

Classificação

Não Supervisionada

Prof. Dr. Rogério Galante Negri

Recordando – notações

- \mathcal{I} uma imagem definida sobre um reticulado $\mathcal{S} \in \mathbb{N}^2$
 - $\mathcal{I}(s) = \mathbf{x}$ denota que o pixel $s \in \mathcal{S}$ possui vetor de atributos $\mathbf{x} \in \mathcal{X}$
 - \mathcal{X} é o Espaço de Atributos de \mathcal{I}
 - $F: \mathcal{X} \rightarrow \mathcal{Y}$ é um classificador que associa elementos de \mathcal{X} a um indicador de classe em $\mathcal{Y} = \{1, 2, \dots, c\}$
 - Seja $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$ o conjunto de classes, o valor do indicador de classe y_i determina a classe ao qual \mathbf{x}_i é associado:
 - Classificadores de aprendizado não supervisionado estimam F através de relações encontrados em
$$\mathcal{D}_u = \{\mathbf{x}_i \in \mathcal{X} : \forall \mathbf{x}_i \exists s \mathcal{I}(s) = \mathbf{x}_i; i = 1, \dots, m'; i \leq \#\mathcal{S}\}$$
 - Não são identificadas classes, mas sim grupos similares
 - Em resumo, \mathcal{D}_u são todos padrões que compõe a imagem \mathcal{I}
-

Algoritmo K-Médias

- Particiona o conjunto de dados (imagem) em k agrupamentos
 - Seja $\mathbf{x}_i = [x_{1i}, x_{2i}, \dots, x_{ni}] \in \mathcal{X}$ um padrão
 - O padrão é o vetor de atributos do pixel
 - Seja $\theta_j \in \mathcal{X}$ um centróide, com $j = 1, \dots, k$
 - θ_j é um “representante do agrupamento”
 - Os padrões \mathbf{x}_i são associados aos centróides via “menor distância”
 - Após associar cada \mathbf{x}_i a um dos centróides, tais centróides são atualizados segundo o valor médio dos \mathbf{x}_i 's associados

 - Após um processo iterativo, haverá convergência

 - A última associação feita define o agrupamento (classificação)
-

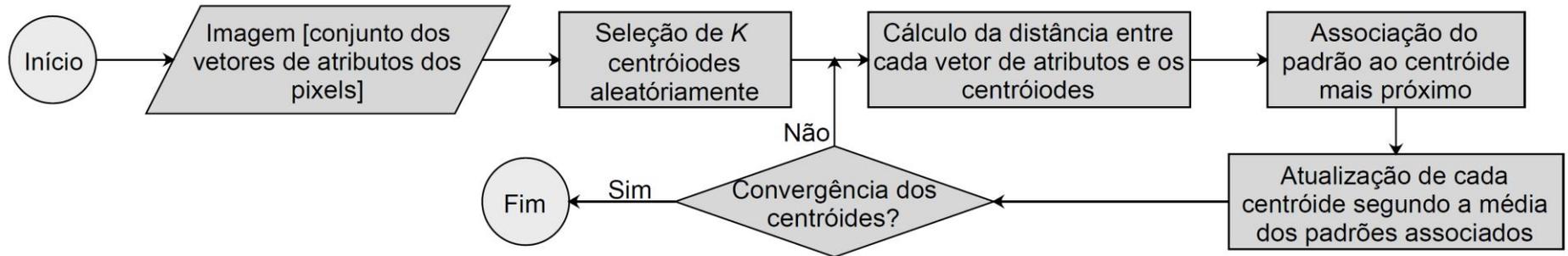
Algoritmo – K-Médias

- i. Define \mathcal{J} (consequentemente $s \in \mathcal{S}$)
 - ii. Define k e inicializa θ_j e $\bar{\theta}_j$; para $j = 1, \dots, k$
 - iii. $T_j \leftarrow 0$; $cont_j \leftarrow 0$; para $j = 1, \dots, k$

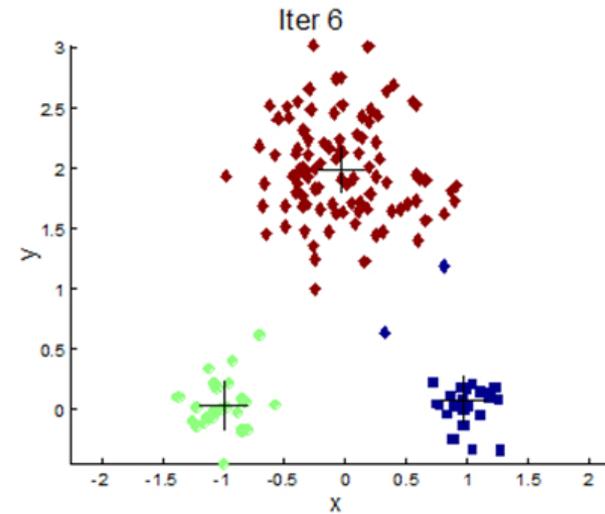
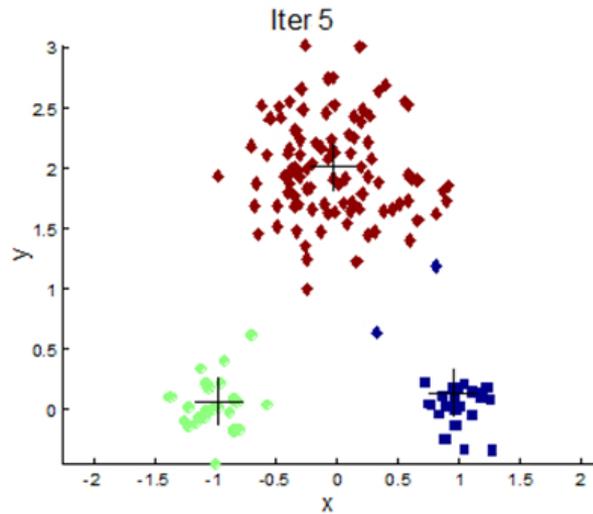
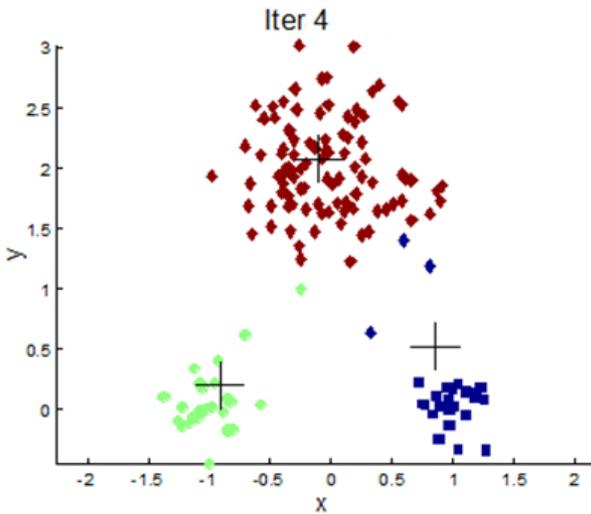
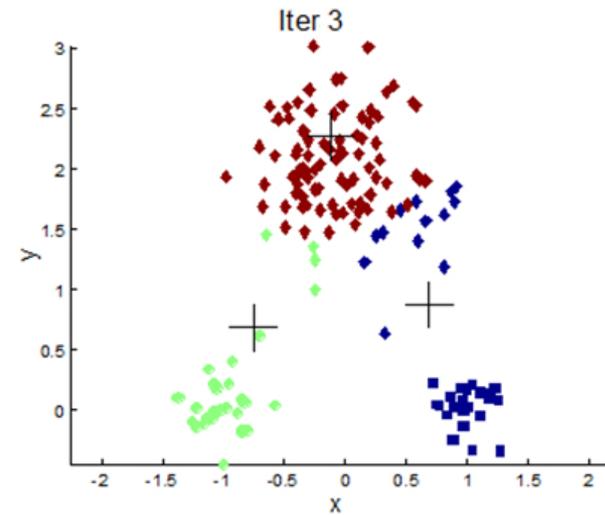
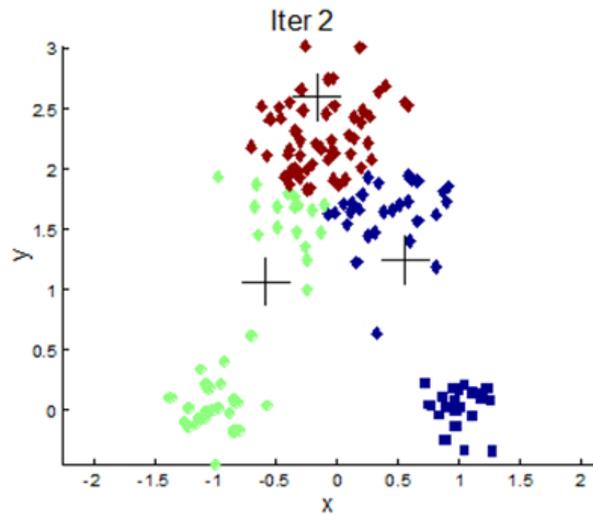
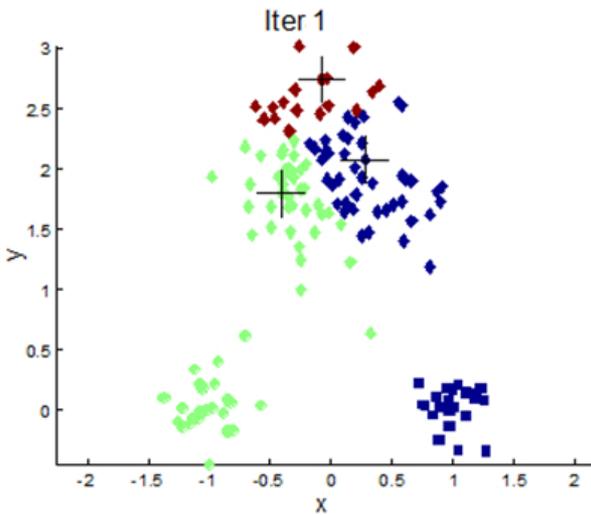
 - iv. Para cada $s_i \in \mathcal{S}$, com $i = 1, \dots, \#\mathcal{S}$, $\mathbf{x}_i \leftarrow \mathcal{J}(s_i)$
 - v. Para cada θ_j , com $j = 1, \dots, k$, $d_j \leftarrow \|\mathbf{x}_i - \theta_j\|$
 - vi. $p \leftarrow \arg \min_{j=1, \dots, k} d_j$
 - vii. $T_p \leftarrow T_p + \mathbf{x}_i$
 - viii. $cont_p \leftarrow cont_p + 1$
 - ix. Para $j = 1, \dots, k$, atualizar: $\theta_j \leftarrow T_j / cont_j$
 - x. Se $\theta_j \neq \bar{\theta}_j$; para algum $j = 1, \dots, k$ então $\bar{\theta}_j \leftarrow \theta_j$ e retorne para (iv)

 - xi. Para cada $s_i \in \mathcal{S}$, com $i = 1, \dots, \#\mathcal{S}$, determina-se $\mathcal{C}(s_i) \leftarrow \arg \min_{j=1, \dots, k} \|\mathbf{x}_i - \theta_j\|$
-

Algoritmo – K-Médias



Iteração-Convergência

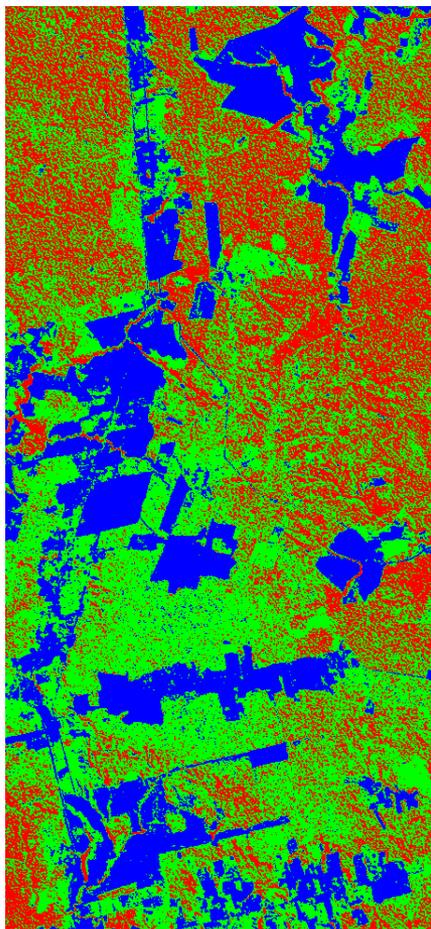


Exemplo – K-Médias

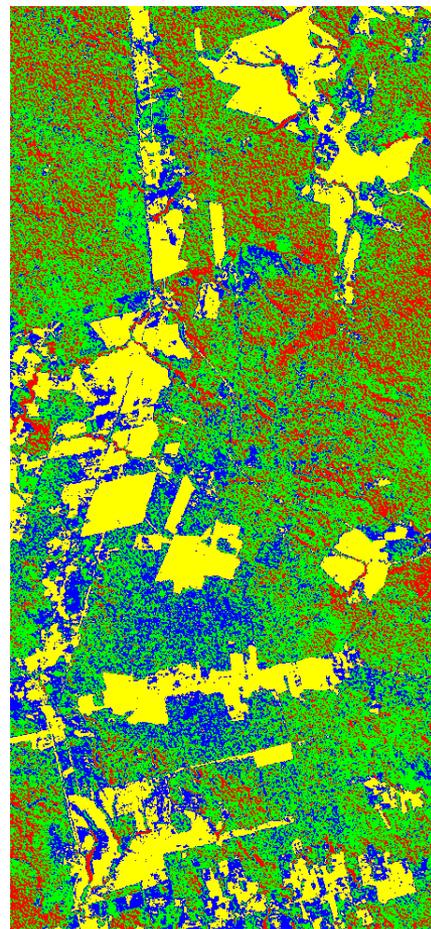
- LANDSAT-5 TM – Set/2010 – FLONA Tapajós-PA



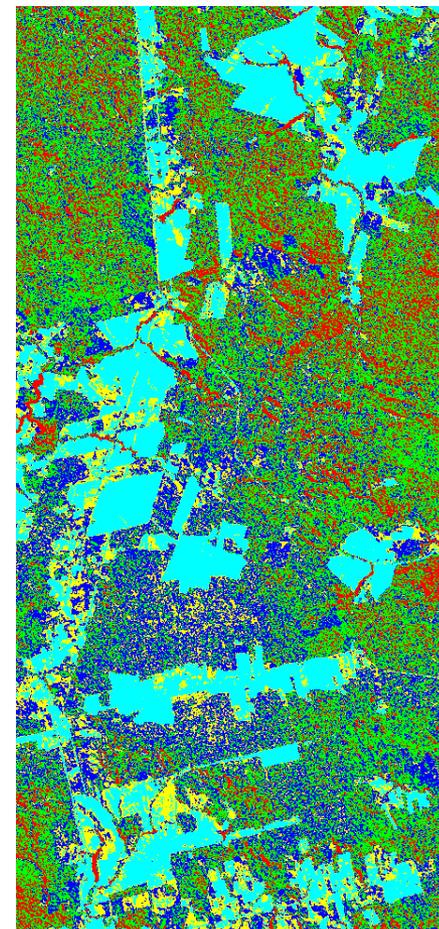
RGB - 543



$K = 3$



$K = 4$



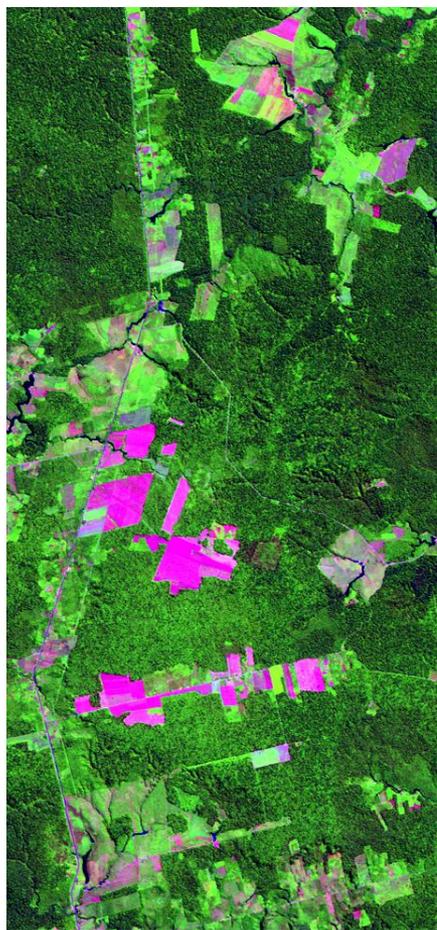
$K = 5$

ISODATA

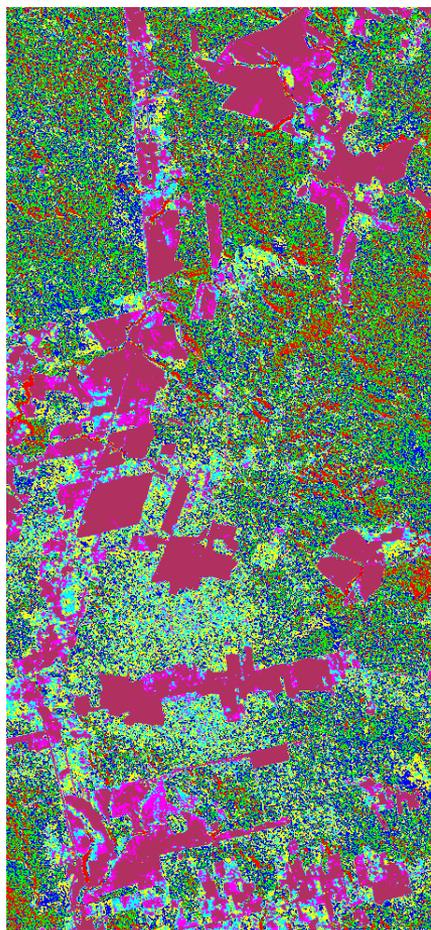
- *Iterative Self-Organizing Data Analysis Technique*
 - Classifica iterativamente os dados
 - Redefinição dos critérios de cada classes
 - Reclassifica a imagem até que os padrões de distância espectral surja
 - Supõe que os agrupamentos no espaço de atributos sejam compactos e esféricos (desvio padrão define a compacidade)
 - Os agrupamentos devem ser separados por uma distância mínima, caso contrário, os agrupamentos são fundidos
 - Se o desvio padrão é maior que um valor estabelecido, os agrupamentos são divididos
 - Maior número de parâmetros em comparação ao *K-Médias*
-

Exemplo – ISODATA

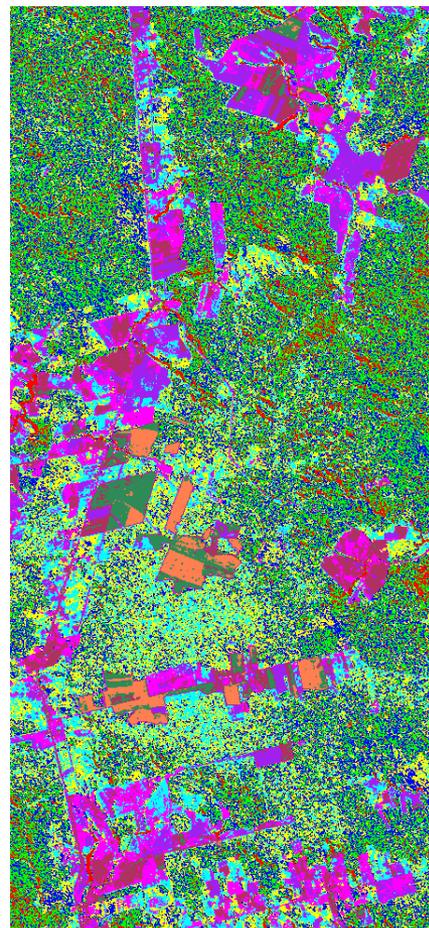
- LANDSAT-5 TM – Set/2010 – FLONA Tapajós-PA



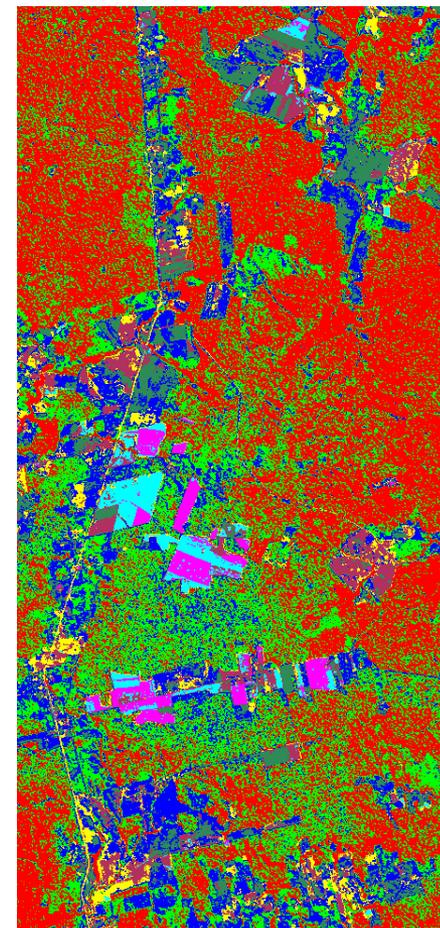
RGB - 543



5 a 10 classes
1 d.p. e dist = 5



5 a 10 classes
2 d.p. e dist = 10



5 a 10 classes
3 d.p. e dist = 15

Bibliografia da aula

• ...

Stu's Views

© Stu All Rights Reserved www.STUS.com

