Sensoriamento Remoto II

Componentes principais

Revisão de matemática

Análise de componentes principais em SR

UFPR – Departamento de Geomática
Prof. Jorge Centeno
2025

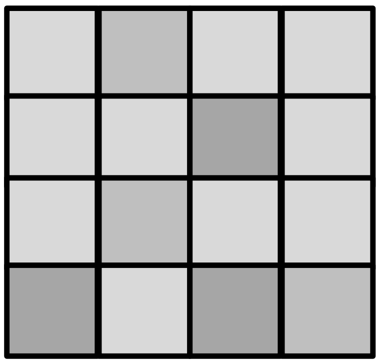
Revisão matemática

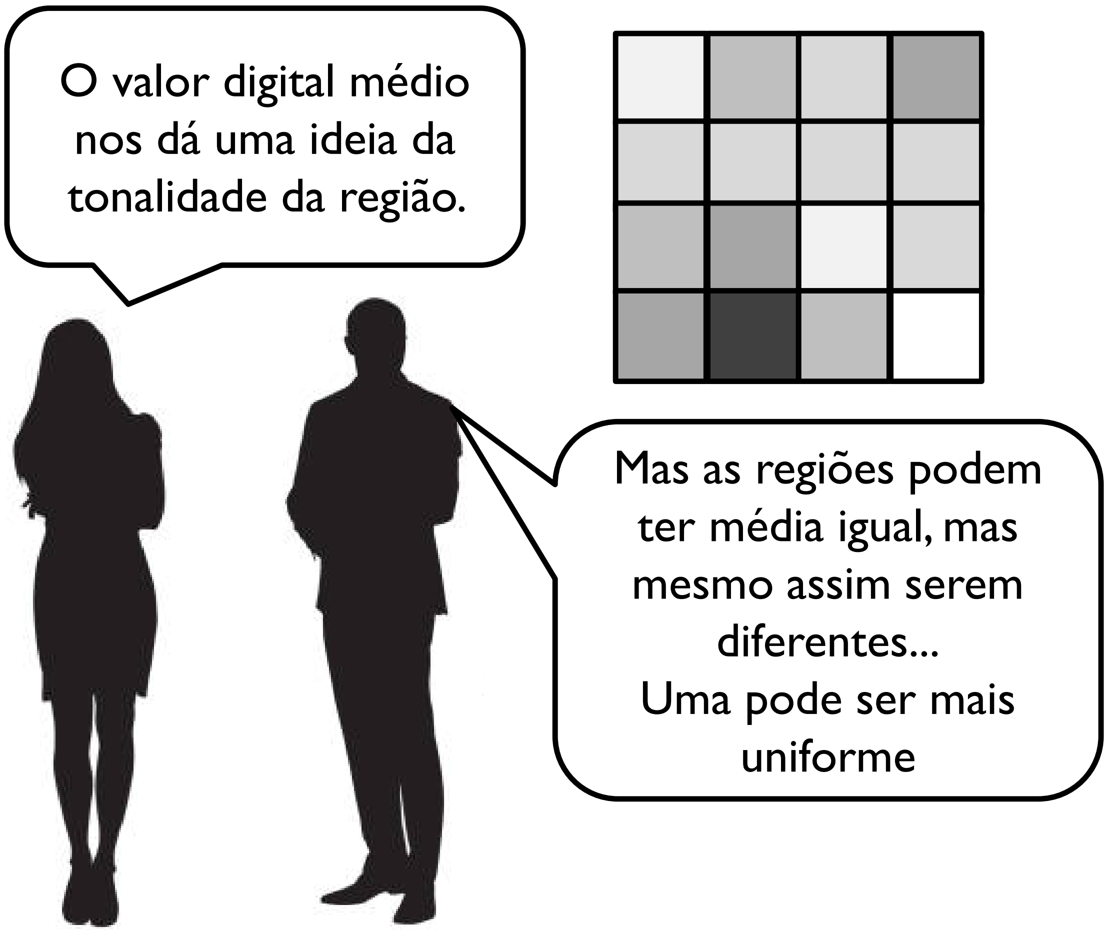
Álgebra Linear

- **Matriz de variância-covariância**
- **Matriz de Correlação**
- **Autovalores e autovetores**

Dados os valores digitais...

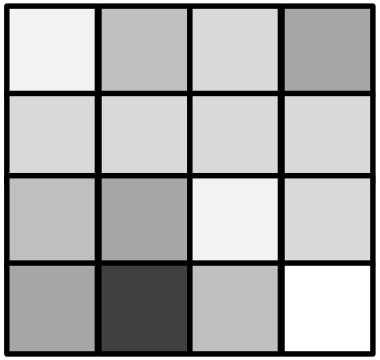
Como descrever a região?

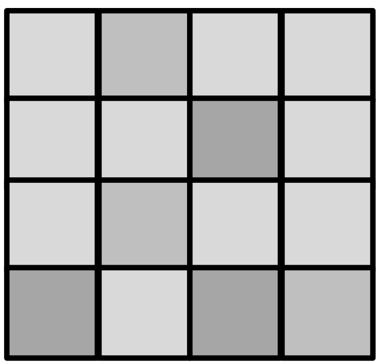


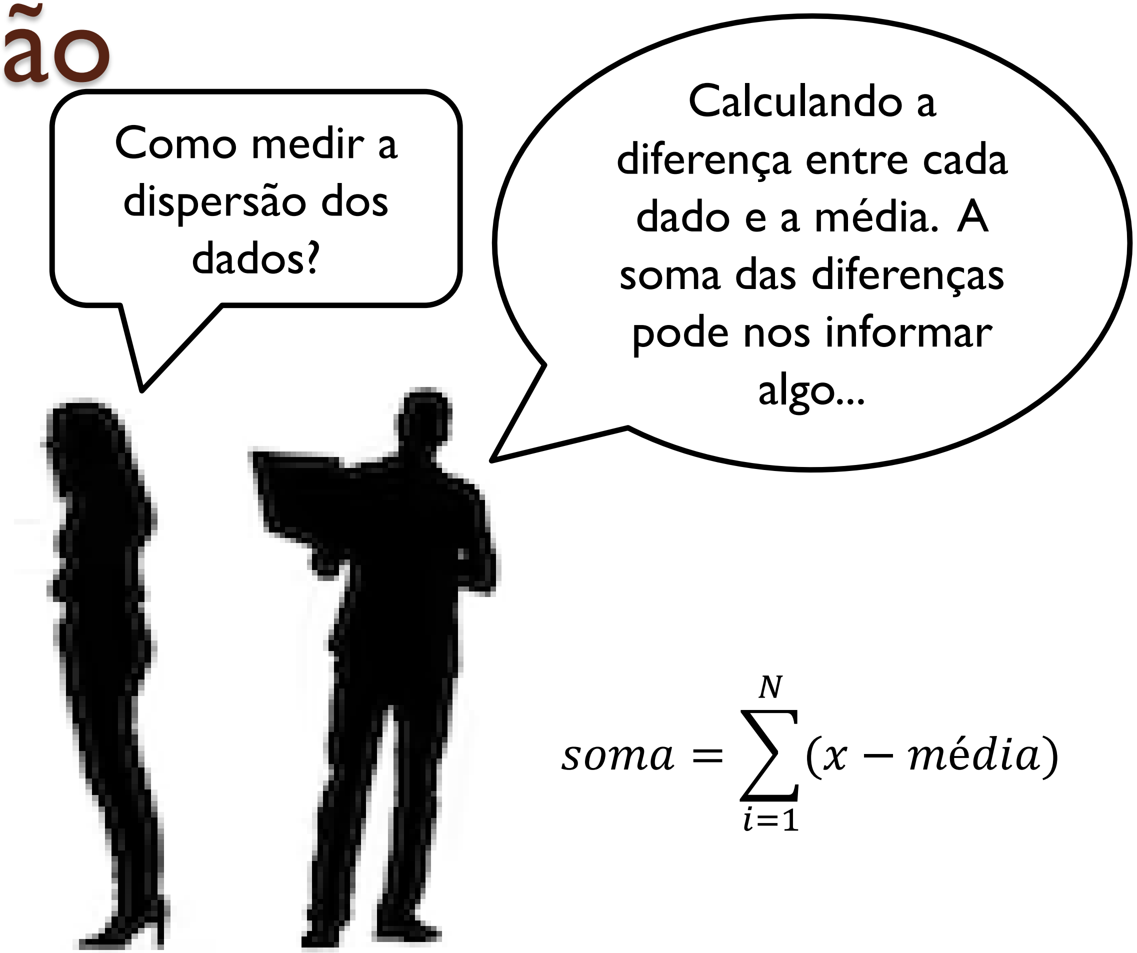


$$m = \frac{\sum_{i=1}^{N} x}{N}$$

A dispersão



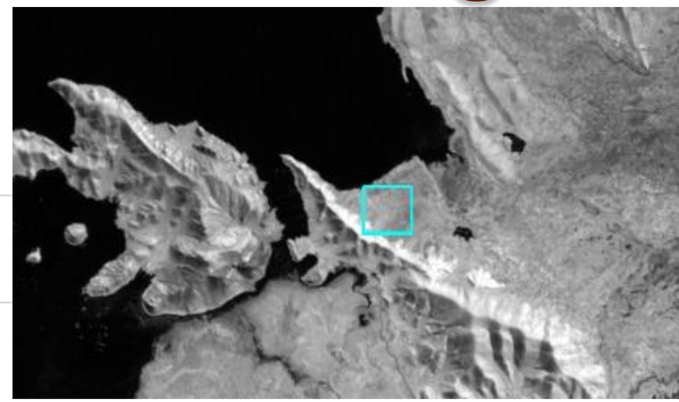




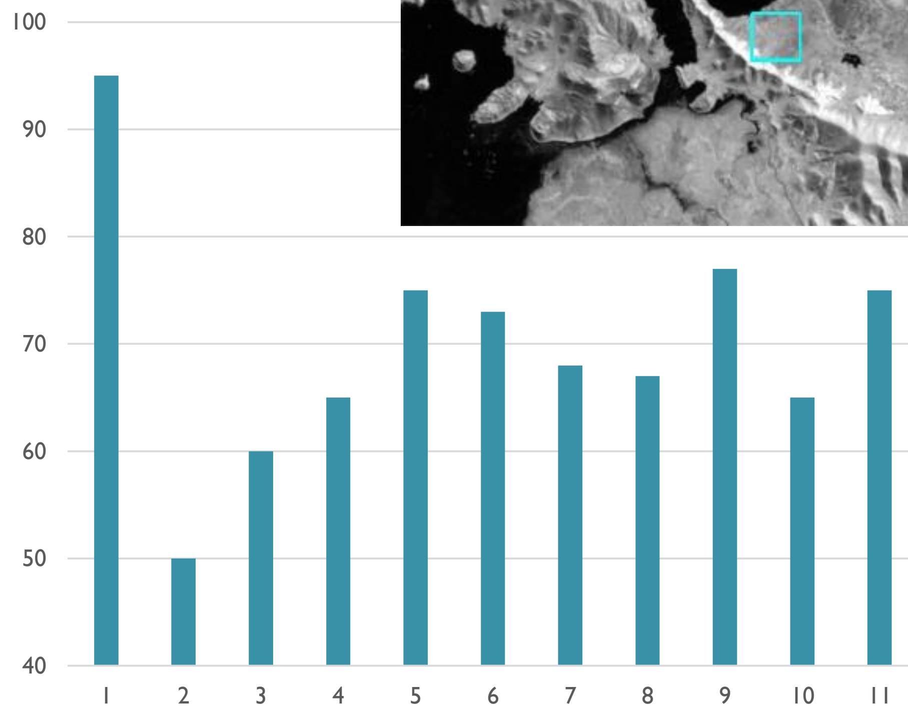
$$soma = \sum_{i=1}^{N} (x - média)$$

Os dados são uniformes?

Valores do pixel em uma região



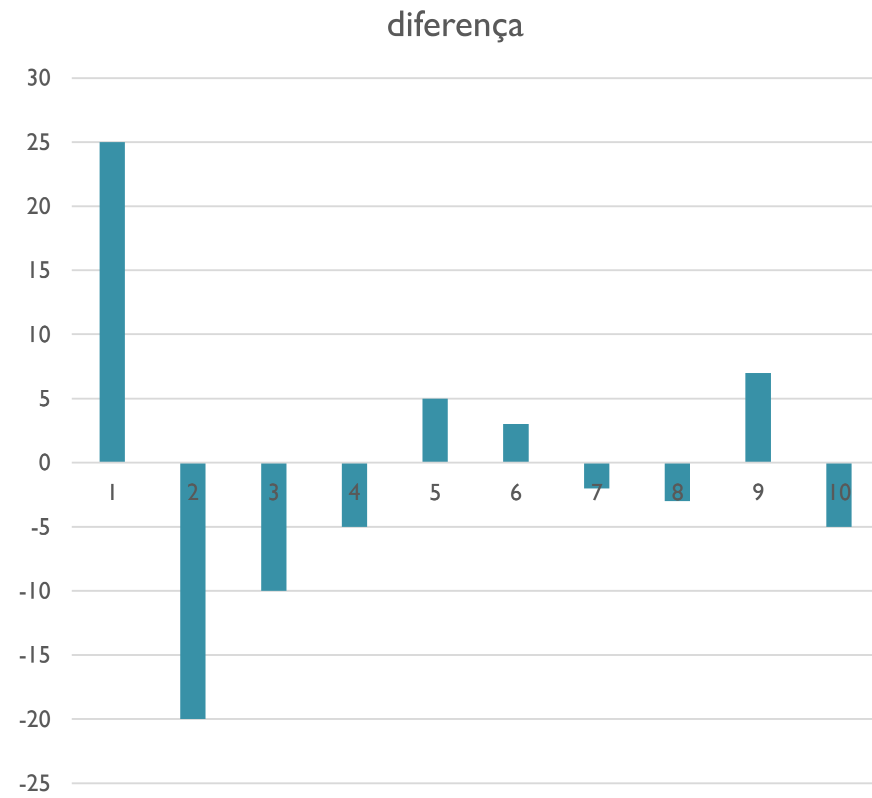
pixel	valores
1	95
2	50
3	60
4	65
5	75
6	73
7	68
8	67
9	77
10	65
11	75



Média=70 mínimo=50 máximo=95 faixa=50-95: 45 valores

Diferença em relação à média

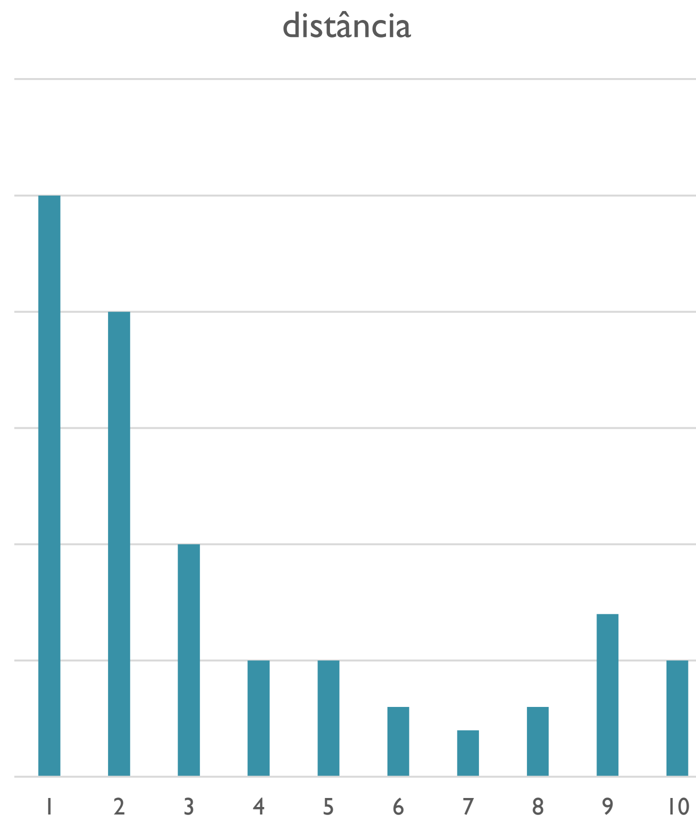
pixel	valores	diferença
1	95	25
2	50	-20
3	60	-10
4	65	-5
5	75	5
6	73	3
7	68	-2
8	67	-3
9	77	7
10	65	-5
11	75	5
media	70	0



Valores se cancelam e a soma é nula! Não informa a dispersão!

Diferença ao quadrado (distância)

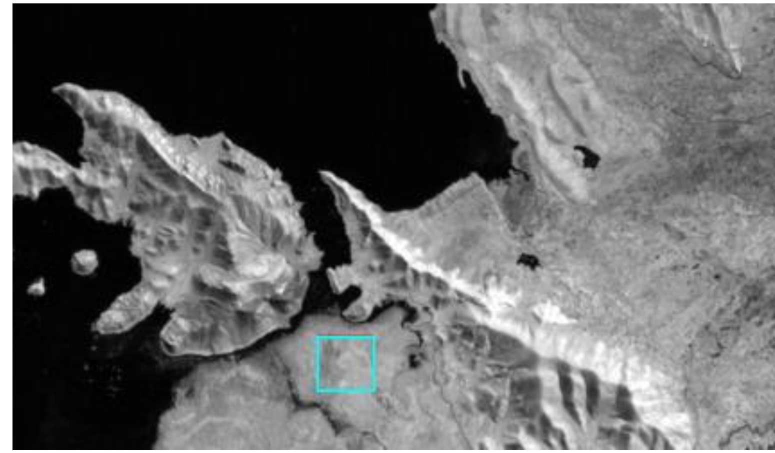
p	valores	dif	dif^2	distância
1	95	25	625	25
2	50	-20	400	20
3	60	-10	100	10
4	65	-5	25	5
5	75	5	25	5
6	73	3	9	3
7	68	-2	4	2
8	67	-3	9	3
9	77	7	49	7
10	65	-5	25	5
11	75	5	25	5
média	70	0	117,82	8,18

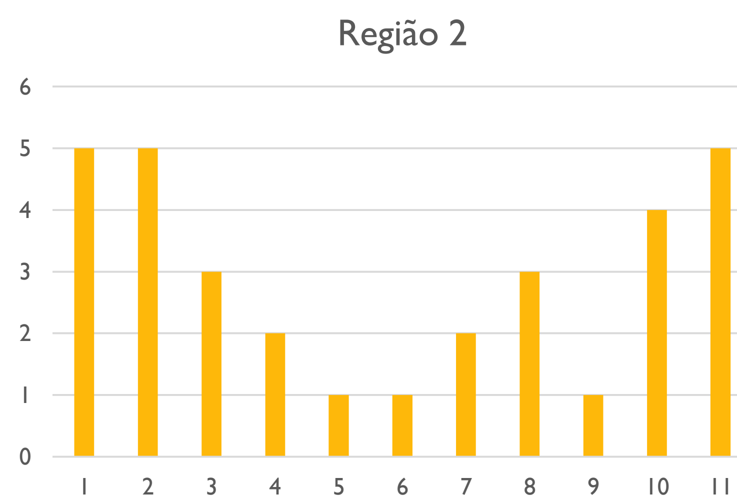


A distância, valor absoluto, tira o efeito do sinal.
Distância à média = 8,18

Outra região

pixel	valores	dif.	dif^2	distância
1	75	5	25	5
2	75	5	25	5
3	73	3	9	3
4	72	2	4	2
5	69	-1	1	1
6	71	1	1	1
7	68	-2	4	2
8	67	-3	9	3
9	71	1	1	1
10	74	4	16	4
11	65	-5	25	5
media	70,90	0,91	10,91	2,1





A dispersão é menor...

A Variância

Dados os N valores x e sua média m

- A variância (s^2) e o desvio padrão s

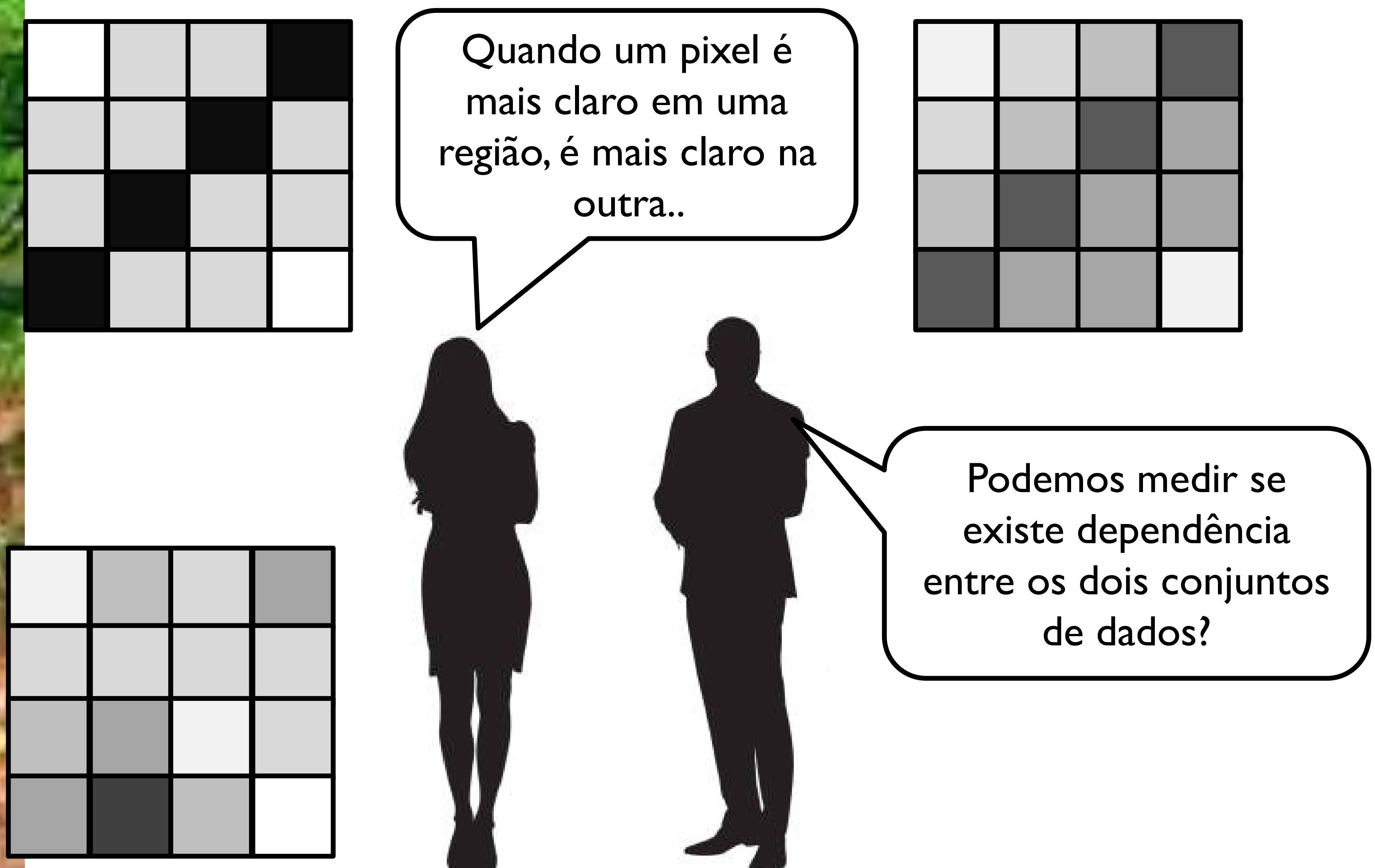
$$s^2 = \frac{\sum (x-m)^2}{N-1}$$

$$s = \sqrt{\frac{\sum (x-m)^2}{N-1}}$$

$(x\text{-}m)^2$ é a distância à média!

E se tiver duas variáveis?

Como descrever se a variação dos valores das duas regiões é dependente?



Covariância

$$\mathrm{cov(x,y)} = \frac{\sum(x - m_x)(y - m_y)}{N - 1}$$

Soma do produto das diferenças em relação à média (considerando o sinal!) de duas variáveis x e y.

- O que ocorre quando ...
- (x- m_x)>0 ao mesmo tempo que (y- m_y)>0
- (x- m_x)<0 ao mesmo tempo que (y- m_y)<0

Covariâncias

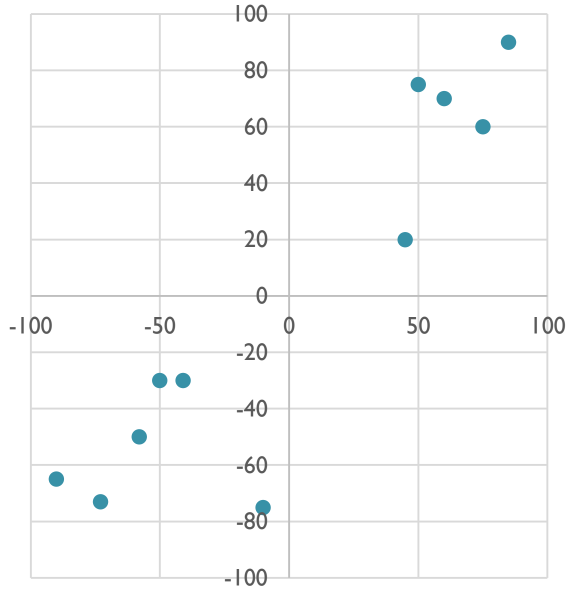
$$\mathrm{cov(x,y)} = \frac{\sum(x - m_x)(y - m_y)}{N - 1}$$

E se os sinais forem trocados?

- (x- m_x)>0 ao mesmo tempo que (y- m_y)<0
- (x- m_x)<0 ao mesmo tempo que (y- m_y)>0

Mesmo sinal vs. sinais trocados

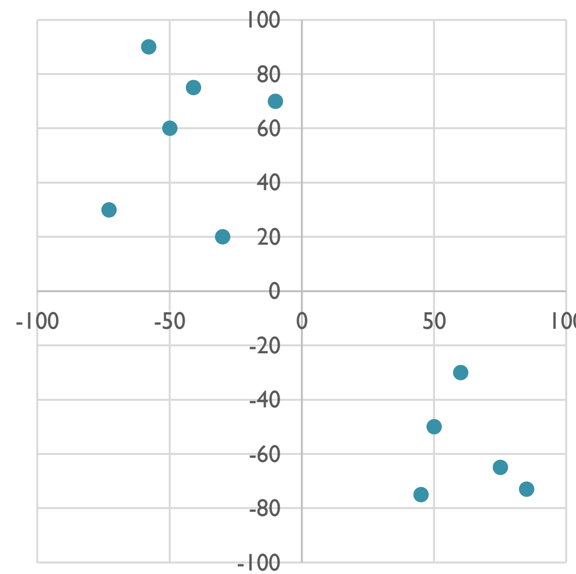
x	y
85	90
50	75
60	70
45	20
75	60
-73	-73
-58	-50
-41	-30
-10	15
-90	-65
-50	-75



Sempre positivo

(-A)* (-B)>0

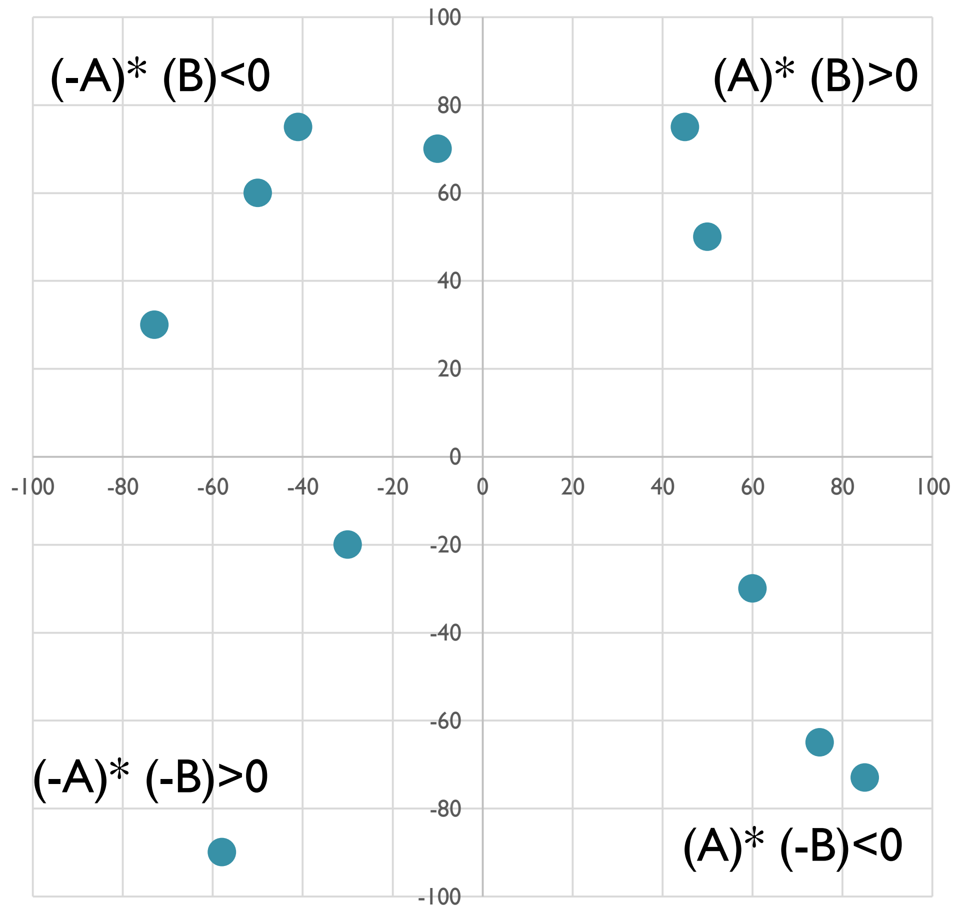
A*B>0



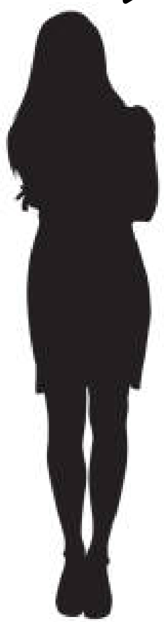
x	y
85	-73
50	-50
60	-30
45	-75
75	-65
-73	15
-58	90
-41	75
-10	70
-90	20
-50	60

E se ocorre de tudo?

x	y
85	-73
50	50
60	-30
45	75
75	-65
-73	30
-58	-90
-41	75
-10	70
-30	-20
-50	60



E se comparo X com X?



- Concordância máxima.

o que resulta da equação da covariância?

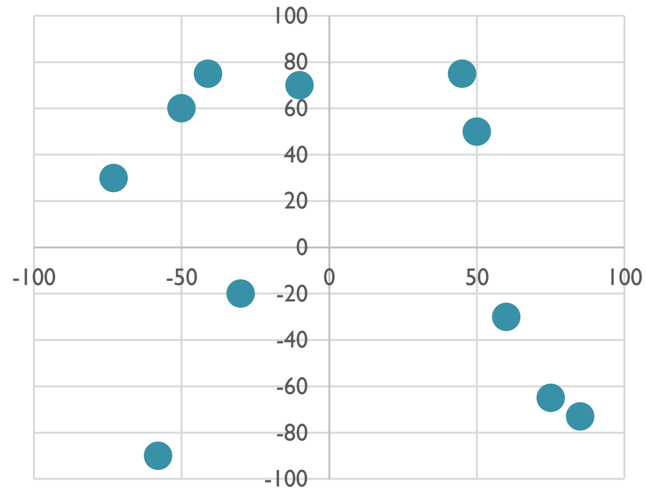
$$COV(X,Y) = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{n-1}$$

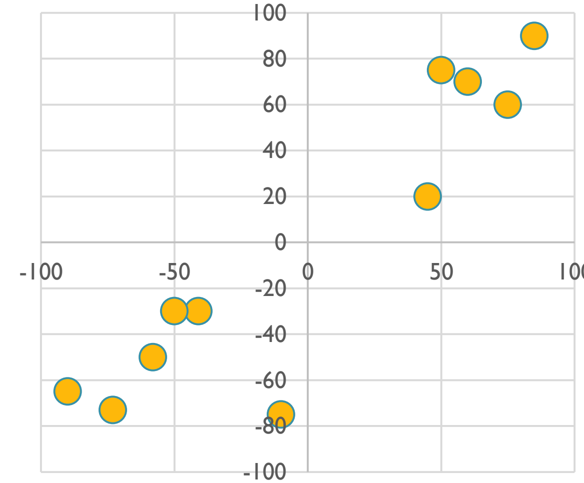
Matriz variância -covariância

- $M = \begin{bmatrix} var(X) & cov(X,Y) & cov(X,Z) \\ cov(X,Y) & var(Y) & COV(Y,Z) \\ cov(X,Z) & cov(Y,Z) & var(Z) \end{bmatrix}$

- Repare na diagonal principal.

Compare a covariância





média de (x-mx)(y-my)=-958,64

média de (x-mx)(y-my)=3504,90

Correlação

$$\mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

- Qual seria a correlação entre X e X?

- O que ocorre se cov(X,Y)=0?
- O que ocorre se cov(X,Y)<0?

Matriz de correlação 3x3

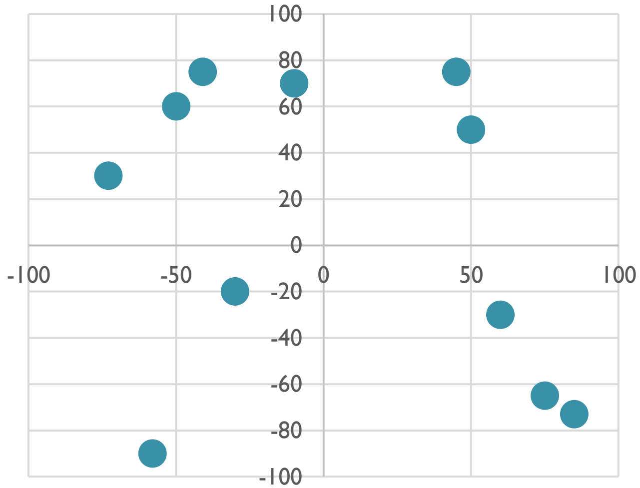
A matriz de correlação é quadrada com dimensão igual ao número de variáveis

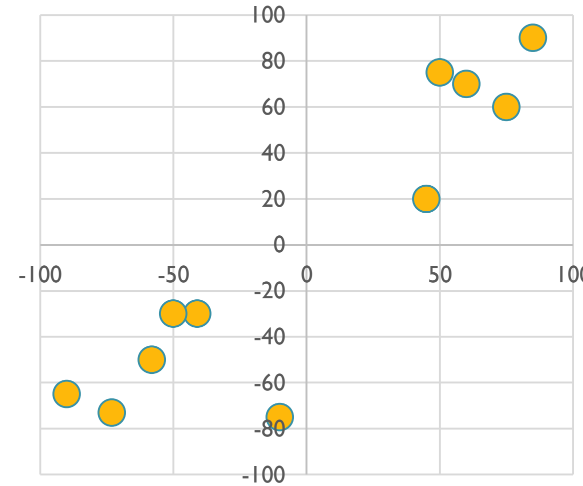
Exemplo para três variáveis

$$Mcorr = \begin{bmatrix} 1 & corr(x,y) & corr(x,z) \\ corr(x,y) & 1 & corr(y,z) \\ corr(x,z) & corr(y,z) & 1 \end{bmatrix}$$

É uma matriz simétrica
É uma matriz quadrada
A diagonal contém apenas 1

Compare a correlação





Cov(x,y)=958,64

Corr= -0,25

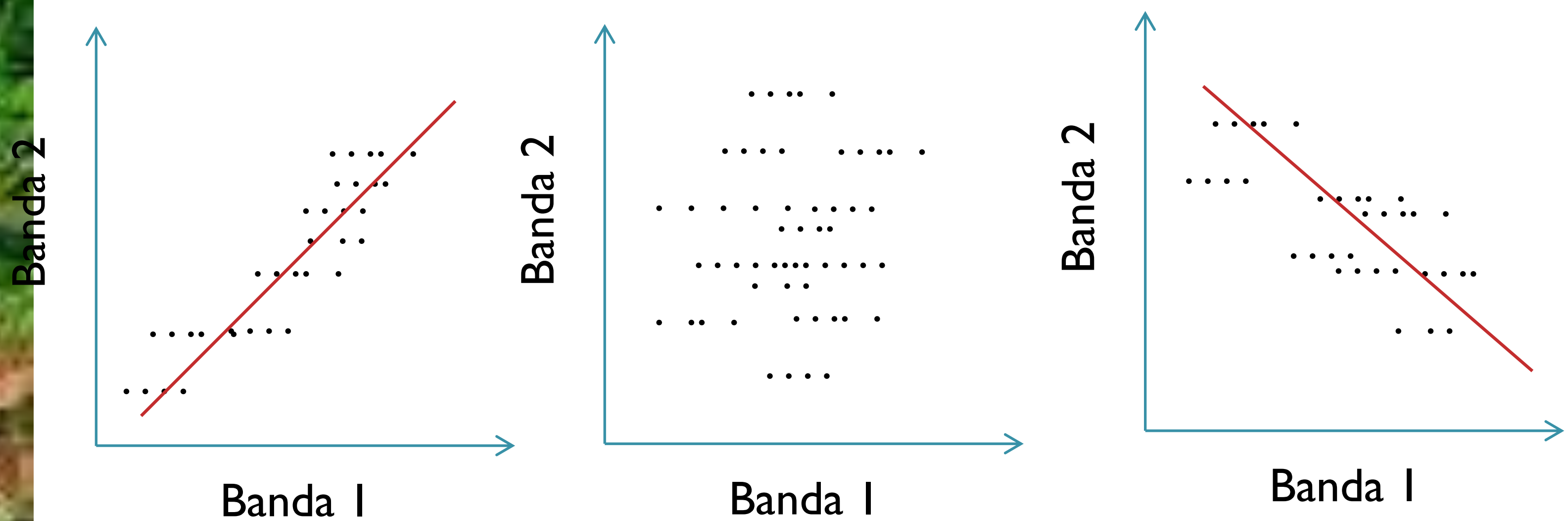
Cov(x,y)=3504,90

Corr= 0,83

Pense na covariância, que deu origem a este valor

Correlação

Correlação positiva,negativa ou zero?



Dispersão no espaço 2D

Como seria a distribuição dos pontos no espaço 2D se conhecemos o tamanho das variâncias e a covariância?

Por exemplo, se...

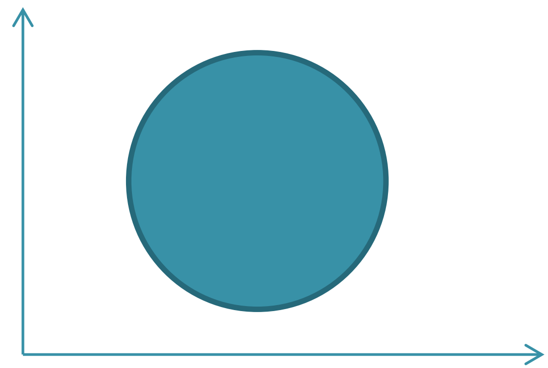
$\sigma_1^2 = \sigma_2^2$ e $\sigma_{12} = 0$?

Ou

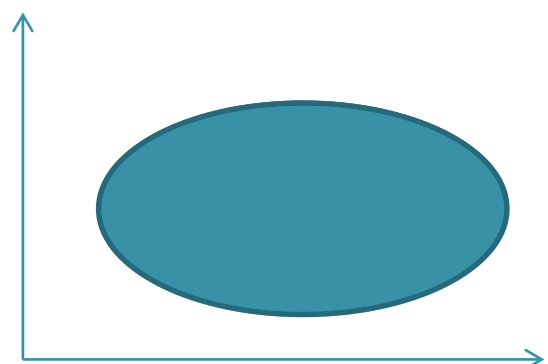
$\sigma_1^2 > \sigma_2^2$ e $\sigma_{12} = 0$?

No caso 2D

$\sigma_1{}^2 = \sigma_2{}^2$ e $\sigma_{12} = 0$?



$\sigma_1{}^2 > \sigma_2{}^2$ e $\sigma_{12} = 0$?



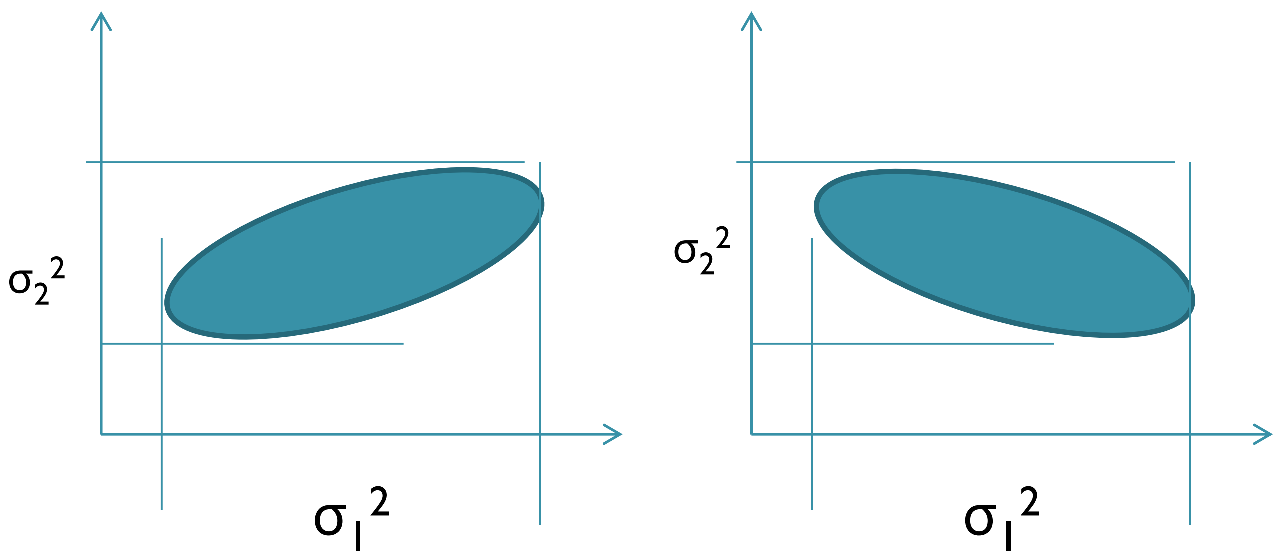
No caso 2D

E se:

$\sigma_1^2 > \sigma_2^2$ e σ_{12} diferente de zero?

No caso 2D

$\sigma_1^2 > \sigma_2^2$ e σ_{12} diferente de zero?



Depende de σ_{12}

Autovalores e Autovetores

Em nosso caso vamos aplicar esta teoria à matriz variância-covariância, que descreve a dispersão dos dados.

Autovalor

Definição: *Seja* ***M*** *uma matriz quadrada* e $\boldsymbol{I}$ *a matriz identidade , os escalares que satisfazem a equação polinomial (que é chamada de equação característica de (**M**):*

$$\det[M - \lambda I] = 0 \quad ou \quad |M - \lambda I| = 0$$

Os escalares **λ** *são chamados de autovalores da matriz M.*

exemplo

$$M = \begin{bmatrix} 510.17 & 559.43 & -335.00 \\ 559.43 & 647.26 & 78.40 \\ -335.00 & 78.40 & 9701.23 \end{bmatrix}$$

$$M - \mathrm{I}*\lambda = \begin{bmatrix} 510.17 & 559.43 & -335.00 \\ 559.43 & 647.26 & 78.40 \\ -335.00 & 78.40 & 9701.23 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$M - \mathrm{I}*\lambda = \begin{bmatrix} 510.17 - \lambda & 559.43 & -335.00 \\ 559.43 & 647.26 - \lambda & 78.40 \\ -335.00 & 78.40 & 9701.23 - \lambda \end{bmatrix}$$

Det(M- I*λ)=0

solução

$$M - \mathrm{I}*\lambda = \begin{bmatrix} 510.17-\lambda & 559.43 & -335.00 \\ 559.43 & 647.26-\lambda & 78.40 \\ -335.00 & 78.40 & 9701.23-\lambda \end{bmatrix}$$

```
Det(M- I*λ )= (510.17- λ)*(647.26- λ)*(9701.23- λ)
              + 559.43*78.40*(-335)
              + 559.43*78.40*(-335)
              - (510.17- λ )* 78.40*78.40
              - (-335)* 647.26- λ *(-335)
              - (9701.23- λ)* 599.43*599.43
```

Det(M- I*λ)=0 resulta em um polinômio de grau N.

Para N=3, três variáveis...

$$a*\lambda^3 + b*\lambda^2 + c*\lambda + d = 0$$

solução

Achar as raízes da equação:

Eq: $a^* \lambda^3 + b^* \lambda^2 + c^* \lambda + d = 0$

- $\lambda 1 =$ 97047.88
- $\lambda 2 =$ 11302.26
- $\lambda 3 =$ 104.96

Estas raízes são os autovalores associados à matriz M

Autovetor

Definição: *Seja uma matriz M com dimensão kxk e λ seus autovalores. Se x é um vetor não nulo tal que:*

- M x = λ *x*

- *então x é chamado de autovetor (vetor característico) da matriz associado ao autovalor λ .*

Exemplo: para matriz quadrada

$$M = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$$

Calcule os autovalores

$$\begin{vmatrix} 6-\lambda & -1 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

$$(6-\lambda)(3-\lambda)+2=0$$

$$18- \ 6\lambda-3\lambda \ +\lambda^2 \ +2=0$$

$$\lambda^2 \ -9\lambda \ +20=0$$

Os autovalores são : 4 e 5

E os autovetores?

M*x = λ*x

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Os autovetores

Para o primeiro Autovalor λ=4

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Desenvolvendo

$$6x_1 - x_2 = 4x_1$$

$$2x_1 + 3\, x_2 = 4x_2$$

$$2\, x_1 = x_2$$

$$2x_1 = x_2$$

Temos apenas uma equação. Existem inúmeras soluções para a igualdade: Uma solução pode ser definida, arbitrando valores para x1 ou x2.

Ex: se x_1=1

x_2=2 ... Vetor solução 1 V1=(1,2)

Os autovetores

Para o segundo Autovalor λ=5

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$6x_1 - x_2 = 5x_2$$

$$2x_1 + 3x_2 = 5x_2$$

$$x_1 = x_2$$

$$2x_1 = 2x_2$$

Ex: se x_2=1

$$x_1 = 1$$

Os dois autovetores são

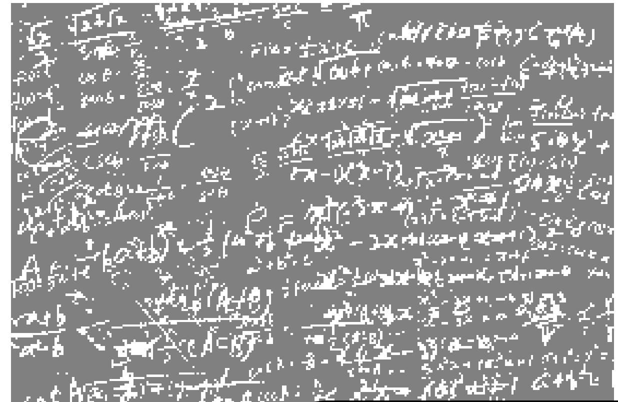
v1=[1,2]

V2=[1,1]

Note que existem outras soluções, múltiplos destes vetores...

v1=[2,4], V2=[9,9] ... etc

A seguir...



Componentes principais

Componentes principais

- **Análise das componentes principais**
- Do ponto de vista estatístico, o objetivo da análise das componentes principais consiste em representar um conjunto de dados usando um novo conjunto de variáveis, combinações lineares das originais.
- Este novo sistema deve ter como propriedade um baixo grau de correlação entre as novas variáveis.
- A transformação das componentes principais consiste basicamente numa rotação dos dados no espaço das bandas.

Exemplo 2D

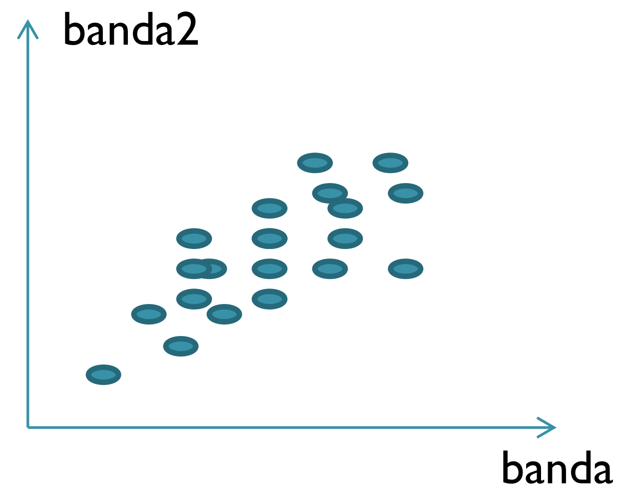
Bandas Originais

Dado um conjunto com duas variáveis, por exemplo os valores digitais em duas bandas espectrais, a dispersão dos valores pode ser descrita com a matriz variância-covariância.

Caso a covariância não seja nula, pode-se dizer que existe dependência linear entre as variáveis (correlacionadas)

$\sigma_1^2 > 0, \sigma_2^2 > 0$

e σ_{12} diferente de zero



Se existe correlação, parte da informação do conjunto é redundante.

Rotação 2D

Aplicando uma transformação de rotação, é possível alinhar novos eixos à dispersão dos dados.

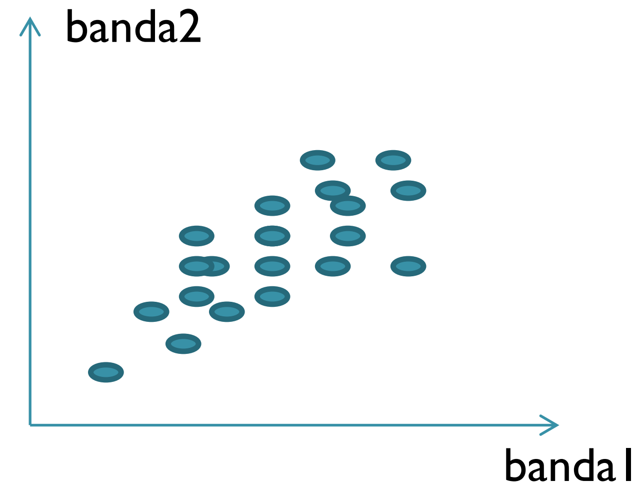
Um eixo passando pela direção de maior dispersão e o segundo perpendicular ao primeiro.

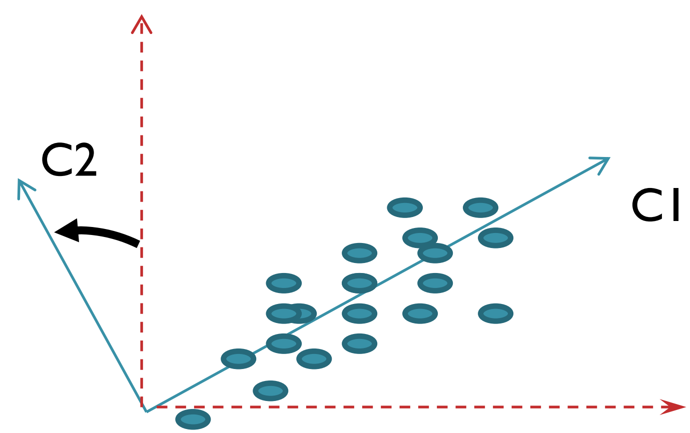
$$\begin{bmatrix} c_1 \\ c2 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} * \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Ou

$$c_1 = r_{11}b_1 + r_{12}b_2$$
$$c_2 = r_{21}b_1 + r_{22}b_2$$

Transformação linear das variáveis originais.





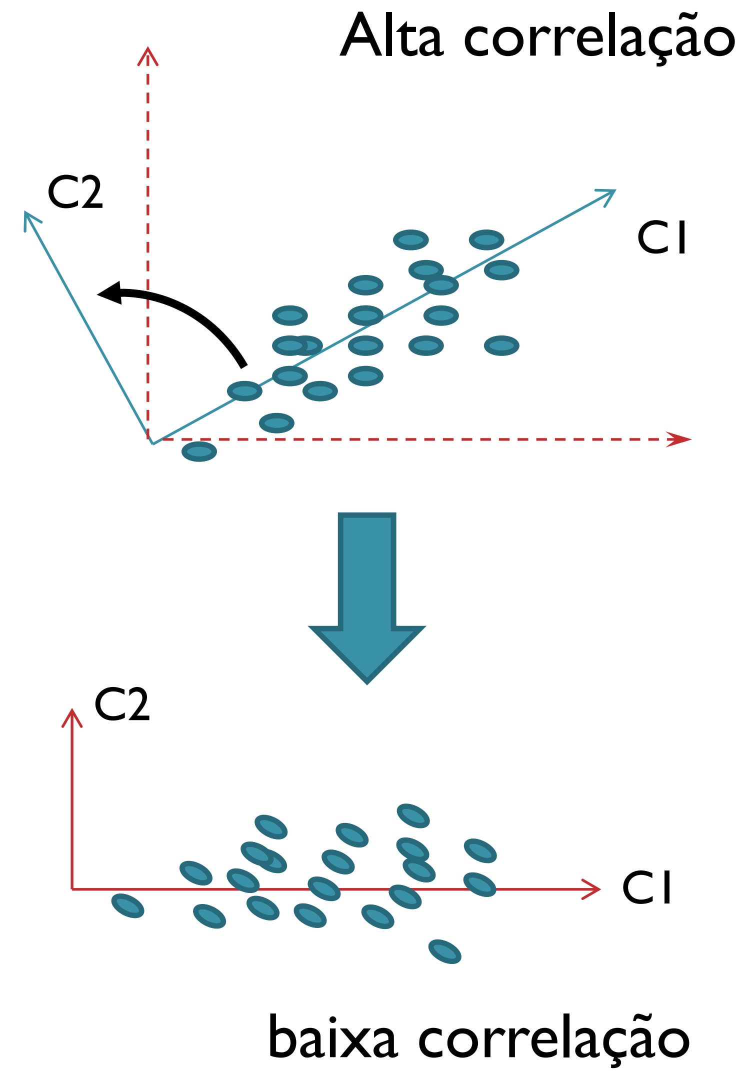
Rotação 2D

O novo sistema apresenta baixa correlação e a variação dos dados é muito melhor explicada pelo eixo novo c1.

A discriminação dos dados seria mais fácil projetando os dados ao longo do eixo principal da dispersão.

$$c_1 = r_{11}b_1 + r_{12}b_2$$
$$c_2 = r_{21}b_1 + r_{22}b_2$$

As novas variáveis são resultado da transformação linear das variáveis originais.



Transformação das CP

A transformação das componentes principais consiste na rotação dos dados no espaço das bandas. Com isto, os dados são representados num novo sistema de bandas (combinação linear das originais), que salientam algumas feições.

$$\begin{bmatrix} c1 \\ c2 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

A solução para determinar os coeficientes necessários para efetuar esta rotação é dada pela análise da matriz de variância-covariância do conjunto de observações. Usando o conceito de autovetores.

Os autovetores são paralelos às direções da dispersão do conjunto original

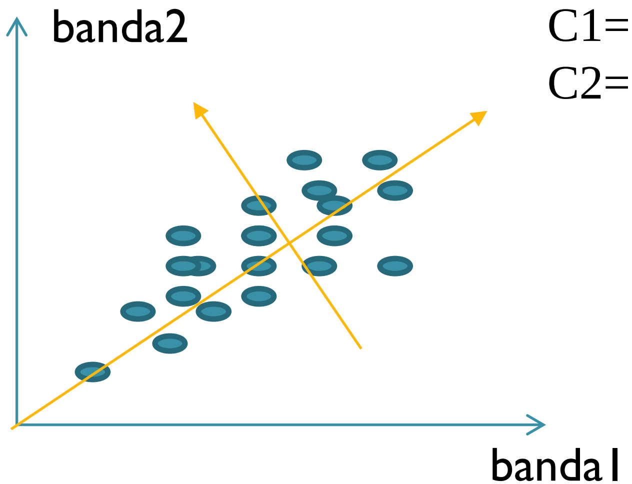
Aplicado a imagens

Calcular a matriz variância-covariância da imagem considerando N bandas.

Calcular os autovetores e autovalores da matriz de variância-covariância.

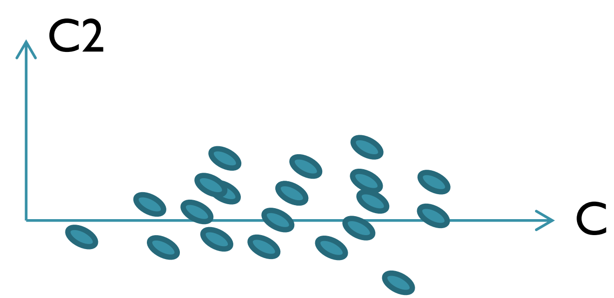
Existem N autovalores e N autovetores associados.

Os autovetores são paralelos às direções da dispersão do conjunto original



$C1=a1*banda1 + a2*banda2$

$C2=a3*banda1 + a4*banda2$



Bandas Originais

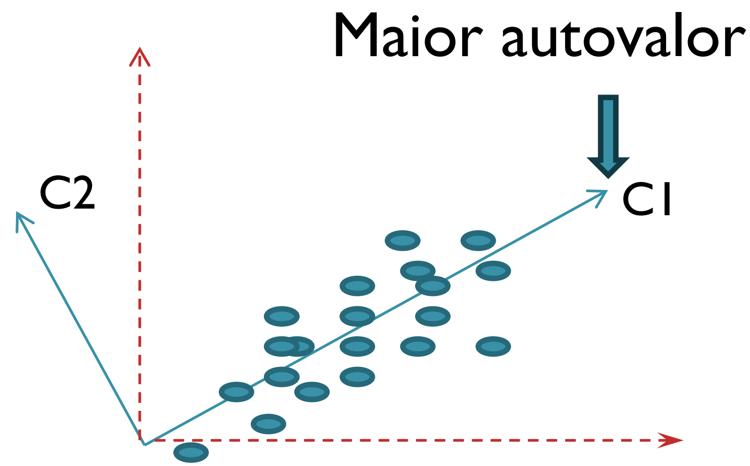
Componentes Principais

Qual autovetor?

Para uma matriz N-dimensional são obtidos N autovetores. A questão é identificar aquele associado à direção de maior dispersão.

Solução:

O teor de informação de cada componente é proporcional ao tamanho do autovalor associado.



Quanto maior o autovalor, mais informação é contida nessa componente. Por isso, ordena-se os autovalores de forma decrescente para identificar as principais componentes calculadas.

Teor de informação

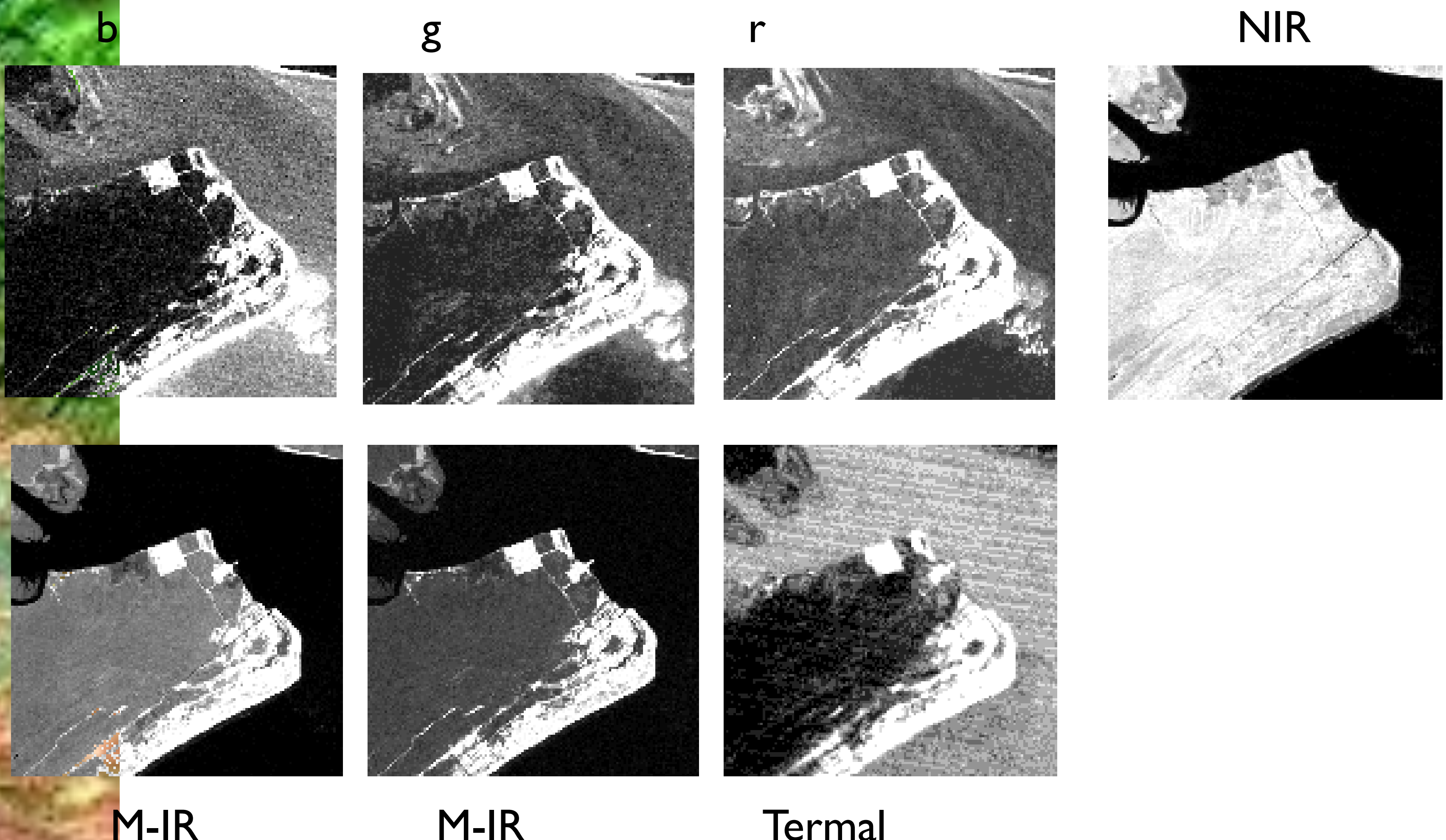
A relação entre os autovalores indica a importância de cada componente e seu respectivo teor de informação.

A variância original explicada por cada componente pode ser calculada como a relação entre o autovalor e a soma de todos os autovalores.

A percentagem da informação original representada por cada banda é dada por:

$$C_i = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i} \cdot 100 = \frac{\lambda_i}{traço(S)} \cdot 100$$

Imagem de 7 bandas

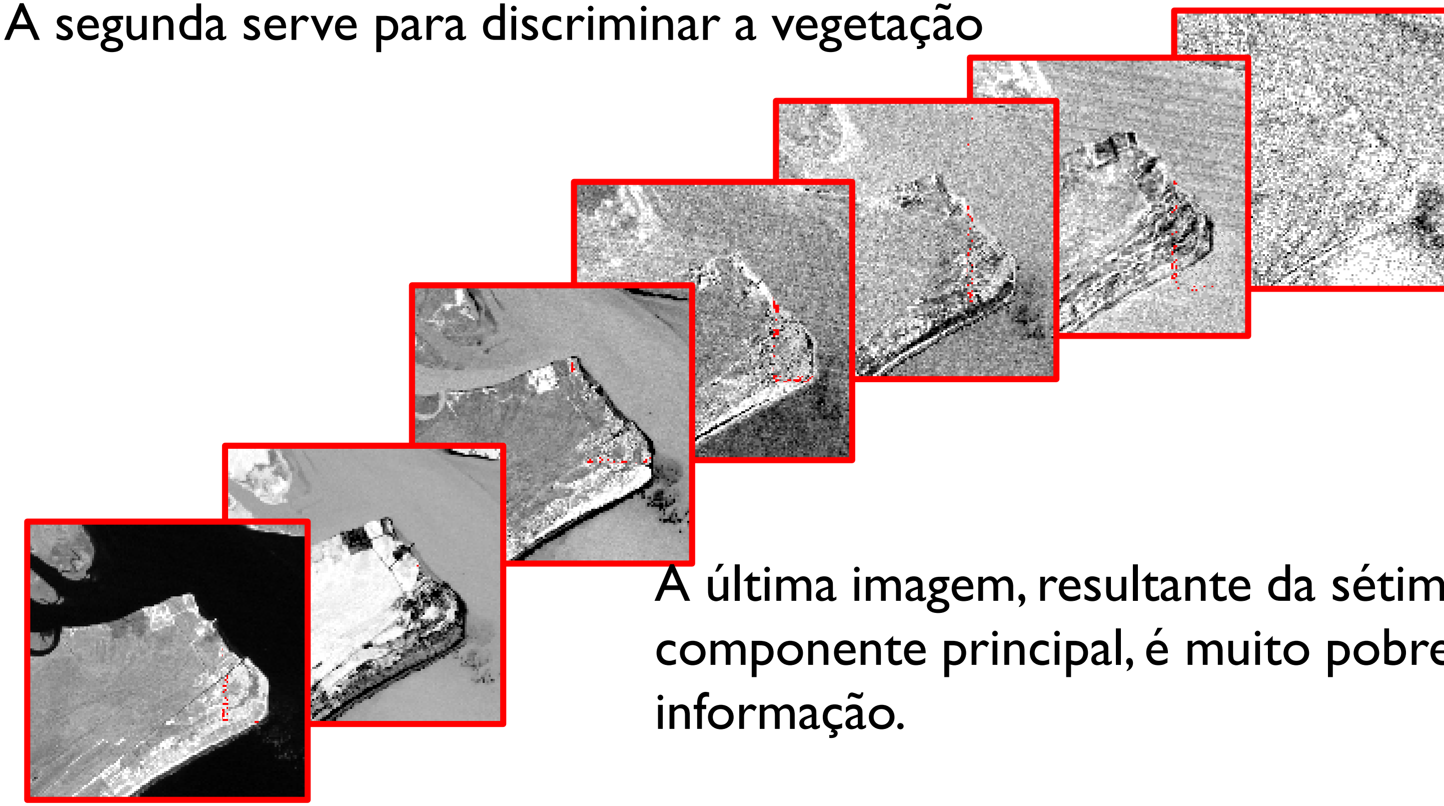


Exemplo....

As primeiras componentes prestam-se melhor para uma análise visual, pois permitem separar melhor os elementos.

A variabilidade dos elementos na imagem é realçada principalmente com a primeira componente principal.

A segunda serve para discriminar a vegetação



A última imagem, resultante da sétima componente principal, é muito pobre em informação.

Analisar os coeficientes da componente

Sendo cada autovetor uma combinação linear das variáveis originais, os coeficientes associados a cada variável original servem para interpretar qual aspecto dos dados originais é melhor descrito por cada componente.

- Ex: se...

  $CP = 0{,}44\ Red - 0{,}45 * NIR + 0{,}01 MIR$

- O que você acha?

C.P.

				PROX	MIR	MIR
1)	0.00	0.00	0.12	0.82	0.49	0.25
2)	-0.35	-0.38	-0.53	0.41	-0.30	-0.41
3	0.31	0.40	0.40	0.34	-0.47	-0.46

PAN

NEG + NEG

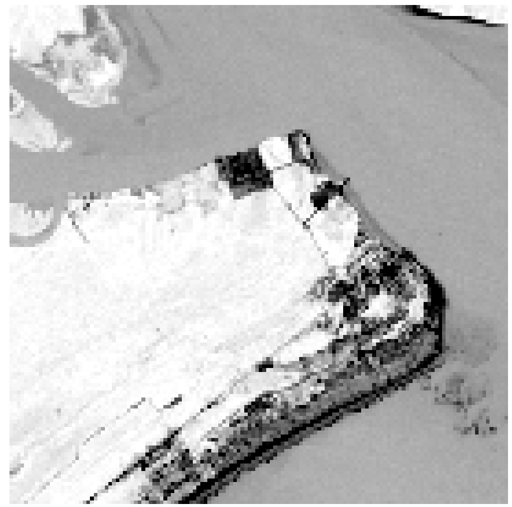
\+ NEG

Traduzindo,

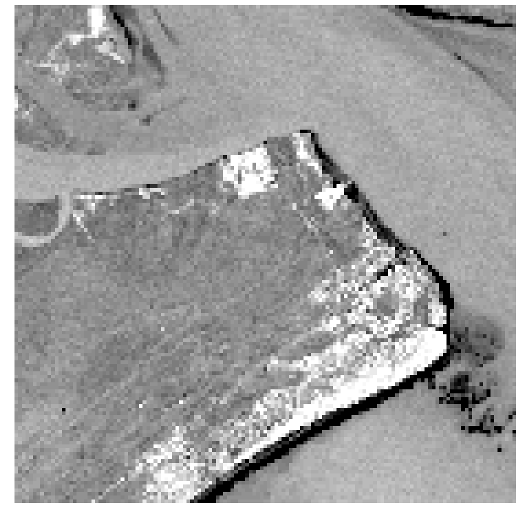
1) A 1ª componente é a soma de diferentes bandas , do vermelho até o infravermelho médio (Uma Pancromática)
2) cp2= IVprox – (VIS-IVmédio): vegetação
3) Medio – VIS - IVprox

3 primeira componentes

1



2



3

Tarefa:

- Com o recorte da imagem de Curitiba e o Mulstispec:

a) calcule a correlação entre as sete primeiras bandas originais do LANDSAT OLI a bordo do Landsat 8.

b) Calcule os autovalores da matriz de correlação de uma área amostral desta imagem. Para isto, marque uma área suficientemente grande como área de treinamento e liste (list stats) as estatísticas desta região.

c) Calcule os autovalores e autovetores da matriz variância-covariância desta imagem (usando Processor + Utilities + Principal Components Analysis)

d) Analise o tamanho relativo dos autovalores e responda:
   a) Qual a percentagem da variância original explicada pela primeira componente principal?
   b) Com quantas bandas podemos representar 99% da informação original?

e) Analise os autovetores que podem representar 99% da informação e responda: é possível dar uma interpretação a cada um deles? Que tipo de informação pode estar contido em cada componente?

f) Aplique a transformação das componentes pincipais à imagerm (Processo+ reformat + Change Image file format + Transfrom data + New Channel from PC eigenvectors)

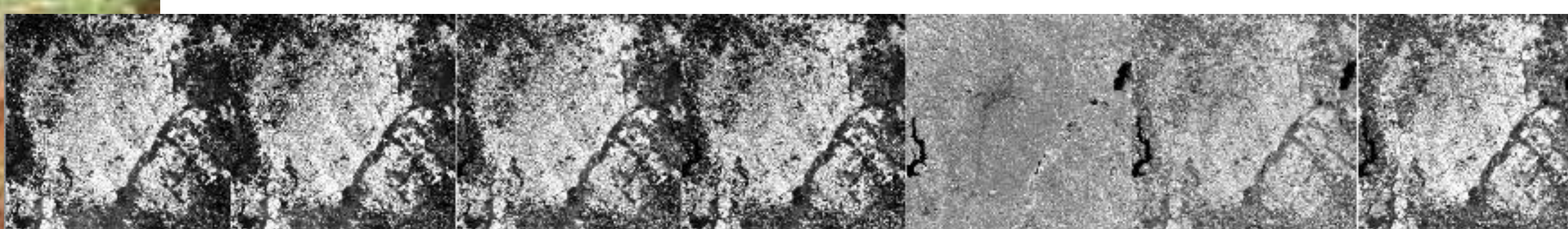
g) Visualize as imagens e componha uma composição colorida com as três primeiras componentes principais.

h) Calcule a matriz de correlação desta nova imagem e compare com aquela obtida na etapa a.

Curitiba: matriz de correlação

Correlation Matrix

Ch	1	2	3	4	5	6	7
1	1.00						
2	0.99	1.00					
3	0.97	0.97	1.00				
4	0.95	0.96	0.98	1.00			
5	-0.12	-0.11	0.01	-0.03	1.00		
6	0.72	0.75	0.81	0.84	0.33	1.00	
7	0.88	0.89	0.91	0.93	0.06	0.93	1.00



Curitiba: Autovalores

Component	Eigenvalue	Percent	Cum. Percent
1	29483.5519	90.34	90.3404
2	2055.7045	6.30	96.6393
3	673.2324	2.06	98.7022
4	241.5155	0.74	99.4422
5	85.0561	0.26	99.7028
6	62.5806	0.19	99.8946
7	34.4106	0.10	100.0000

CP\banda	1	2	3	4	5	6	7	%
1	0.44	0.45	0.42	0.44	-0.00	0.26	0.39	90
2	-0.33	-0.28	-0.05	0.00	0.56	0.60	0.35	6
3	0.19	0.17	0.27	-0.05	0.74	-0.18	-0.51	2
4	0.44	0.28	-0.34	-0.70	0.11	0.07	0.29	0.7
5	0.10	0.28	-0.63	0.34	-0.03	0.43	-0.44	0.2
6	-0.11	0.09	0.47	-0.42	-0.32	0.55	-0.39	0.1
7	-0.66	0.71	-0.02	-0.06	0.06	-0.16	0.12	0.1

Ex:

CP1 representa 90% da informação e é calculada assim:

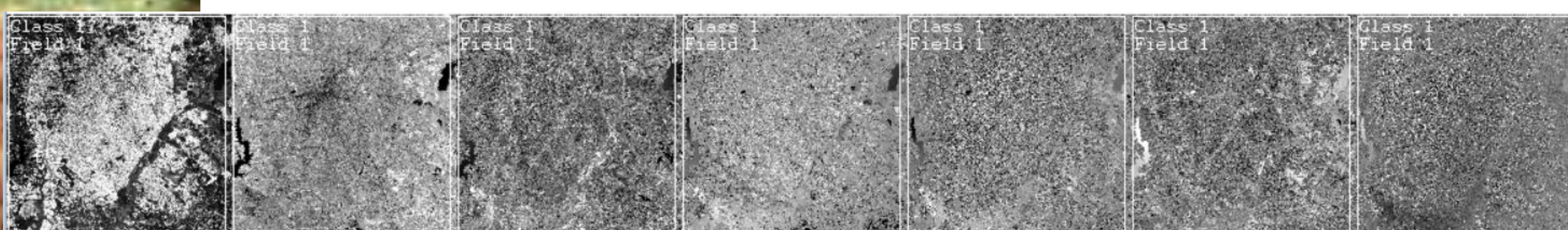
CP1= 0.44**B1** +0.45**B2** +0.42**B3** +0.44**B4** +0.26**B6** +0.39**B7**

Ex:

CP2 representa 6% da informação e é calculada assim:

CP2= -0.33**B1**-0.28**B2** + 0.56**B5**+ 0.60**B6** + 0.35**B7**

Channel	1	2	3	4	5	6	7
1	1.00						
2	-0.01	1.00					
3	-0.00	-0.01	1.00				
4	0.00	0.00	-0.01	1.00			
5	0.00	0.00	0.01	0.00	1.00		
6	-0.00	-0.00	0.01	0.00	-0.00	1.00	
7	-0.00	-0.00	-0.00	0.01	0.01	0.00	1.00



3 primeiras componentes