D_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ representa uma distribuição p-variada (em princípio qualquer) com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$.
- A matriz de correlações será denotada, como antes, por $\boldsymbol{\rho}$.
- Objetivo, encontrar $\boldsymbol{A}_{(p \times p)}$ (não estocástica), a fim de obter $\boldsymbol{Y} = \boldsymbol{AX} = (Y_1, ..., Y_p)'$, de sorte que $\mathcal{V}(Y_1) \geq \mathcal{V}(Y_2) \geq ...\mathcal{V}(Y_p)$ e que $Corre(Y_i, Y_j) = 0 \ \forall i \neq j, i, j = 1, ..., p$, e que $\mathcal{V}(Y_i)$ seja máximo $i = 1, 2, .., p$.

Cont.

- Seja $\boldsymbol{a}_i'$ a i-ésima linha da matriz $\boldsymbol{A}$. Temos que

$$
\begin{aligned}
Y_1 &= \boldsymbol{a}_1'\boldsymbol{X} = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \\
Y_2 &= \boldsymbol{a}_2'\boldsymbol{X} = a_{21}X_1 + a_{22}X_2 + ... + a_{2p}X_p \\
\vdots &= \vdots \qquad\qquad\qquad \vdots \\
Y_p &= \boldsymbol{a}_p'\boldsymbol{X} = a_{p1}X_1 + a_{p2}X_2 + ... + a_{pp}X_p
\end{aligned}
$$

Cont.

$$\boldsymbol{Y} = \boldsymbol{AX} = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_p' \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

- Portanto $\mathcal{V}(Y_i) = \boldsymbol{a}_i' \boldsymbol{\Sigma} \boldsymbol{a}_i$ e $Cov(Y_i, Y_j) = \boldsymbol{a}_i' \boldsymbol{\Sigma} \boldsymbol{a}_j$.
- Claramente $\mathcal{V}(Y_i)$ pode ser aumentada de modo ilimitado multiplicando-se cada vetor $\boldsymbol{a}_i$ por uma constante positiva (o que não afetaria o fato de que $\boldsymbol{a}_i' \boldsymbol{\Sigma} \boldsymbol{a}_j = 0$). Assim, iremos restringir nossa atenção a vetores com norma unitária ($\boldsymbol{a}_i' \boldsymbol{a}_i = 1$).

Procedimento

- Primeira componente principal: combinação linear $\boldsymbol{a}_1'\boldsymbol{X}$ que maximiza $\mathcal{V}(\boldsymbol{a}_1'\boldsymbol{X})$ sujeito à $\boldsymbol{a}_1'\boldsymbol{a}_1 = 1$.
- Segunda componente principal: combinação linear $\boldsymbol{a}_2'\boldsymbol{X}$ que maximiza $\mathcal{V}(\boldsymbol{a}_2'\boldsymbol{X})$ sujeito à $\boldsymbol{a}_2'\boldsymbol{a}_2 = 1$ e $Cov(\boldsymbol{a}_1'\boldsymbol{X}, \boldsymbol{a}_2'\boldsymbol{X}) = 0$.
- i-ésima componente principal: combinação linear $\boldsymbol{a}_i'\boldsymbol{X}$ que maximiza $\mathcal{V}(\boldsymbol{a}_i'\boldsymbol{X})$ sujeito à $\boldsymbol{a}_i'\boldsymbol{a}_i = 1$ e $Cov(\boldsymbol{a}_i'\boldsymbol{X}, \boldsymbol{a}_k'\boldsymbol{X}) = 0$, para $k < i$.

Resultado

- Sejam $(\lambda_1, \boldsymbol{e}_1)$, $(\lambda_2, \boldsymbol{e}_2)$,...,$(\lambda_p, \boldsymbol{e}_p)$, em que $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$ são os autovalores associados à $\boldsymbol{\Sigma}$, e $\boldsymbol{e}_i$ os respectivos auto-vetores ortonormalizados, ou seja $\boldsymbol{e}_i'\boldsymbol{e}_i = 1$ e $\boldsymbol{e}_i'\boldsymbol{e}_j = 0, \forall i \neq j, i, j = 1, 2, ..., p$.
- A i-ésima componente principal é dada por

$$\boldsymbol{Y}_i = \boldsymbol{e}_i'\boldsymbol{X}, i = 1, 2, ..., p. \tag{1}$$

- Assim, temos que

$$\mathcal{V}(Y_i) = \boldsymbol{e}_i'\boldsymbol{\Sigma}\boldsymbol{e}_i = \lambda_i, i = 1, 2, ..., p \tag{2}$$

$$Cov(Y_i, Y_j) = \boldsymbol{e}_i'\boldsymbol{\Sigma}\boldsymbol{e}_j = 0, \forall i \neq j, i, j = 1, 2, ..., p \tag{3}$$

Prova

- Com relação à (2) e (3), seja $\boldsymbol{E}$ uma matriz em que as linhas são os auto-vetores $\boldsymbol{e}_1, \boldsymbol{e}_2, .., \boldsymbol{e}_p$ (definidos anteriormente), ou seja (ortonormalizados)

$$\boldsymbol{E} = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1p} \\ e_{21} & e_{22} & \dots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{p1} & e_{p2} & \dots & e_{pp} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1' \\ \boldsymbol{e}_2' \\ \vdots \\ \boldsymbol{e}_p' \end{bmatrix}$$

- Vimos, anteriormente, que $\boldsymbol{\Sigma}$ for uma matriz positiva definida então $\boldsymbol{\Sigma} = \boldsymbol{E}'\boldsymbol{\Lambda}\boldsymbol{E}$, em que $\boldsymbol{\Lambda} = diag(\lambda_1, ...\lambda_p)$ (autovalores), e as colunas da matriz $\boldsymbol{E}'$ são formadas pelos respectivos autovetores

- Assim, temos que

$$Cov(\boldsymbol{Y}) = Cov(\boldsymbol{EX}) = \boldsymbol{E}Cov(\boldsymbol{X})\boldsymbol{E}' = \boldsymbol{E\Sigma E}' = \boldsymbol{EE}'\Lambda\boldsymbol{EE}' = \Lambda$$

pois,

$$\boldsymbol{EE}' = \begin{bmatrix} \boldsymbol{e}_1' \\ \boldsymbol{e}_2' \\ \vdots \\ \boldsymbol{e}_p' \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_p \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1'\boldsymbol{e}_1 & \boldsymbol{e}_1'\boldsymbol{e}_2 & \dots & \boldsymbol{e}_1'\boldsymbol{e}_p \\ \boldsymbol{e}_2'\boldsymbol{e}_1 & \boldsymbol{e}_2'\boldsymbol{e}_2 & \dots & \boldsymbol{e}_2'\boldsymbol{e}_p \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{e}_p'\boldsymbol{e}_1 & \boldsymbol{e}_p'\boldsymbol{e}_2 & \dots & \boldsymbol{e}_p'\boldsymbol{e}_p \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \boldsymbol{I}_p$$

Prova

- A prova de (1), vem do que fato de que $max_{\boldsymbol{a}\neq\boldsymbol{0}} \frac{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{a}}$ é obtido quando $\boldsymbol{a} = \boldsymbol{e}_1$ e o $min_{\boldsymbol{a}\neq\boldsymbol{0}} \frac{\boldsymbol{a}'\Sigma\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{a}}$ é obtido quando $\boldsymbol{a} = \boldsymbol{e}_p$ (veja página 80, do livro do Johnson and Wichern (2007), 7^a edição). Como $\boldsymbol{e}_1'\boldsymbol{e}_1 = 1$, temos que $\max_{\boldsymbol{a}\neq\boldsymbol{0}} \frac{\boldsymbol{a}'\Sigma\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{a}} = \max_{\boldsymbol{a}\neq\boldsymbol{0}} \boldsymbol{a}'\Sigma\boldsymbol{a}$ e o mesmo vale para o mínimo.
- Similarmente, $\max_{\boldsymbol{a}\neq\boldsymbol{0}\perp\boldsymbol{e}_1,\boldsymbol{e}_2,..,\boldsymbol{e}_k} \frac{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{a}}$ é obtido quando $\boldsymbol{a} = \boldsymbol{e}_{k+1}$.

Componentes principais e variáveis originais

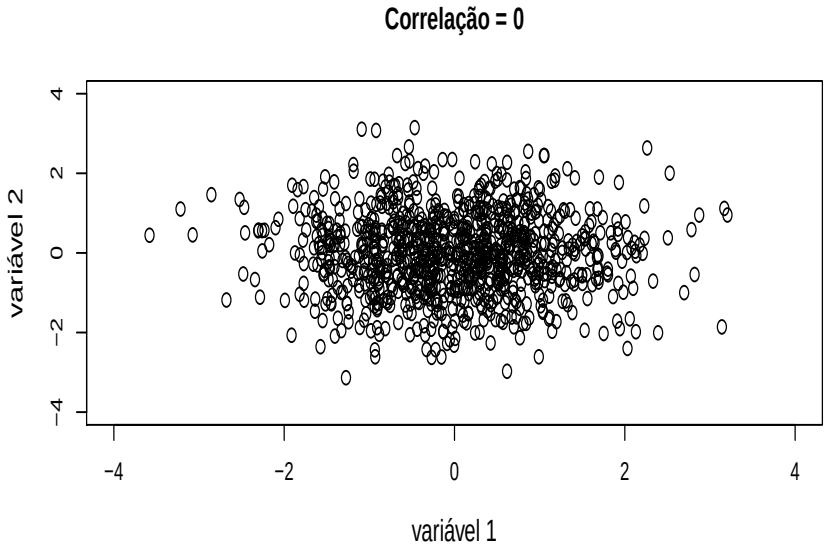
- Defina $\boldsymbol{a}_j = [0\ 0 \ldots \underbrace{1}_{\text{posição } j} \ldots 0\ 0]$
- Temos que $Cov(Y_i, X_j) = Cov(\boldsymbol{e}_i'\boldsymbol{X}, \boldsymbol{a}_j'\boldsymbol{X}) = \boldsymbol{e}_i' Cov(\boldsymbol{X}, \boldsymbol{X})\boldsymbol{a}_j = \boldsymbol{e}_i'\boldsymbol{\Sigma}\boldsymbol{a}_j = \lambda_i \boldsymbol{e}_i'\boldsymbol{a}_j = \lambda_i e_{ij}$. (Provar que $\boldsymbol{e}_i'\boldsymbol{\Sigma} = \lambda_i \boldsymbol{e}_i'$).
- Assim $Corre(Y_i, X_j) = \frac{Cov(Y_i, X_j)}{\sqrt{\mathcal{V}(Y_i)}\sqrt{\mathcal{V}(X_j)}} = \frac{\lambda_i e_{ij}}{\sqrt{\lambda_i}\sigma_j} = \frac{\sqrt{\lambda_i} e_{ij}}{\sigma_j}$, em que $\sigma_j = DP(X_j)$.
- Pode-se mostrar que $tr(\boldsymbol{\Sigma}) = \sum_{i=1}^{p} \mathcal{V}(X_i) = \sum_{i=1}^{p} \lambda_i = \sum_{i=1}^{p} \mathcal{V}(Y_i)$

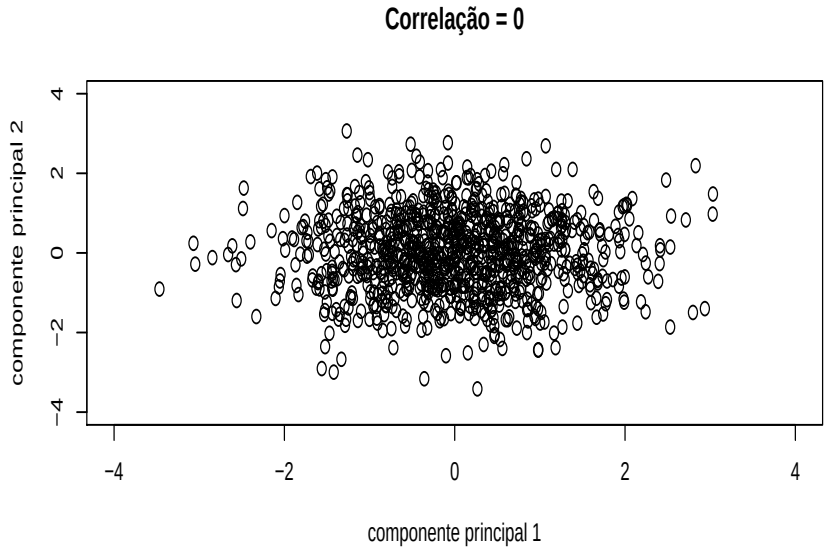
Comentários

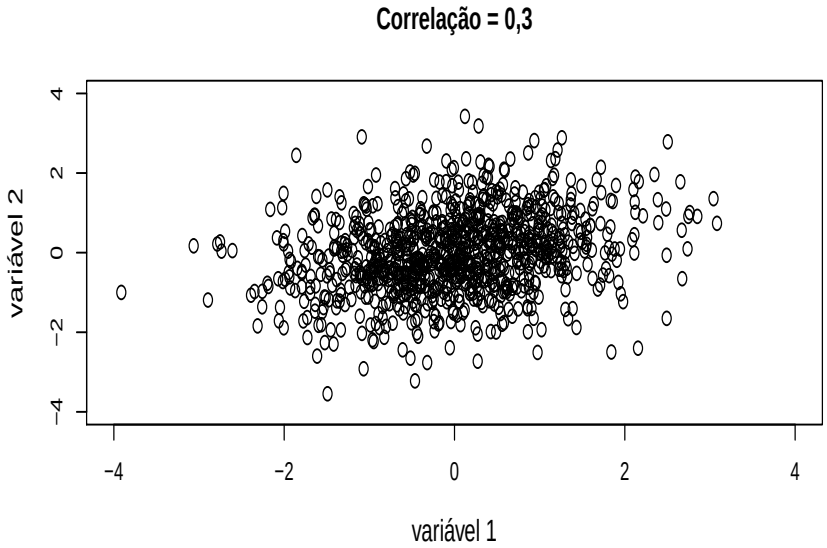
- Há problemas em se trabalhar com a matriz de covariâncias ($\mathbf{\Sigma}$). As componentes principais tendem a ser influenciadas pela variabilidade (escala) das variáveis.
- Quanto maior a variabilidade de uma determinada variável, mais ele influenciará as componentes.
- Alternativa: trabalhar com a matriz de correlações ($\boldsymbol{\rho}$) (o que é equivalente a trabalhar com variáveis com variância unitária).
- Nesse caso as variáveis, portanto as componentes principais, serão adimensionais.

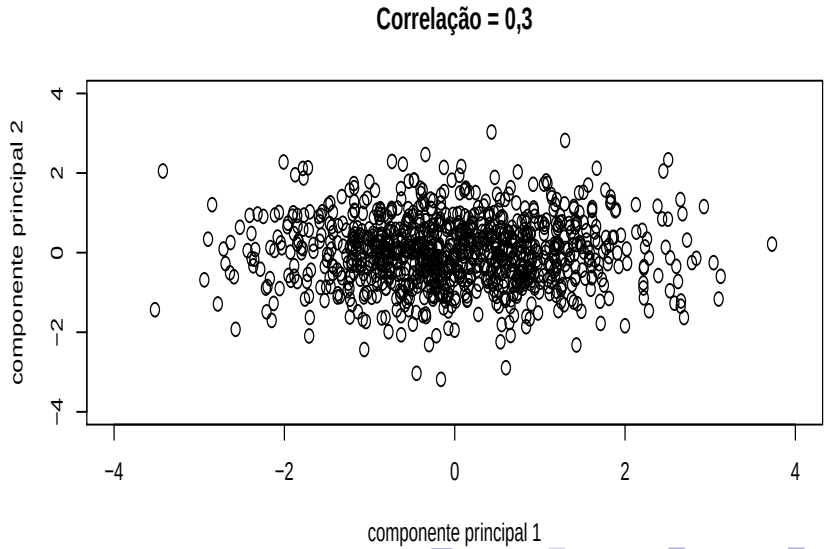
Comentários

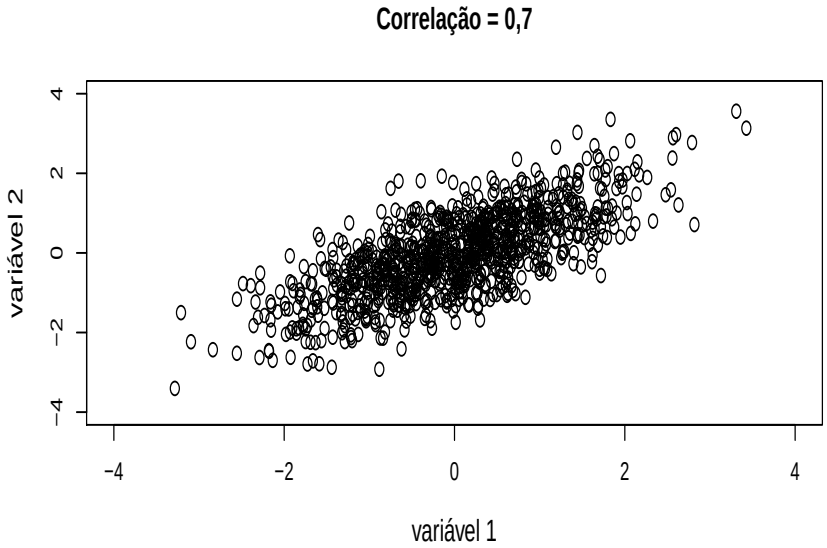
- O procedimento de obtenção das componentes principais continua o mesmo mas, nesse caso, trabalharemos com a matriz de correlações.
- Temos que $Cov(Y_i, X_j) = Cov(\boldsymbol{e}_i'\boldsymbol{X}, \boldsymbol{a}_j'\boldsymbol{X}) = \boldsymbol{e}_i' Cov(\boldsymbol{X}, \boldsymbol{X})\boldsymbol{a}_j = \boldsymbol{e}_i'\boldsymbol{\rho}\boldsymbol{a}_j = \lambda_i \boldsymbol{e}_i'\boldsymbol{a}_j = \lambda_i e_{ij}$. (Provar que $\boldsymbol{e}_i'\boldsymbol{\rho} = \lambda_i \boldsymbol{e}_i'$).
- Além disso, $Corre(Y_i, X_j) = \frac{Cov(Y_i, X_j)}{\sqrt{\mathcal{V}(Y_i)}\sqrt{\mathcal{V}(X_j)}} = \frac{\lambda_i e_{ij}}{\sqrt{\lambda_i}1} = \sqrt{\lambda_i} e_{ij}$.
- Pode-se também trabalhar com as variáveis padronizadas (média 0 e variância 1).
- Como determinar o número de componentes principais e como utilizá-las?

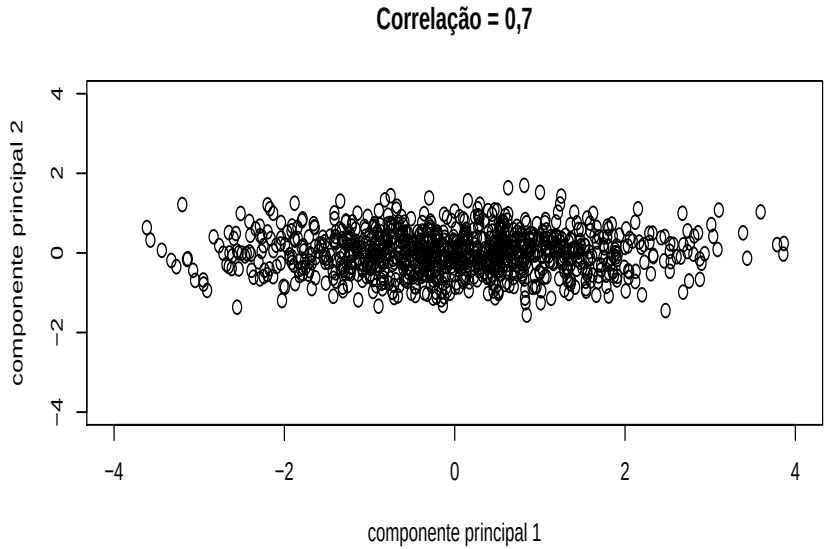

Correlação = 0
variável 2
variável 1
−4
−2
0
2
4

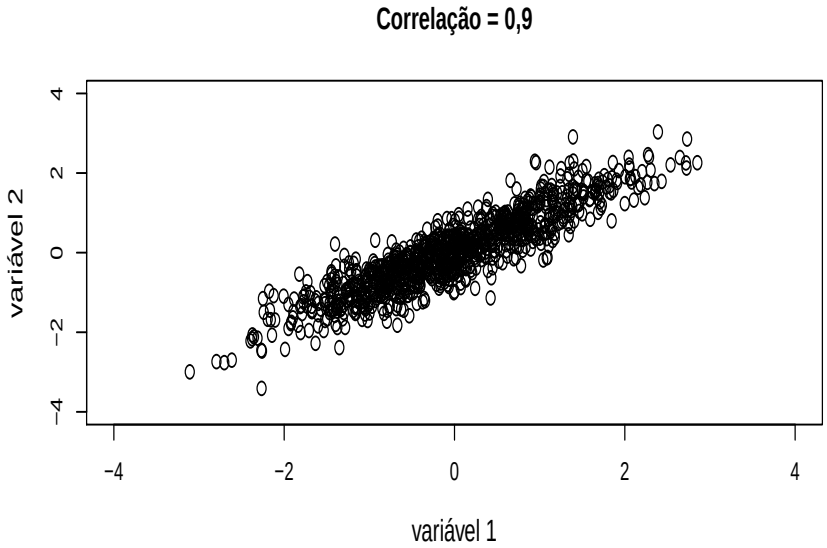

Correlação = 0
componente principal 2
componente principal 1
−4
−2
0
2
4

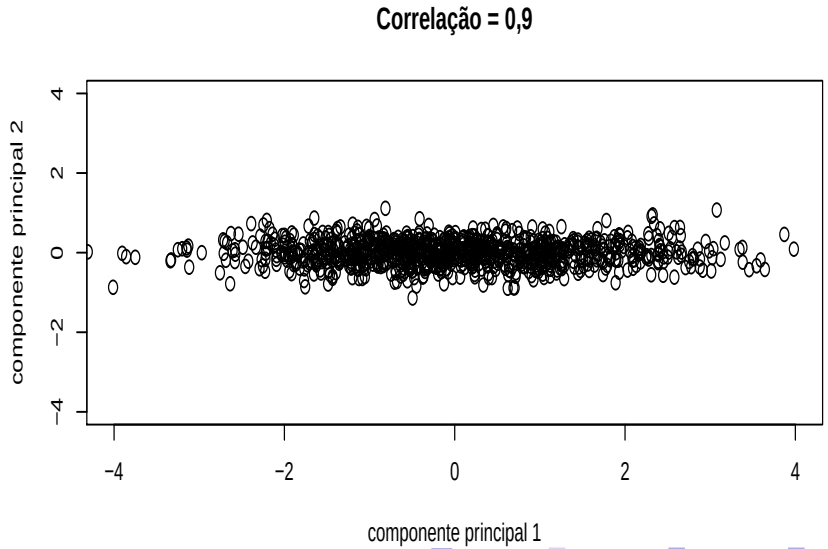

Correlação = 0,3
variável 2
variável 1
−4
−2
0
2
4

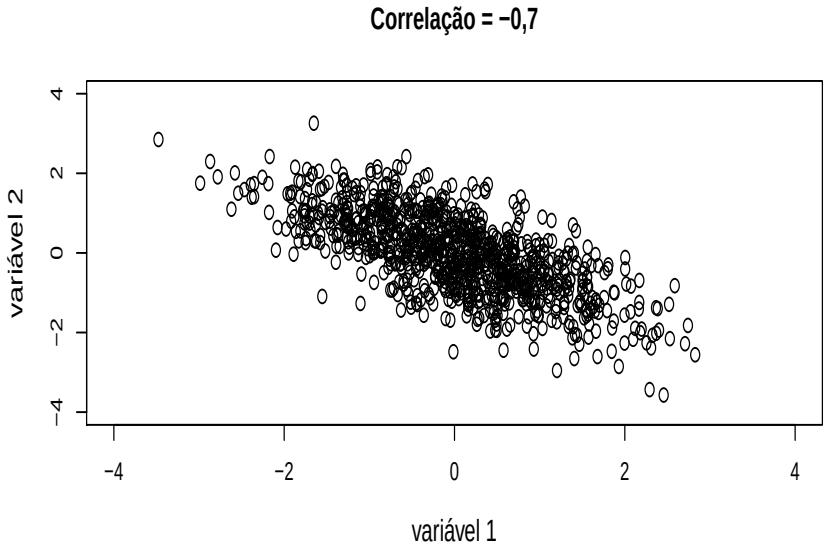

Correlação = 0,3
componente principal 2
componente principal 1
−4
−2
0
2
4

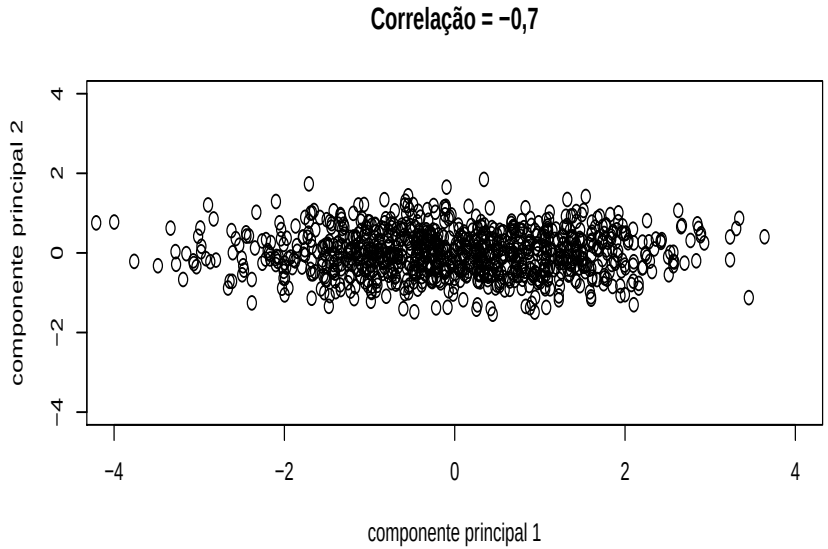

Correlação = 0,7
variável 2
variável 1
−4
−2
0
2
4

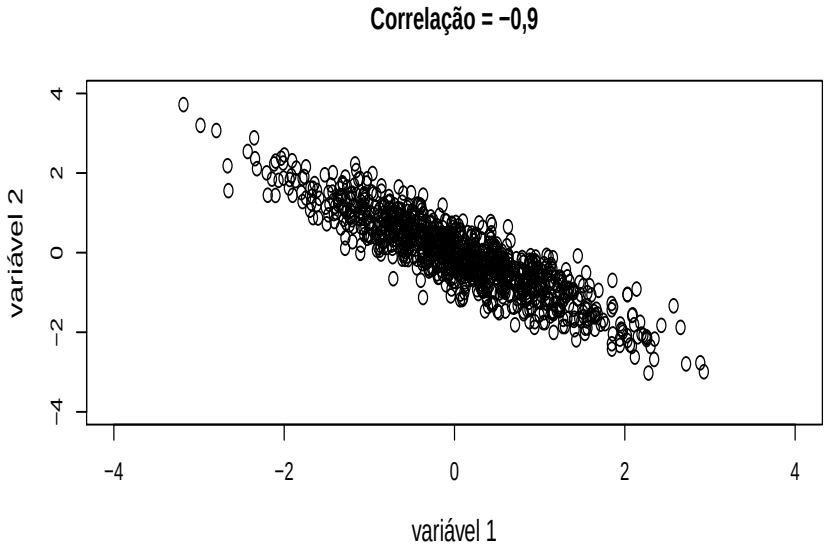

Correlação = 0,7
componente principal 2
componente principal 1
−4
−2
0
2
4

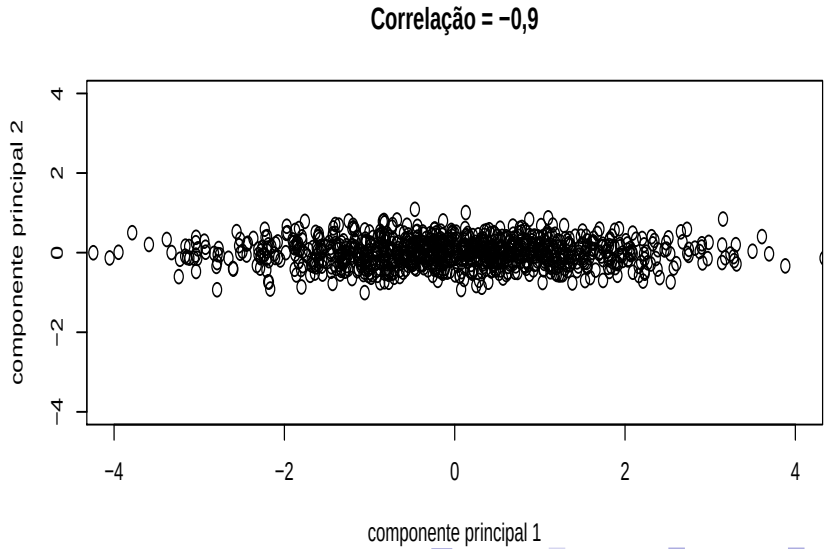

Correlação = 0,9
variável 2
variável 1
−4
−2
0
2
4


Correlação = 0,9
componente principal 2
componente principal 1
−4
−2
0
2
4


Correlação = −0,7
variável 2
variável 1
−4
−2
0
2
4


Correlação = −0,7
componente principal 2
componente principal 1
−4
−2
0
2
4


Correlação = −0,9
variável 2
variável 1
−4
−2
0
2
4


Correlação = −0,9
componente principal 2
componente principal 1
−4
−2
0
2
4

Estimação das componentes principais

- Dada uma matriz de dados $\boldsymbol{X}_{(n\times p)}$

Indivíduo	Variável 1	Variável 2	...	Variável p
1	X_{11}	X_{12}	...	X_{1p}
2	X_{21}	X_{22}	...	X_{2p}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	X_{n1}	X_{n2}	...	X_{np}

estima-se a matriz de covariâncias $(\widetilde{\boldsymbol{\Sigma}})$ ou a matriz de correlações amostrais $(\widetilde{\boldsymbol{\rho}})$, conforme visto anteriormente.

Estimação das componentes principais

- Obtem-se os auto-valores e auto-vetores ortonormalizados, digamos $(\widetilde{\lambda}_1, \widetilde{\boldsymbol{e}}_1)$, $(\widetilde{\lambda}_2, \widetilde{\boldsymbol{e}}_2)$,...,$(\widetilde{\lambda}_p, \widetilde{\boldsymbol{e}}_p)$ a partir de $\widetilde{\boldsymbol{\Sigma}}$ ou $\widetilde{\boldsymbol{\rho}}$.
- Assim $\widetilde{\mathcal{V}}(Y_{ij}) = \widetilde{\lambda}_i$.
- Cada componente principal será calculada como $y_{ij} = \widetilde{\boldsymbol{e}}_j \boldsymbol{x}_i$, em que $\boldsymbol{x}_i, i = 1, .., n, j = 1, ..., p$ é a i-ésima linha da matriz de dados observada.
- Alternativamente, podemos estimar as componentes através de $y_{ij} = \widetilde{\boldsymbol{e}}_j \boldsymbol{z}_i$, em que $\boldsymbol{z}_i$ corresponde às variáveis padronizadas (default do R).

Exemplo 4: índices de criminalidade de estados americanos

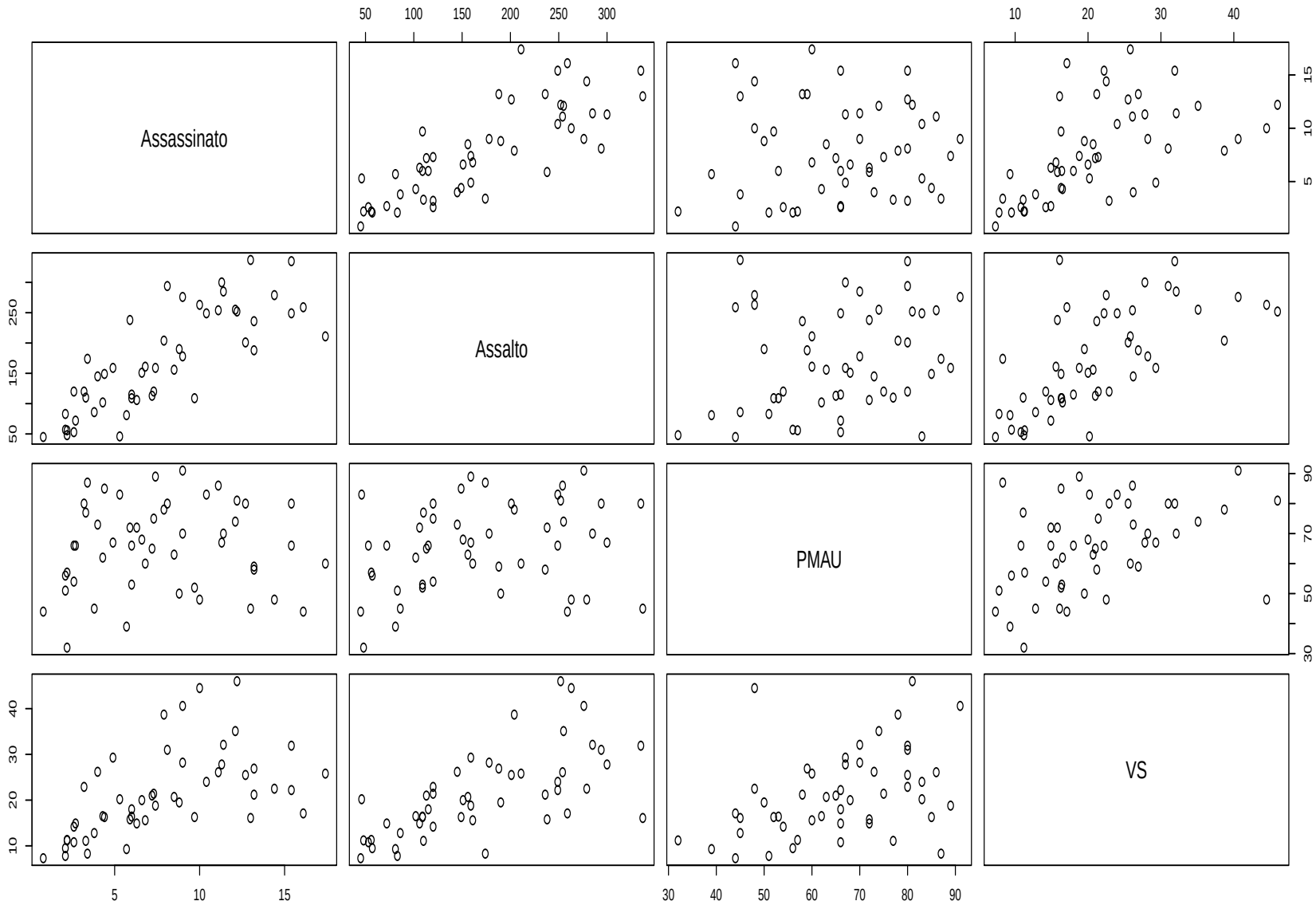
- Número de prisões efetuadas (por 100.000 habitantes) em 50 estados americanos em 1973.
- Variáveis: assalto, violência sexual (VS), assassinato e porcentagem de moradores na área urbana (PMAU).

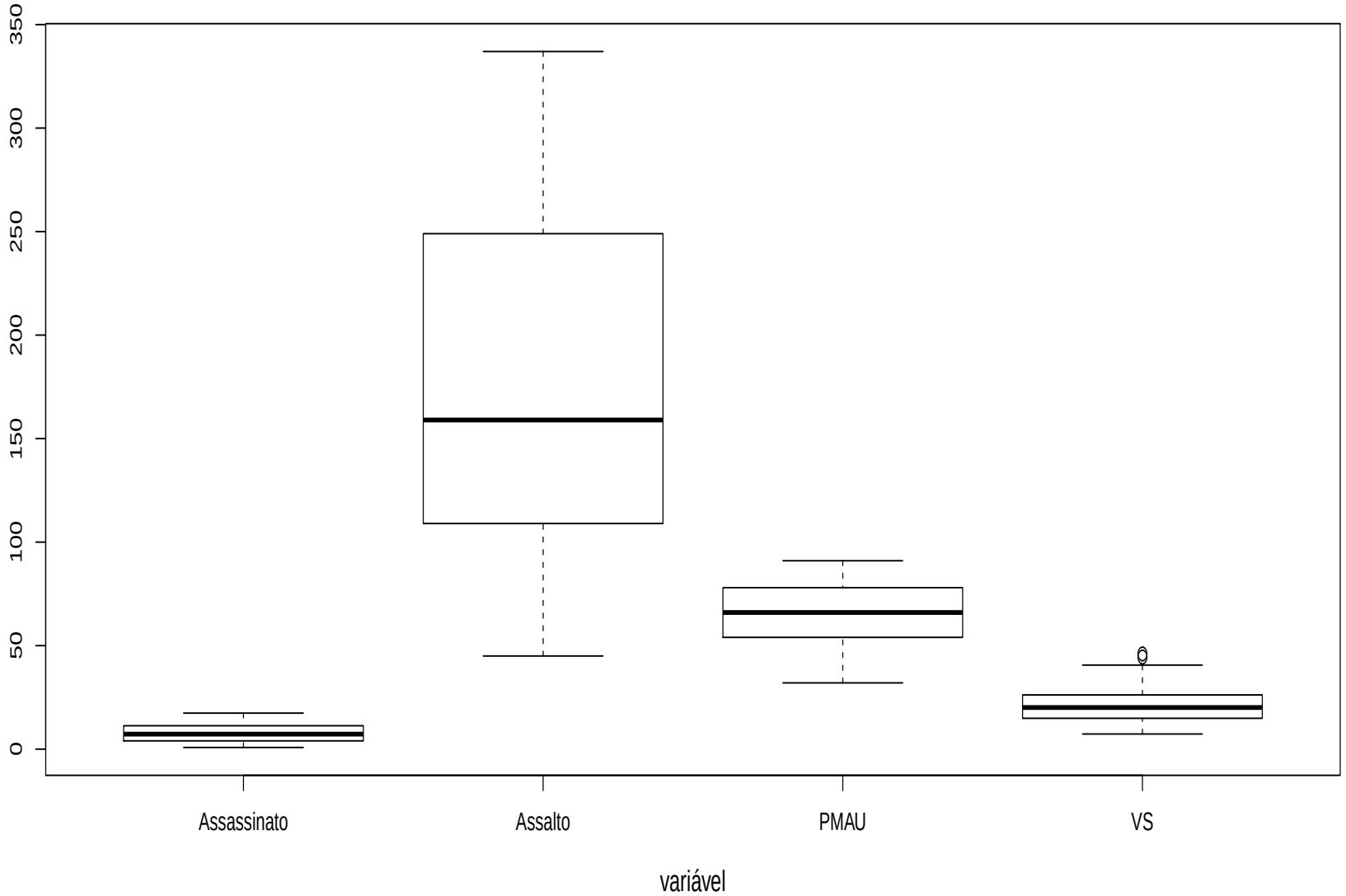
Estado	**Assassinato**	**Assalto**	**PMAU**	**VS**
Alabama	13,2	236,0	58	21,2
Alaska	10,0	263,0	48	44,5
Arizona	8,1	294,0	80	31,0
⋮	⋮	⋮	⋮	⋮

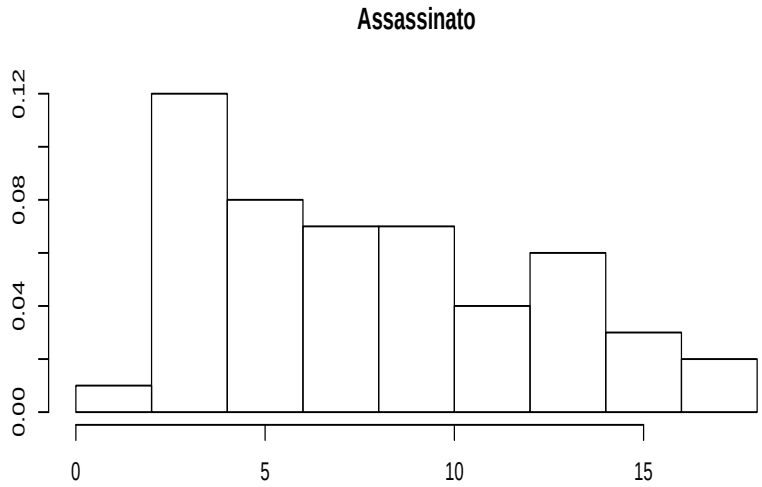
- Digitar *USArrests* no programa *R*.

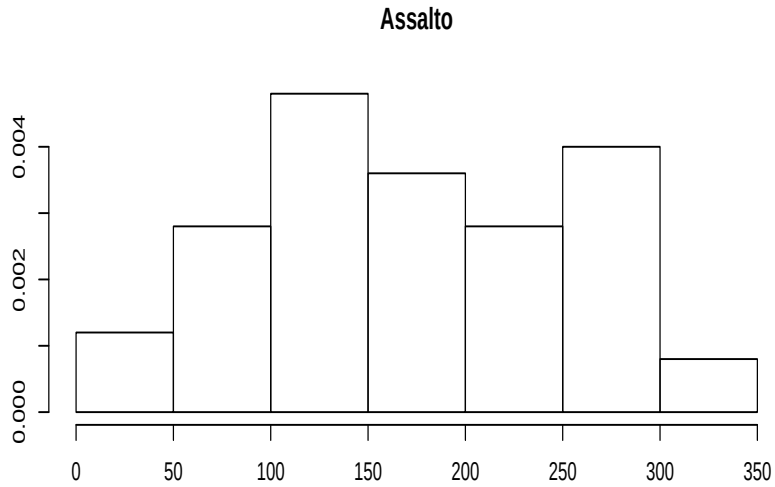
- Medidas descritivas

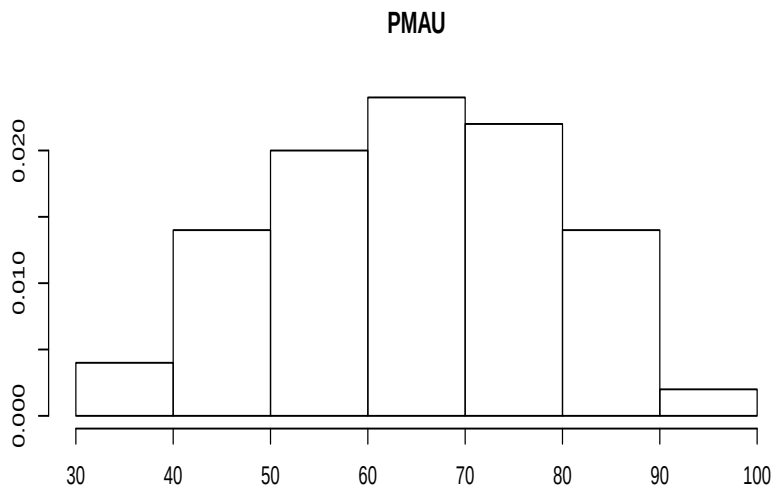
Estatística	**Assassinato**	**Assalto**	**PMAU**	**VS**
Média	7,79	170,76	65,54	21,23
Var.	18,97	**6.945,17**	209,52	87,73
DP	4,36	83,34	14,47	9,37
CV(%)	55,93	48,80	22,09	44,11
Mínimo	0,80	45,00	32,00	7,30
Mediana	7,25	159,00	66,00	20,10
Máximo	17,40	337,00	91,00	46,00

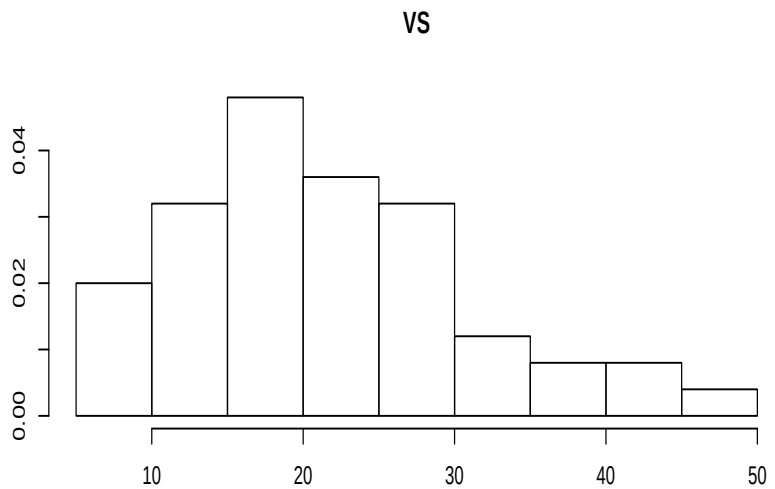

Assassinato
Assalto
PMAU
VS
50
100
150
200
250
300
10
20
30
40
5
10
15
50
150
250
10
20
30
40
30
50
70
90
5
10
15
30
40
50
60
70
80
90

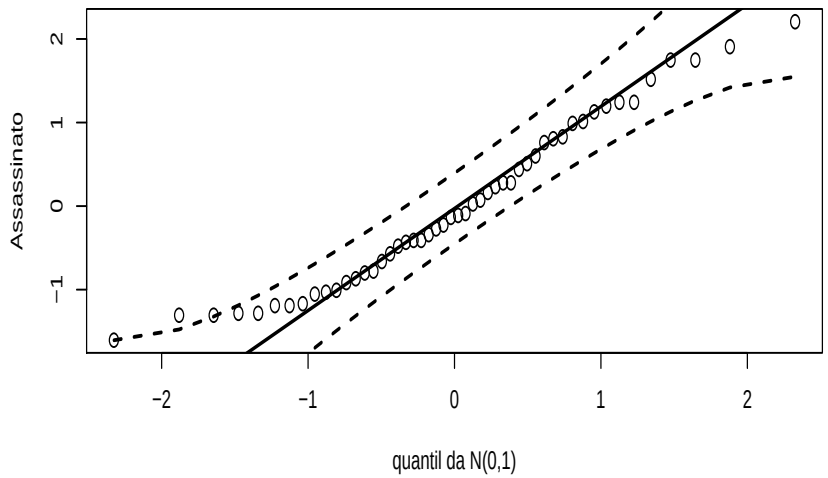

0
50
100
150
200
250
300
350
Assassinato
Assalto
PMAU
VS
variável

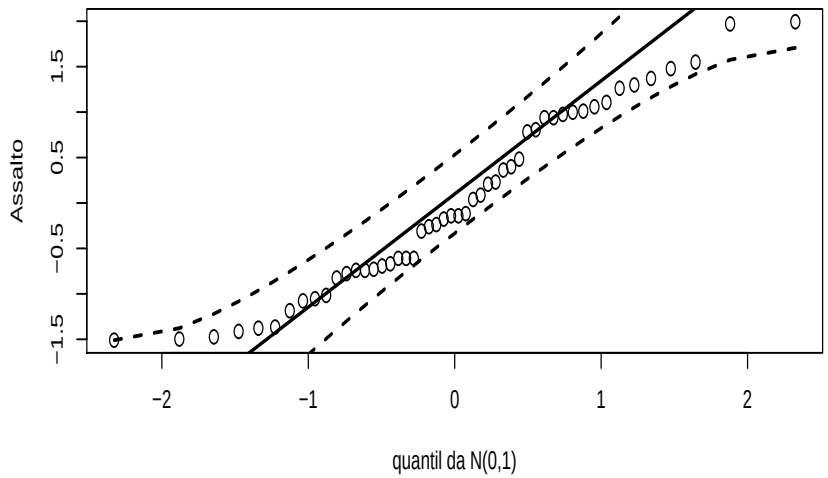

Assassinato
0.00
0.04
0.08
0.12
0
5
10
15

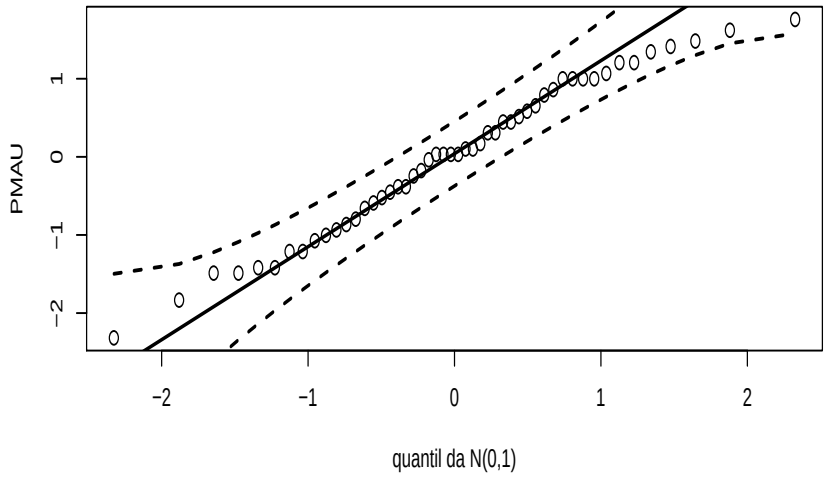

Assalto
0.000
0.002
0.004
0
50
100
150
200
250
300
350

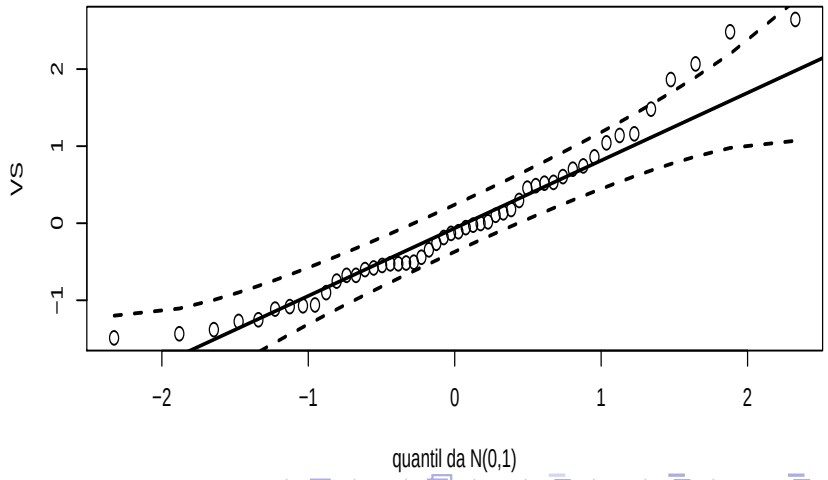

PMAU
0.000
0.010
0.020
30
40
50
60
70
80
90
100


VS
0.00
0.02
0.04
10
20
30
40
50


Assassinato
quantil da N(0,1)


Assalto
quantil da N(0,1)


PMAU
quantil da N(0,1)


VS
quantil da N(0,1)

- Matriz de covariâncias

	Assassinato	**Assalto**	**PMAU**	**VS**
Assassinato	18,97	291,06	4,38	22,99
Assalto	291,06	**6945,16**	312,27	519,26
PMAU	4,38	312,27	209,51	55,76
VS	22,99	519,26	55,76	87,72

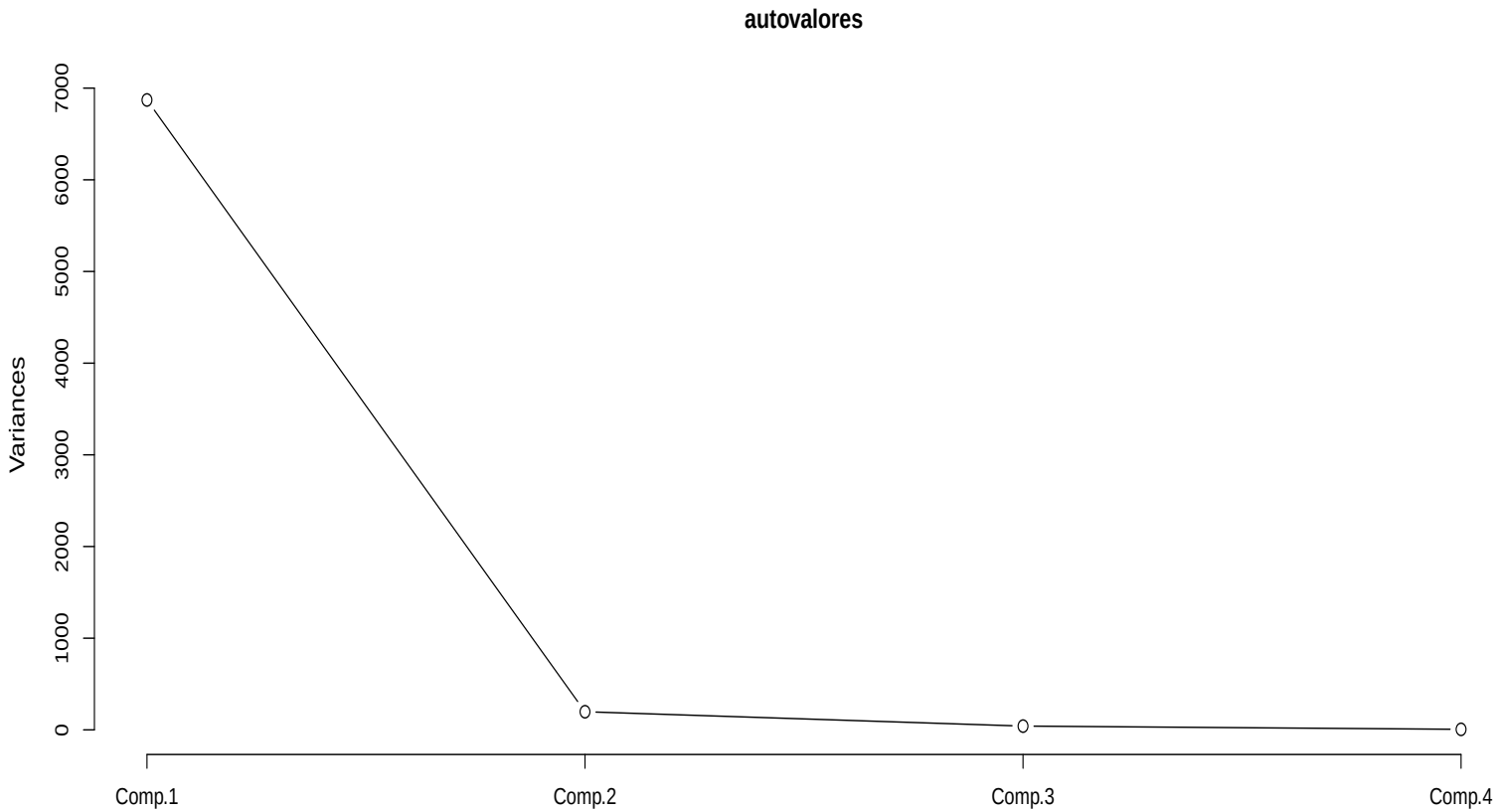
- Autovalores

  7011,11 201,99 42,11 6,16

- Autovetores

-0,04170432	0,04482166	0,07989066	0,99492173
-0,99522128	0,05876003	-0,06756974	-0,03893830
-0,04633575	-0,97685748	-0,20054629	0,05816914
-0,07515550	-0,20071807	0,97408059	-0,07232502

0,04170432	-0,04482166	-0,07989066	-0,99492173
0,99522128	-0,05876003	0,06756974	0,03893830
0,04633575	0,97685748	0,20054629	-0,05816914
0,07515550	0,20071807	-0,97408059	0,07232502


autovalores
Variances
0
1000
2000
3000
4000
5000
6000
7000
Comp.1
Comp.2
Comp.3
Comp.4

- Variância explicada

	Comp. 1	**Comp. 2**	**Comp. 3**	**Comp. 4**
PVE (%)	96,55	2,78	< 0.01	< 0.01
PVEA (%)	96,55	99,33	99,91	100,00

- Componentes principais

	Comp. 1	**Comp. 2**
Assassinato	0,04 (0,80)	-0,04 (-0,14)
Assalto	0,99 (0,99)	-0,06 (<-0,01)
PMAU	0,04 (0,26)	0,98 (0,96)
VS	0,07 (0,67)	0,20 (0,30)

- Utilizar a matriz de correlações.

- Matriz de correlações

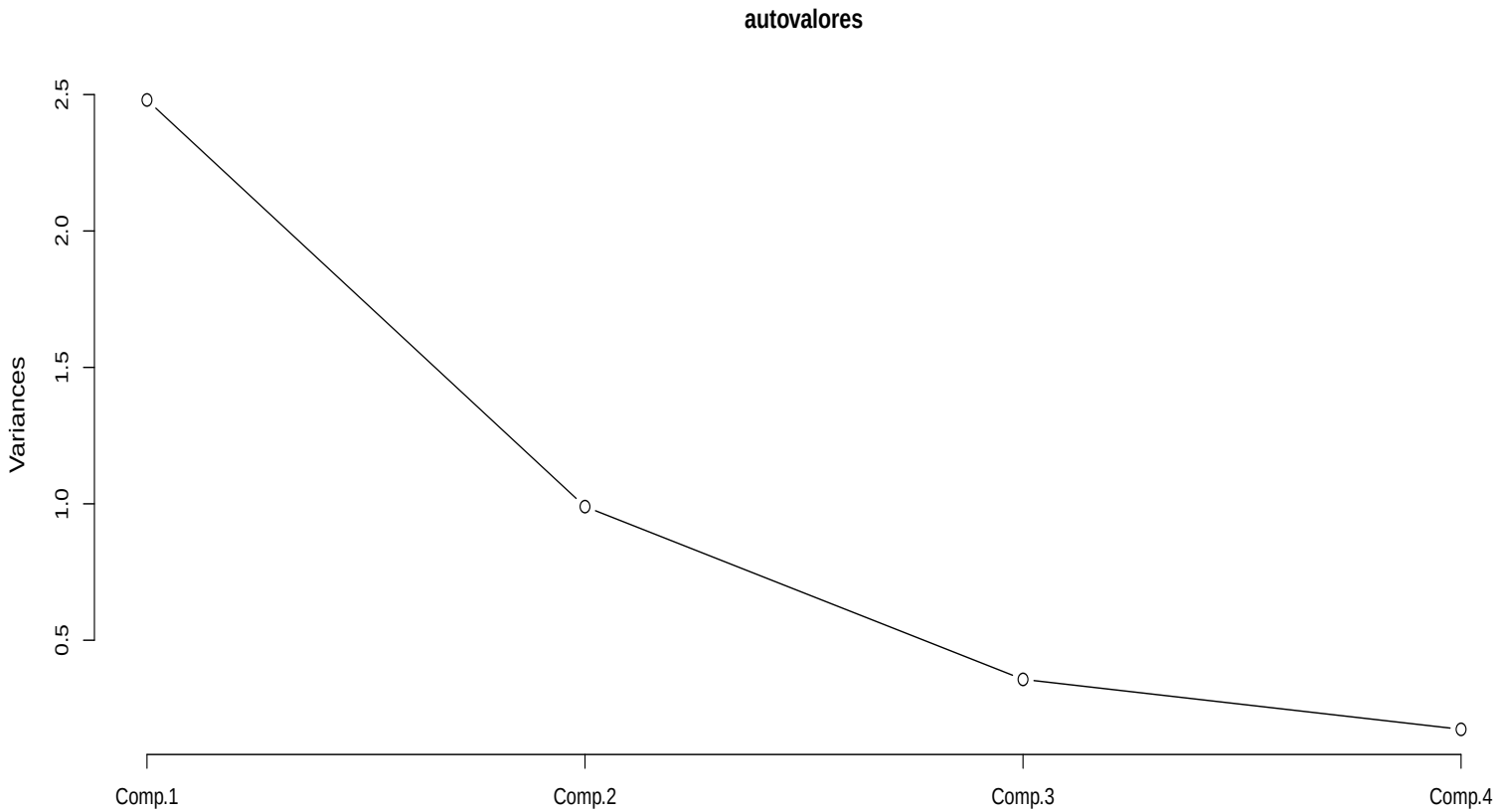
	Assassinato	**Assalto**	**PMAU**	**VS**
Assassinato	1,00	0,80	0,06	0,56
Assalto	0,80	1,00	0,25	0,66
PMAU	0,07	0,25	1,00	0,41
VS	0,56	0,66	0,41	1,00

- Autovalores

  2,48 0,98 0,35 0,17

- Autovetores

0,5358995	-0,4181809	0,3412327	-0,64922780
0,5831836	-0,1879856	0,2681484	0,74340748
0,2781909	0,8728062	0,3780158	-0,13387773
0,5434321	0,1673186	-0,8177779	-0,08902432


autovalores
Variances
0.5
1.0
1.5
2.0
2.5
Comp.1
Comp.2
Comp.3
Comp.4

- Variância explicada

	Comp. 1	**Comp. 2**	**Comp. 3**	**Comp. 4**
PVE (%)	62,00	24,74	8,91	4,32
PVEA (%)	62,00	86,75	95,67	100,00

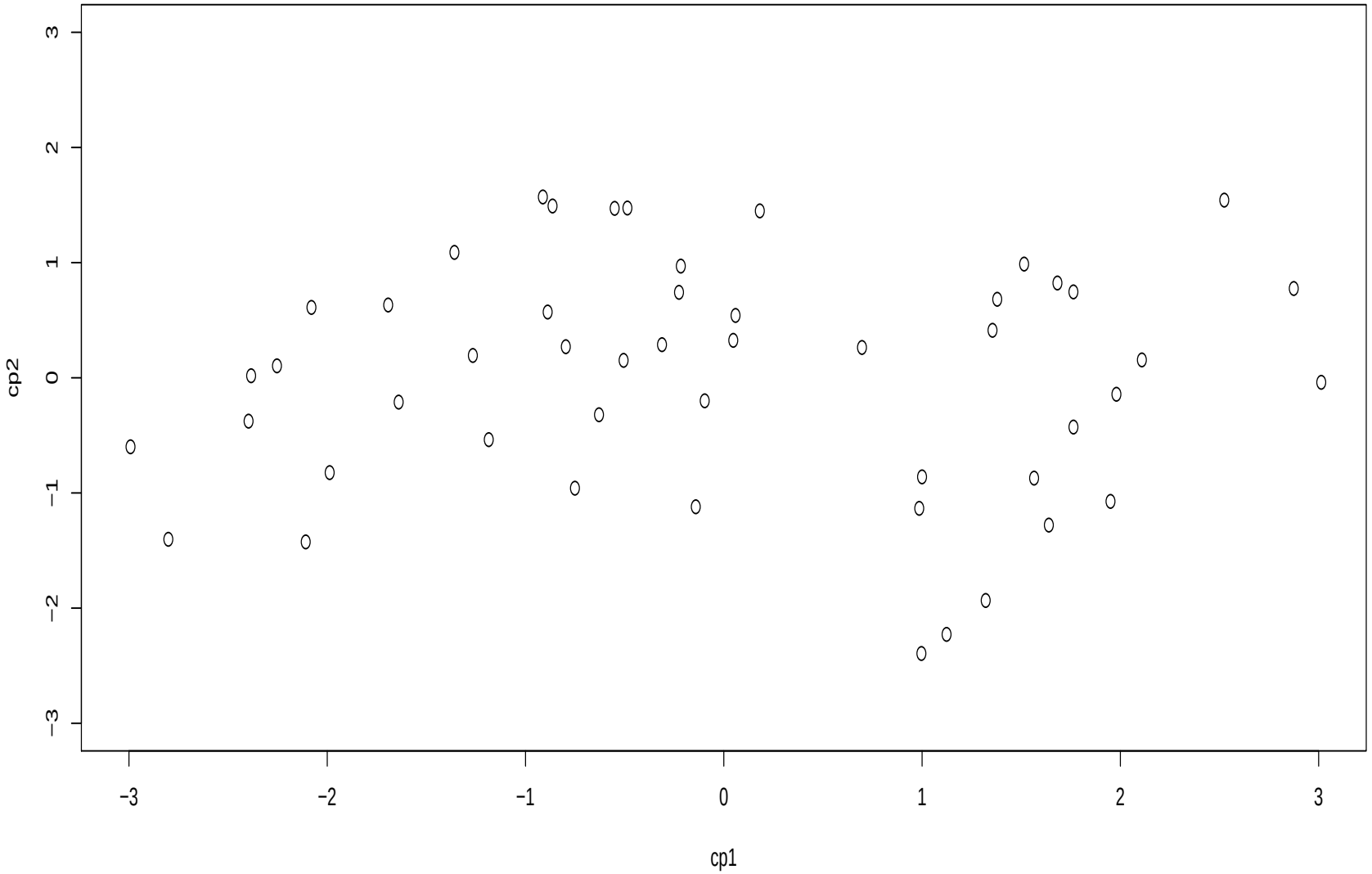
- Componentes principais

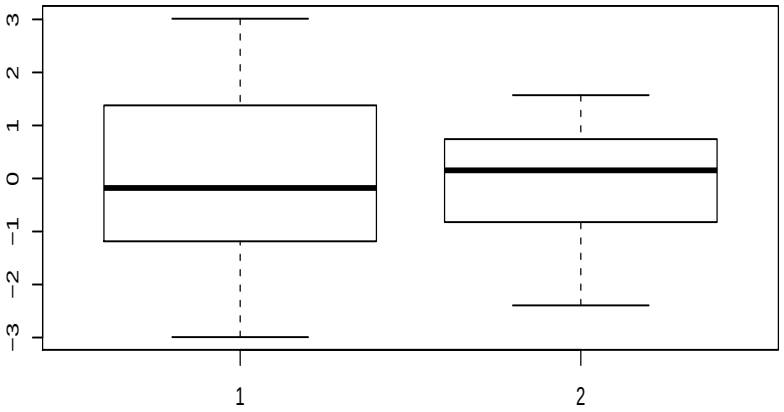
	Comp. 1	**Comp. 2**
Assassinato	0,53 (0,84)	-0,41(-0,41)
Assalto	0,58 (0,91)	-0,18(-0,19)
PMAU	0,27 (0,43)	0,86(0,87)
VS	0,54 (0,85)	0,16(0,17)

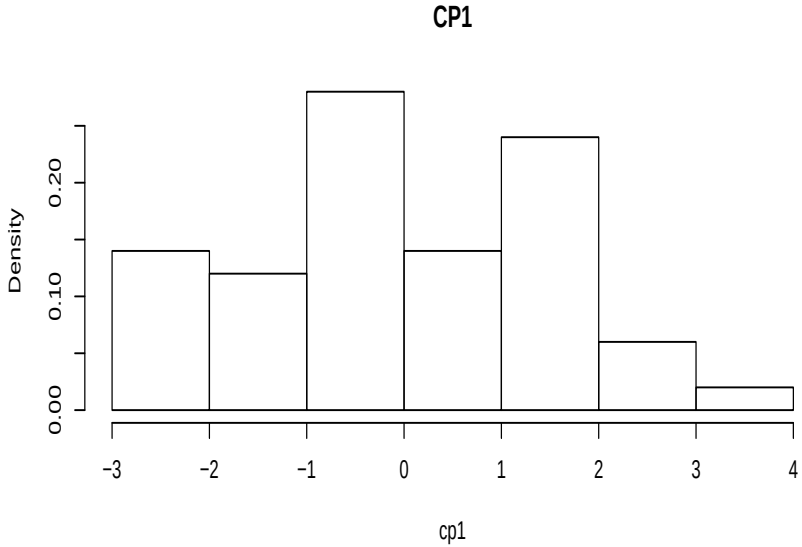
- Primeira componente: “escore ponderado” entre as variáveis, com maiores pesos para as variáveis criminais.
- Segunda componente: é um constraste entre PMAU e VS e as outras variáveis criminais, com maior peso para a variável PMAU.

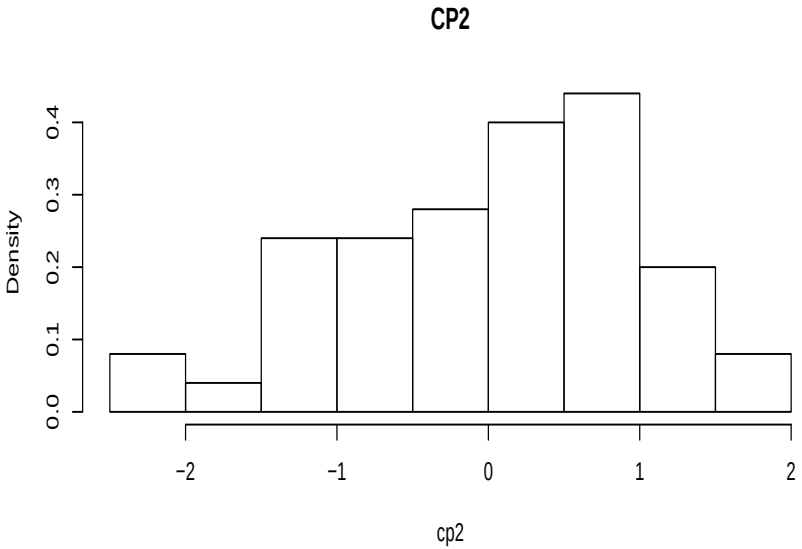
- Função “princomp”. Resultado salvo num objeto chamado result.

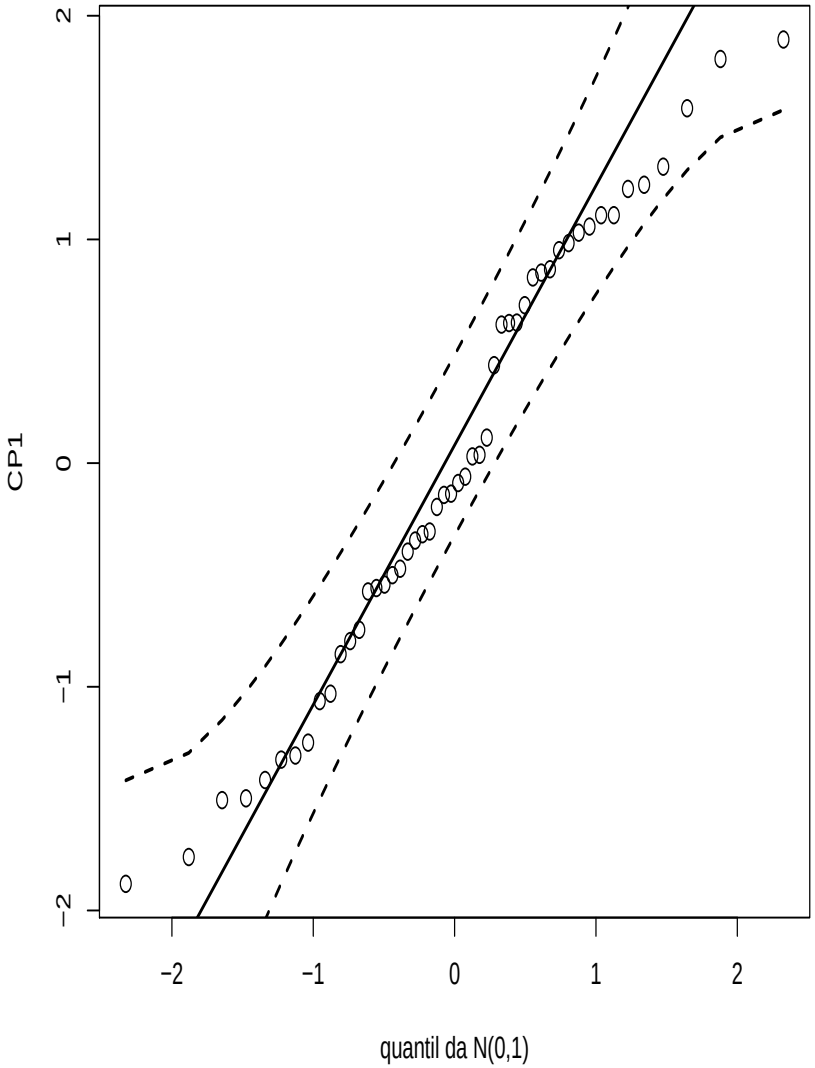
Comando	**Significado**
result$loadings	coeficientes que geram as CP's (geralmente os autovetores negativos)
result$scores	estimativa das componentes principais
screeplot(result)	screeplot

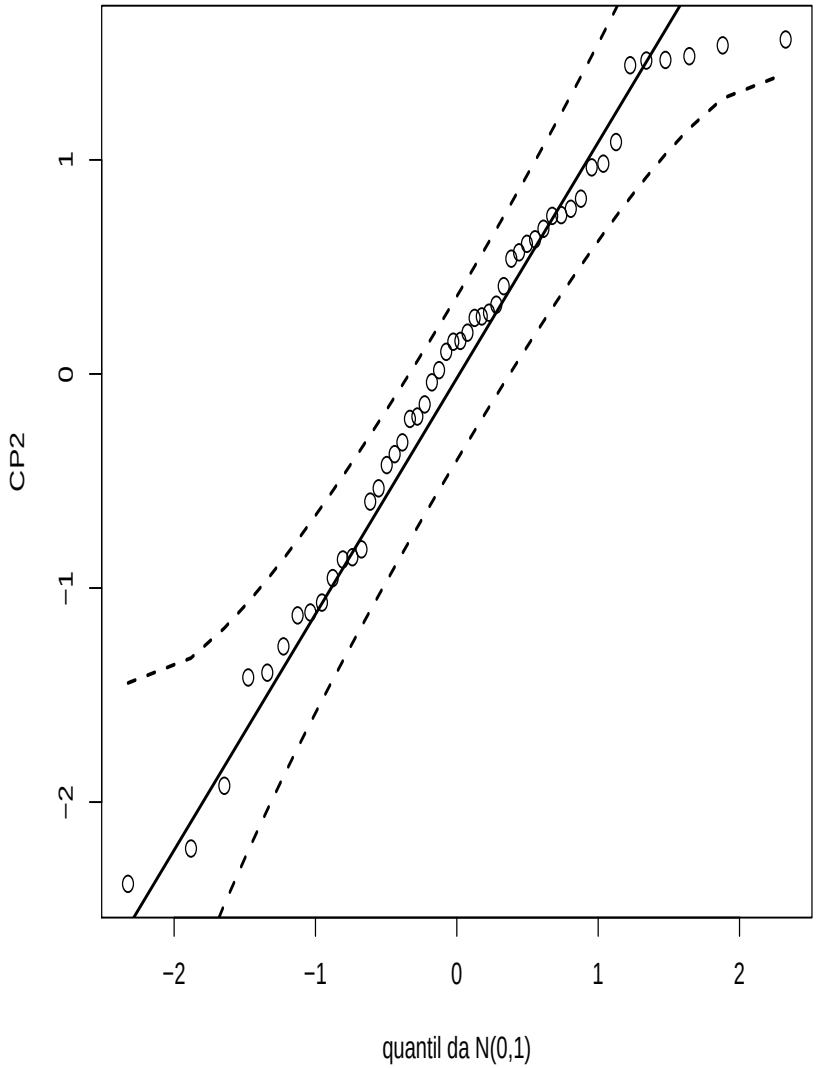

cp2
cp1
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3

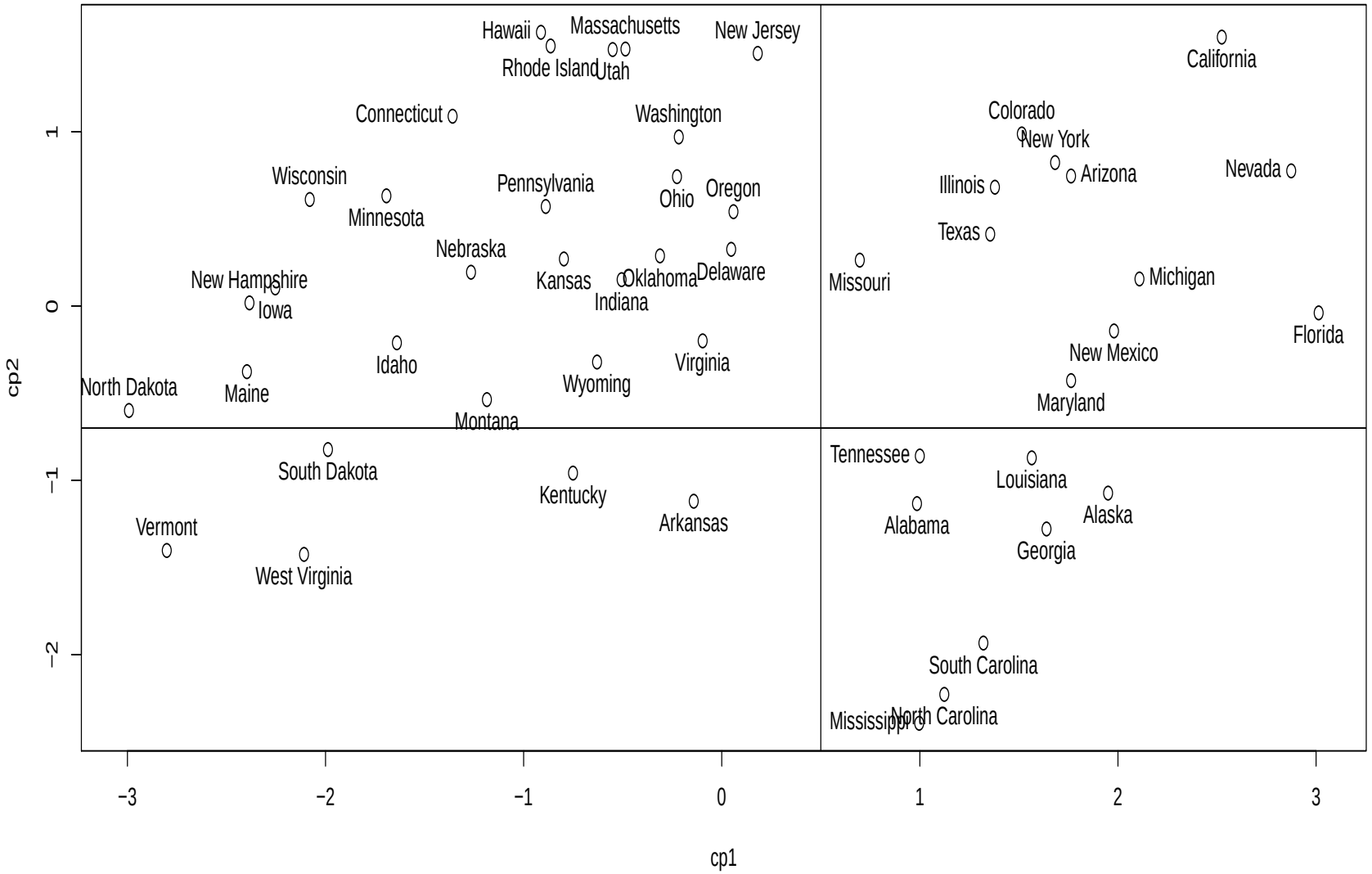

3
2
1
0
−1
−2
−3
1
2
CP

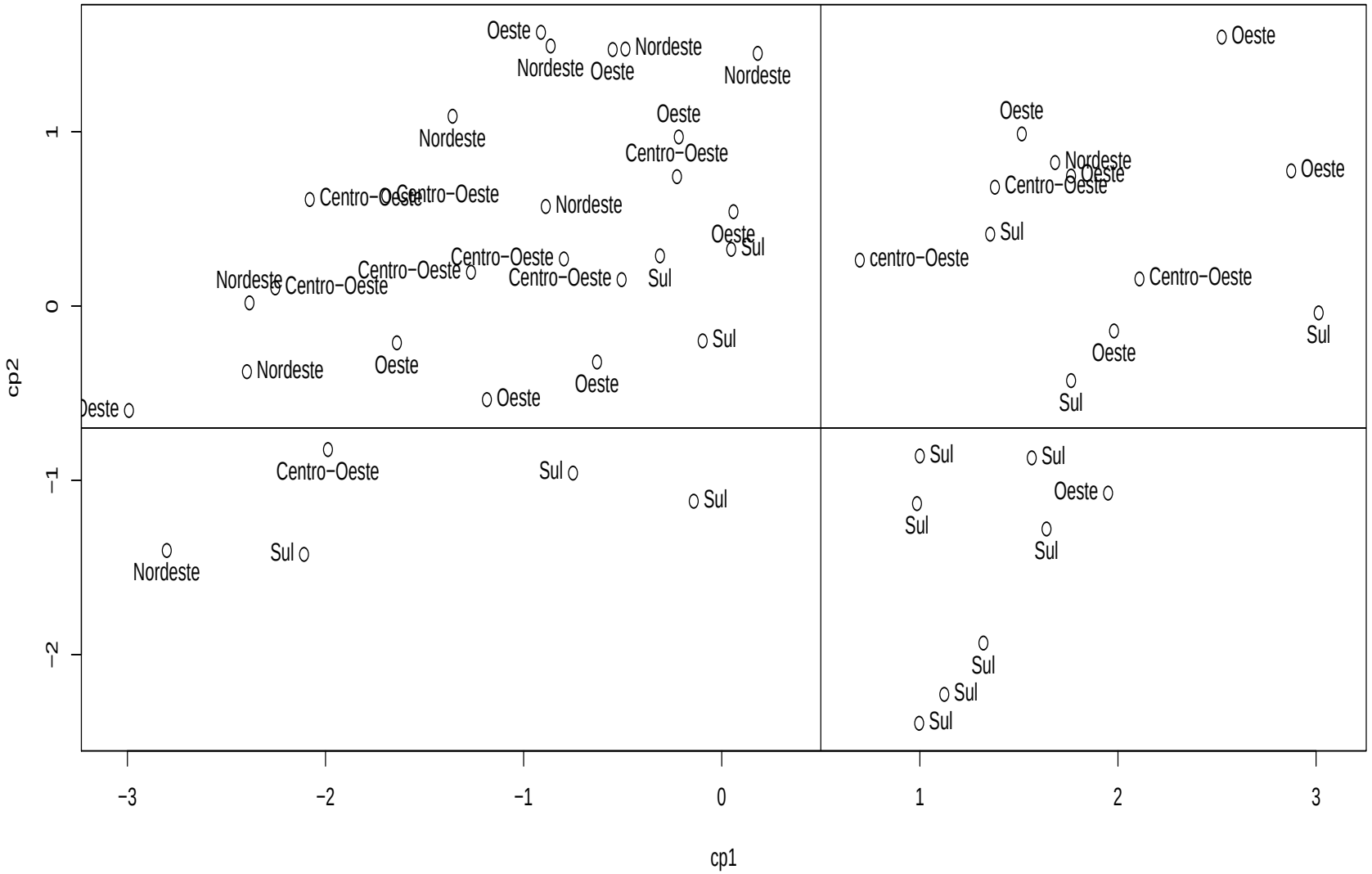

CP1
Density
0.00
0.10
0.20
−3
−2
−1
0
1
2
3
4
cp1

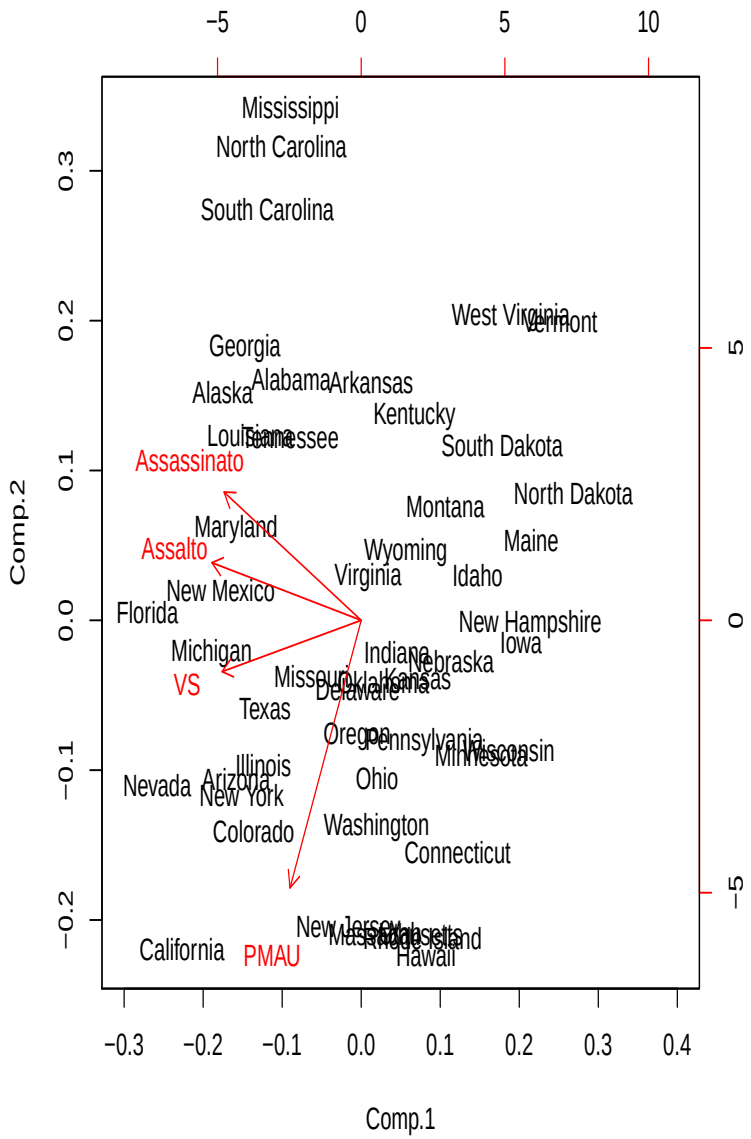

CP2
Density
0.0
0.1
0.2
0.3
0.4
−2
−1
0
1
2
cp2


CP1
quantil da N(0,1)


CP2
quantil da N(0,1)


Hawaii
Massachusetts
New Jersey
Rhode Island
Utah
California
Connecticut
Washington
Colorado
New York
Wisconsin
Pennsylvania
Ohio
Oregon
Arizona
Nevada
Illinois
Minnesota
Texas
Nebraska
Missouri
Michigan
New Hampshire
Kansas
Oklahoma
Delaware
Iowa
Indiana
Florida
Idaho
Virginia
New Mexico
North Dakota
Maine
Wyoming
Maryland
Montana
South Dakota
Tennessee
Louisiana
Kentucky
Alaska
Vermont
Alabama
Arkansas
Georgia
West Virginia
South Carolina
Mississippi
North Carolina
cp1
cp2
−3
−2
−1
0
1
2
3
1
0
−1
−2


Oeste
Nordeste
Centro−Oeste
Sul
centro−Oeste
cp1
cp2
−3
−2
−1
0
1
2
3
−2
−1
0
1


−5
0
5
10
Mississippi
North Carolina
South Carolina
West Virginia
Vermont
Georgia
Alabama
Arkansas
Alaska
Kentucky
Louisiana
Tennessee
South Dakota
Assassinato
North Dakota
Montana
Maryland
Maine
Assalto
Wyoming
Virginia
Idaho
New Mexico
Florida
New Hampshire
Iowa
Michigan
Indiana
Nebraska
VS
Missouri
Kansas
Delaware
Oklahoma
Texas
Oregon
Pennsylvania
Wisconsin
Minnesota
Illinois
Nevada
Arizona
Ohio
New York
Colorado
Washington
Connecticut
New Jersey
Massachusetts
Rhode Island
Hawaii
California
PMAU
0.3
0.2
0.1
0.0
−0.1
−0.2
Comp.2
5
0
−5
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
Comp.1

Biplot

- Um gráfico com quatro eixos.
- Os pontos representam as observações enquanto que os vetores (setas) representam as variáveis.
- Os eixos da parte de baixo e da esquerda se referem aos valores das duas componentes principais divididos pelos respectivos desvios-padrão (das componentes) vezes a raiz quadrada do tamanho da amostra ($y_{ij}^* = \frac{y_{ij}}{\sqrt{\lambda_i n}}$).
- Os eixos das partes de cima e da direita representam os valores dos coeficientes das componentes multiplicados pelos respectivos desvios-padrão (das componentes) vezes a raiz quadrada do tamanho da amostra ($e_{ij}^* = e_{ij}\sqrt{\lambda_i n}$).