TÉCNICO LISBOA

Estimar Posições através de Dados com Ruído

Optimização e Algoritmos

*"The world's most valuable resource is no longer oil, but **data**"*

\- *The Economist*, 2017

Visão geral:

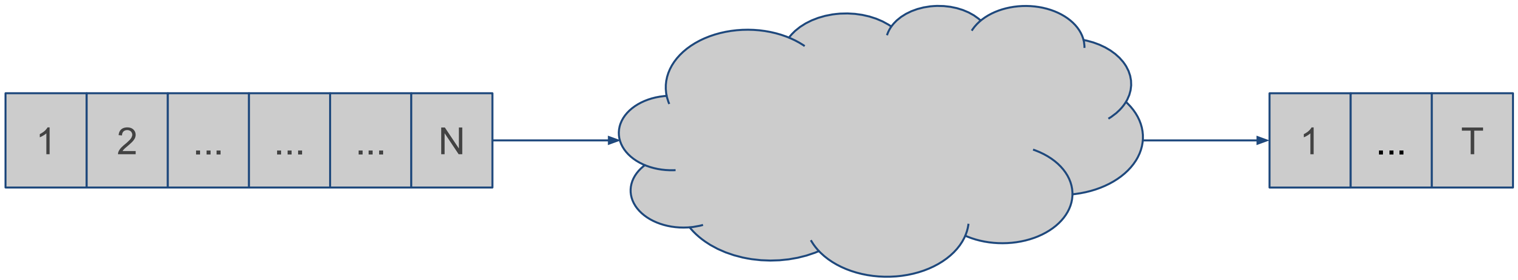
- **Análise da Parte I**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais
- **Análise da Parte II**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais

Visão geral:

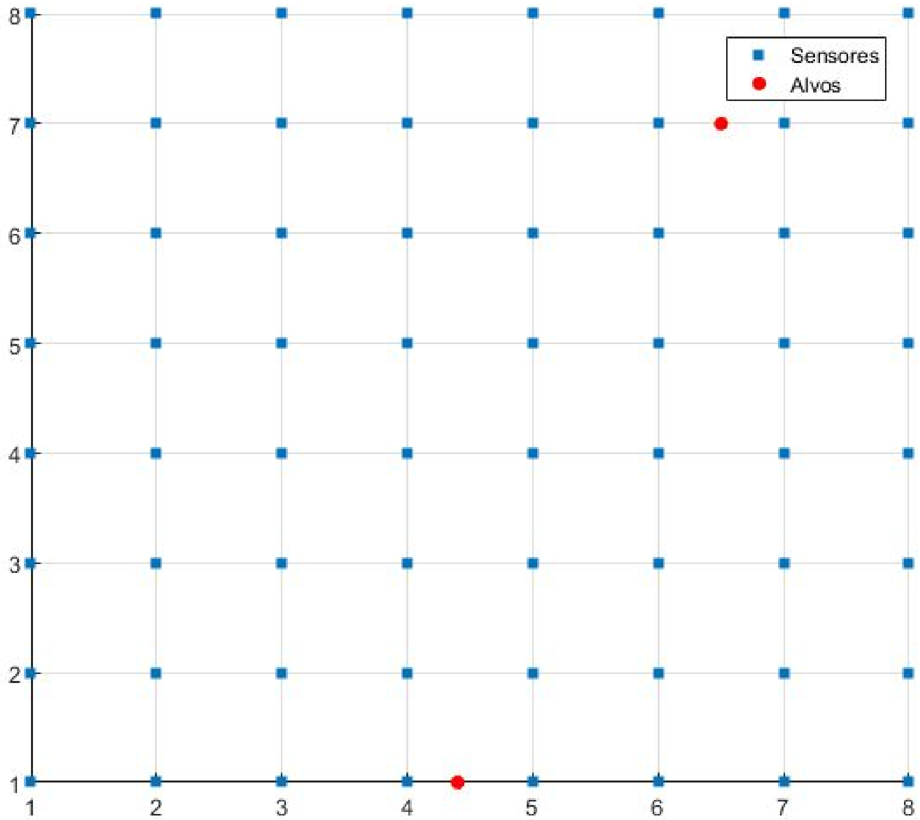
- **Análise da Parte I**
  - **Descrição do Problema**
  - Desenvolvimento
  - Resultados Experimentais
- **Análise da Parte II**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais

Descrição da Parte I

- **N = 64** sensores numa malha 8x8
- Cada sensor (cuja posição é s_n) mede a distância d_n ao alvo mais próximo
- As medições são perturbadas por ruído
- Existem **T = 2** alvos presentes na malha
- **Objectivo:** Descobrir a posição x_n dos **T** alvos



Descrição da Parte I



Visão geral:

- **Análise da Parte I**
  - Descrição do Problema
  - **Desenvolvimento**
  - Resultados Experimentais
- **Análise da Parte II**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais

Desenvolvimento

Formulação do Problema:

$$f(x_1, \ldots, x_n, y_1, \ldots, y_n) = \sum_{n=1}^{N} (\alpha_n^T \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \beta_n)^2 + \rho \sum_{1 \leq n \leq N} \sum_{n+1 \leq m \leq N, m \sim n} \phi\left(\begin{bmatrix} x_n \\ y_n \end{bmatrix}, \begin{bmatrix} x_m \\ y_m \end{bmatrix} \right)$$

Define-se:

$$\alpha_n := \begin{bmatrix} -2s_n \\ 1 \end{bmatrix} \in \mathbf{R}^3, (n = 1, \ldots, N),$$

$$\beta_n := d_n^2 - \|s_n\|^2 \text{ e } \|x_n\|^2 = y_n \text{ para } n = 1, \ldots, N$$

$$x_n \in \mathbf{R}^2 \text{ e } y_n \in \mathbf{R} \text{ para } n = 1, \ldots, N$$

Define-se vizinho :

$$m \sim n \text{ se } ||s_n - s_m|| = 1$$

Desenvolvimento

Formulação do Problema:

$$f(x_1,\ldots,x_n,y_1,\ldots,y_n) = \boxed{\sum_{n=1}^{N}(\alpha_n^T \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \beta_n)^2} + \rho \sum_{1\leq n\leq N} \sum_{n+1\leq m\leq N, m\sim n} \phi\left(\begin{bmatrix} x_n \\ y_n \end{bmatrix}, \begin{bmatrix} x_m \\ y_m \end{bmatrix}\right)$$

Define-se:

$$\alpha_n := \begin{bmatrix} -2s_n \\ 1 \end{bmatrix} \in \mathbf{R}^3, (n = 1,\ldots,N),$$

$$\beta_n := d_n^2 - \|s_n\|^2 \text{ e } \|x_n\|^2 = y_n \text{ para } n = 1,\ldots,N$$

$$x_n \in \mathbf{R}^2 \text{ e } y_n \in \mathbf{R} \text{ para } n = 1,\ldots,N$$

Define-se vizinho :

$$m \sim n \text{ se } ||s_n - s_m|| = 1$$

Desenvolvimento

$$d_n \simeq ||s_n - x_n|| \iff d_n^2 \simeq ||s_n||^2 - 2s_n^T x_n + ||x_n||^2$$

$$\begin{cases} \begin{bmatrix} -2s_n^T & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \simeq d_n^2 - ||s_n||^2 \\ y_n = ||x_n||^2 \end{cases}$$

Desenvolvimento

Função de Custo:

$$f(x_1, \ldots, x_n, y_1, \ldots, y_n) = \sum_{n=1}^{N} (\alpha_n^T \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \beta_n)^2 + \boxed{\rho \sum_{1 \leq n \leq N} \sum_{n+1 \leq m \leq N, m \sim n} \phi\left(\begin{bmatrix} x_n \\ y_n \end{bmatrix}, \begin{bmatrix} x_m \\ y_m \end{bmatrix} \right)}$$

Define-se:

$$\alpha_n := \begin{bmatrix} -2s_n \\ 1 \end{bmatrix} \in \mathbf{R}^3, (n = 1, \ldots, N),$$

$$\beta_n := d_n^2 - \|s_n\|^2 \text{ e } \|x_n\|^2 = y_n \text{ para } n = 1, \ldots, N$$

$$x_n \in \mathbf{R}^2 \text{ e } y_n \in \mathbf{R} \text{ para } n = 1, \ldots, N$$

Define-se vizinho :

$$m \sim n \text{ se } ||s_n - s_m|| = 1$$

Desenvolvimento

Formulação do problema:

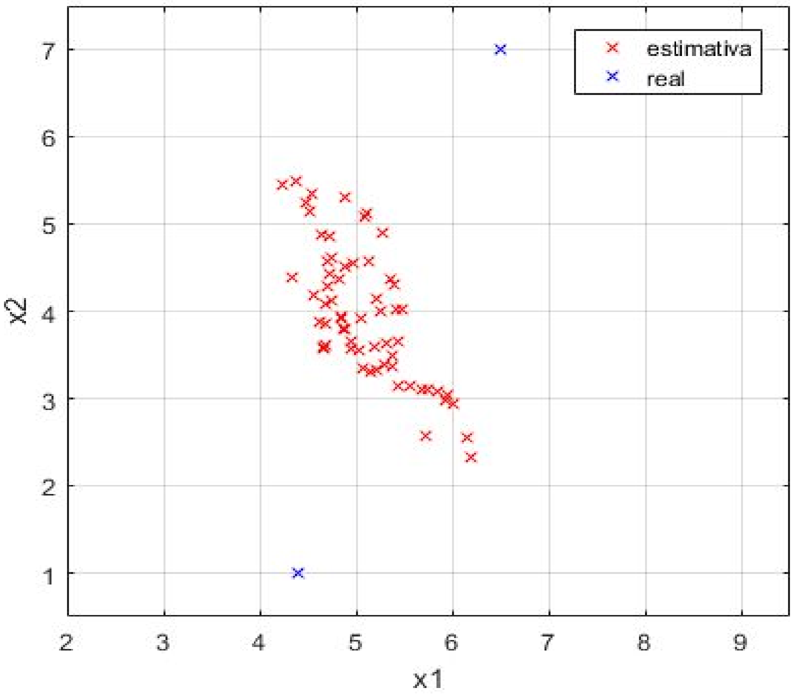
$$\begin{array}{ll} \underset{x_1,\ldots,x_N,y_1,\ldots,y_N}{\text{minimize}} & f(x_1,\ldots,x_n,y_1,\ldots,y_n) \\ \text{subject to} & \|x_n\|^2 \leq y_n, n = 1,\ldots,N \end{array}$$

Desenvolvimento - 1º Passo

$$\phi\left(\begin{bmatrix} x_n \\ y_n \end{bmatrix}, \begin{bmatrix} x_m \\ y_m \end{bmatrix}\right) = \left\| \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \begin{bmatrix} x_m \\ y_m \end{bmatrix} \right\|$$

Resultado - 1º Passo

Solução com $\rho = 10$:



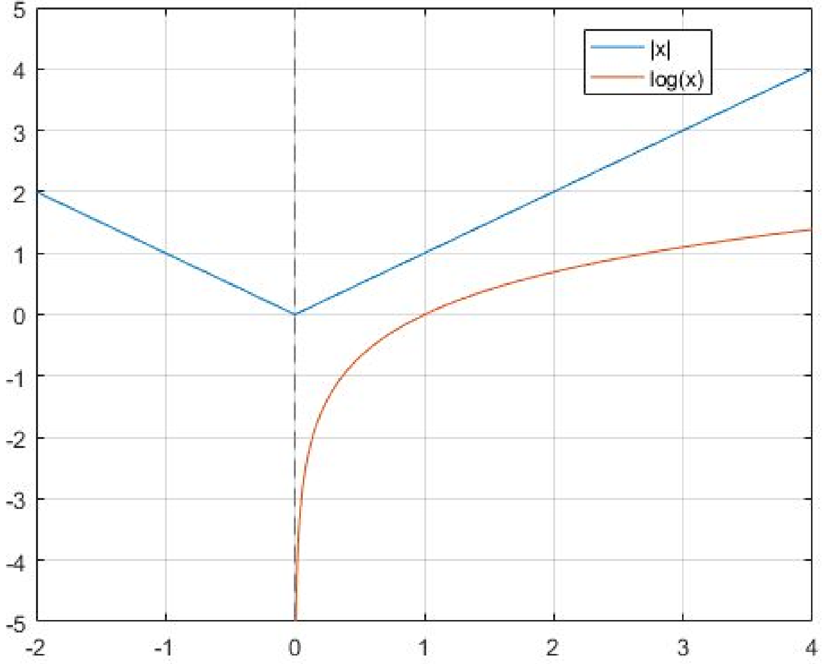
Desenvolvimento - 2º Passo

Maior lucro para estimativas concordantes de sensores vizinhos:

$$\phi\left(\begin{bmatrix} x_n \\ y_n \end{bmatrix}, \begin{bmatrix} x_m \\ y_m \end{bmatrix}\right) = log\left\|\begin{bmatrix} x_n \\ y_n \end{bmatrix} - \begin{bmatrix} x_m \\ y_m \end{bmatrix}\right\|$$

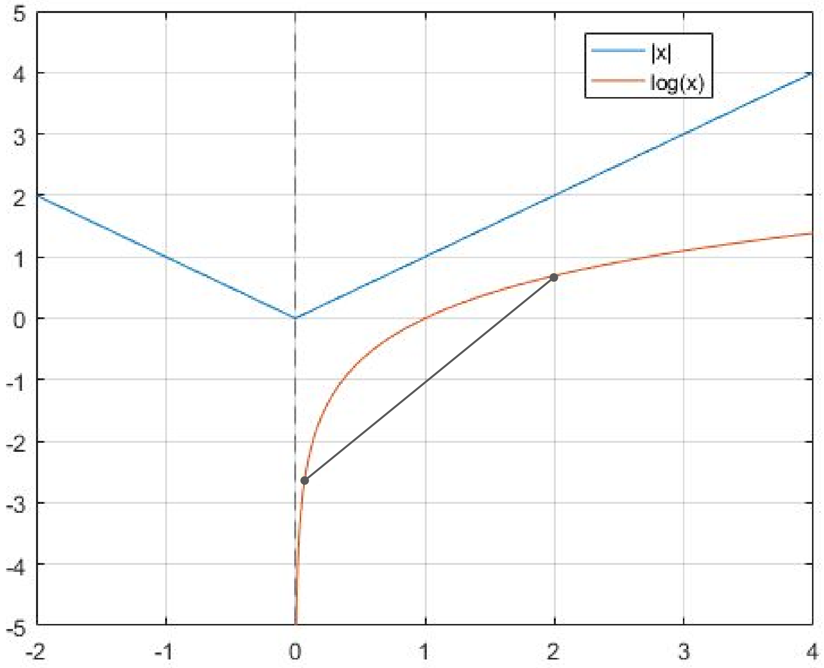
Desenvolvimento - 2º Passo

Exemplo em $\mathbb{R}$:



Desenvolvimento - 2º Passo

Exemplo em $\mathbb{R}$:



logaritmo é não convexo

Desenvolvimento - 2º Passo

Método MM para transformar o problema não convexo num processo de linearização iterativo:

$$(x_n^{r+1}, y_n^{r+1})_{n=1}^N = \underset{x_1,\ldots,x_N,y_1,\ldots,y_N}{\operatorname{argmin}} \quad f^{(r)}(x_1,\ldots,x_n,y_1,\ldots,y_n)$$

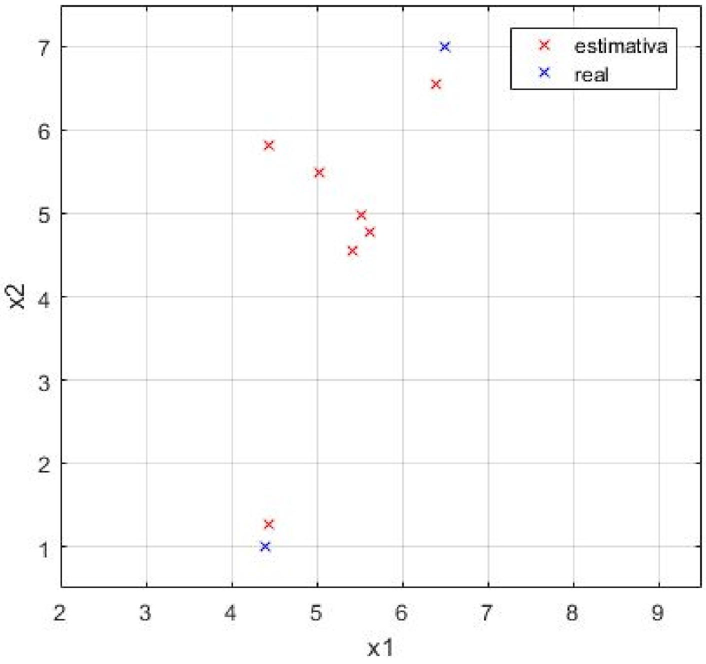
$$\text{subject to } \|x_n\|^2 \leq y_n, n = 1,\ldots,N,$$

$$f^{(r)}(x_1,\ldots,x_n,y_1,\ldots,y_n) = \sum_{n=1}^{N}(\alpha_n^T \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \beta_n)^2 + \rho \sum_{1\leq n\leq N} \sum_{n+1\leq m\leq N, m\sim n} \left(\left\| \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \begin{bmatrix} x_m \\ y_m \end{bmatrix} \right\|\right) \cdot \left(\left\| \begin{bmatrix} x_n^{(r)} \\ y_n^{(r)} \end{bmatrix} - \begin{bmatrix} x_m^{(r)} \\ y_m^{(r)} \end{bmatrix} \right\| + \epsilon\right)^{-1}$$

$\simeq 0$

Resultados - 2º Passo

Solução com $\rho = 10$:



Desenvolvimento - 3º Passo

Aperfeiçoamento dos resultados:

- Seleciona-se 2 pontos, por exemplo: $(\bar{x}_p, \bar{y}_p)$ e $(\bar{x}_q, \bar{y}_q)$
- Divide-se as estimativas em 3 conjuntos: os pontos mais próximos de $(\bar{x}_p, \bar{y}_p)$ e $(\bar{x}_q, \bar{y}_q)$
- Utiliza-se cada conjunto para determinar um ponto, a partir de:

$$(\hat{x}_t, \hat{y}_t) := \underset{x,y}{\operatorname{argmin}} \quad \textstyle\sum_{n=1, n\in\Omega_t}^{N} \left(\alpha_n^T \begin{bmatrix} x \\ y \end{bmatrix} - \beta_n\right)^2$$

$$\text{subject to} \qquad \|x\|^2 \leq y,$$

- Determinar o erro das 2 estimativas obtidas:

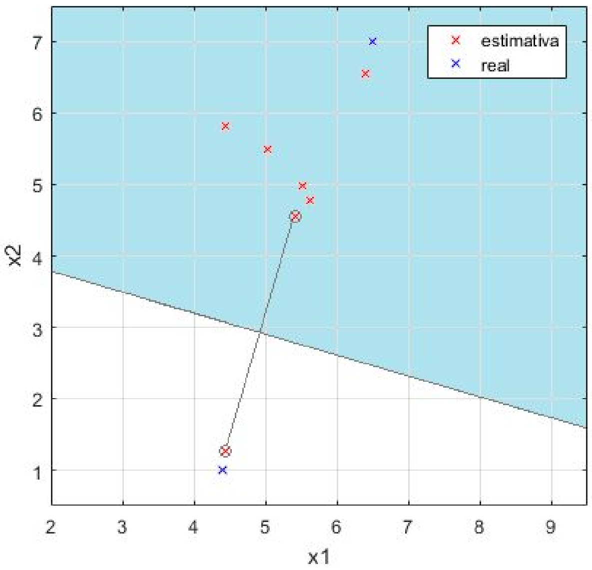
$$e := \textstyle\sum_{t=1}^{T} \sum_{n=1, n\in\Omega_t}^{N} \left(\|s_n - \hat{x}_t\| - d_n\right)^2.$$

- Repete-se o processo para todas as outras combinações de $(\bar{x}, \bar{y})$
- Escolhe-se as 2 estimativas que garantam menor erro

Desenvolvimento - 3º Passo

Das K estimativas obtidas $(\bar{x_k}, \bar{y_k})$ escolhem-se 2 (número de alvos) e define-se:

$$\Omega_1 := \{n = 1, .., N : \|s_n - \bar{x_1}\| \leq \|s_n - \bar{x_2}\|\}$$
$$\Omega_2 := \{n = 1, .., N : \|s_n - \bar{x_2}\| \leq \|s_n - \bar{x_1}\|\}$$



Desenvolvimento - 3º Passo

Para cada conjunto Ω_t, t={1,2} efetua-se a minimização:

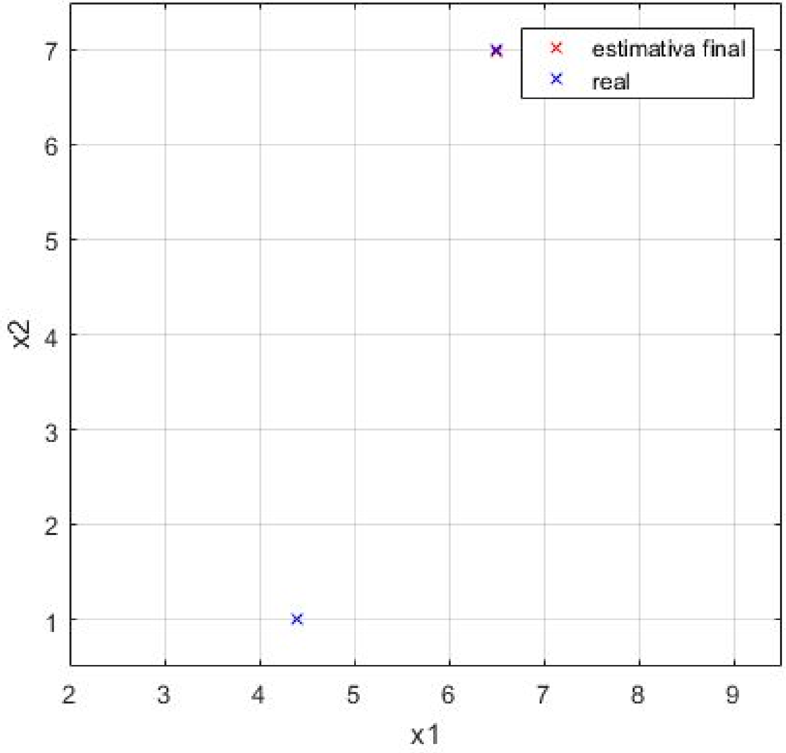
$$(\hat{x}_t, \hat{y}_t) := \underset{x,y}{\operatorname{argmin}} \quad \sum_{n=1, n\in\Omega_t}^{N} \left(\alpha_n^T \begin{bmatrix} x \\ y \end{bmatrix} - \beta_n\right)^2$$

$$\text{subject to} \qquad \|x\|^2 \leq y,$$

Critério de escolha das estimativas finais:

$$e := \sum_{t=1}^{T} \sum_{n=1, n\in\Omega_t}^{N} (\|s_n - \hat{x}_t\| - d_n)^2 .$$

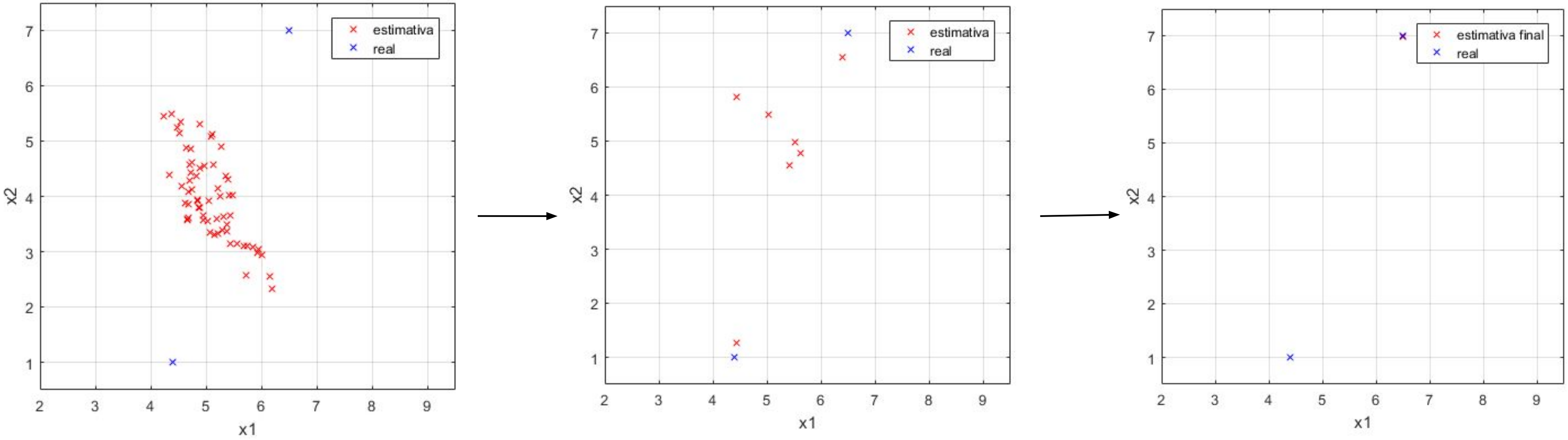
Desenvolvimento - 3º Passo



Visão geral:

- **Análise da Parte I**
  - Descrição do Problema
  - Desenvolvimento
  - **Resultados Experimentais**
- **Análise da Parte II**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais

Resultados Experimentais



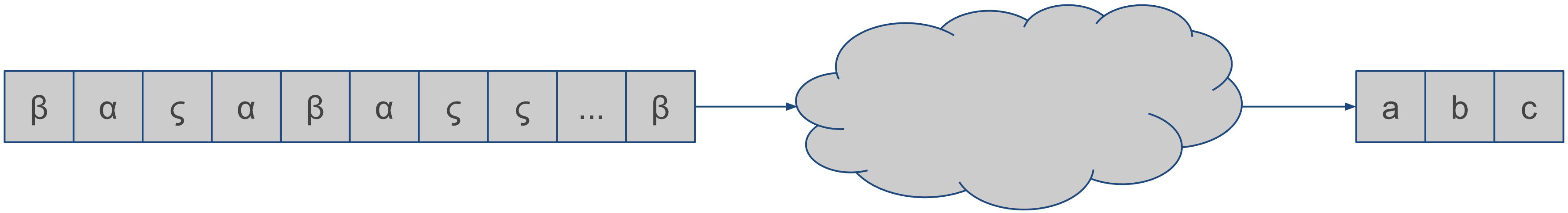
$$\textbf{\textit{Mean Squared Error (MSE)}} = \frac{1}{N}\sum_{n=1}^{N}\left|\left|x_n - x_{real}\right|\right|^2 = 0.001$$

Visão geral:

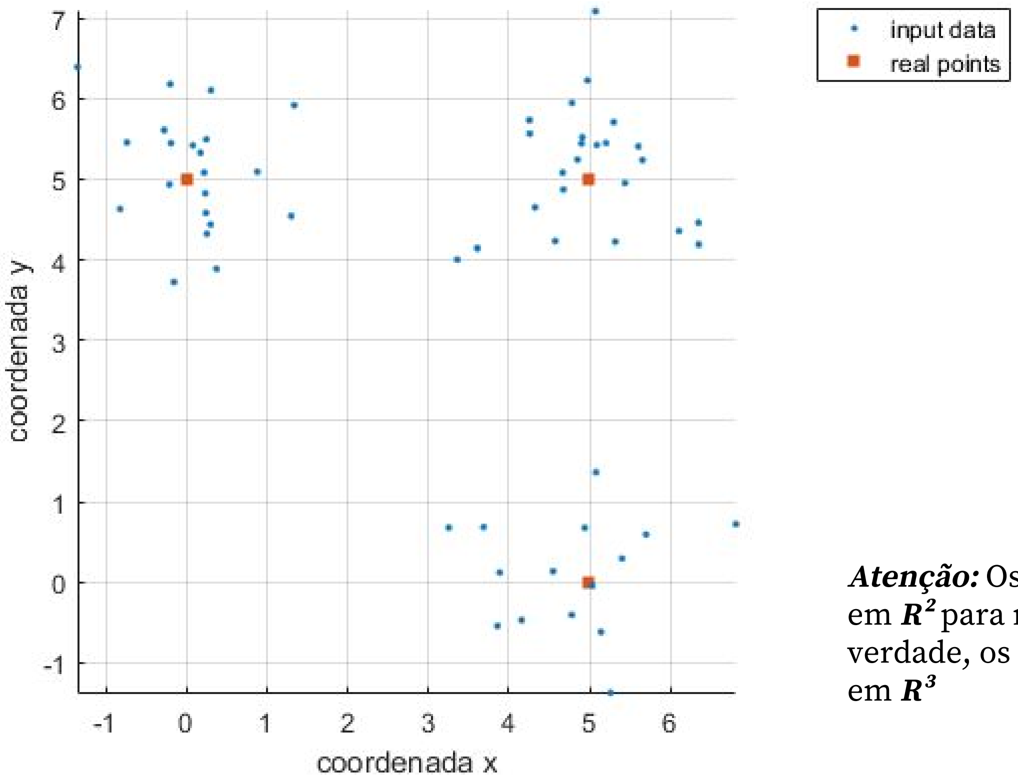
- **Análise da Parte I**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais
- **Análise da Parte II**
  - **Descrição do Problema**
  - Desenvolvimento
  - Resultados Experimentais

Descrição da Parte II

- ***a***, ***b*** e ***c*** são três pontos (ou classes) definidos em R^3
- **N = 60** pontos uniformemente distribuídos entre ***a***, ***b*** e ***c***
- Cada um sujeito a ruído gaussiano (média nula), com fator multiplicativo ξ
- Denomina-se os N pontos por ***α***, ***β*** e ***ς***, caso tenham origem em ***a***, ***b*** e ***c***, respectivamente
- **Objectivo:** Estimar a posição de ***a***, ***b*** e ***c***



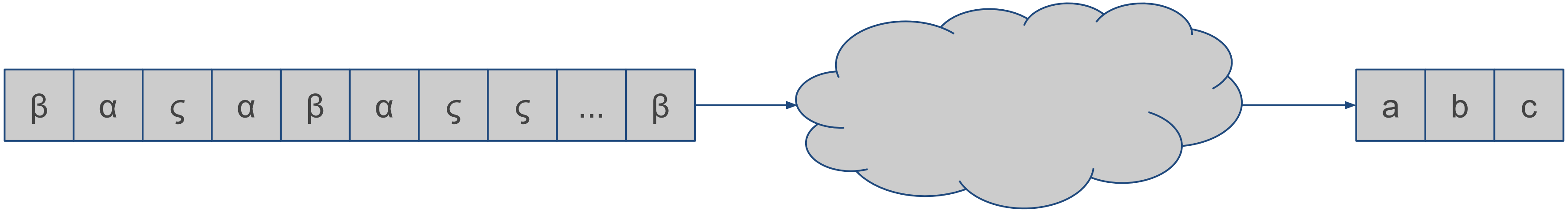
Descrição da Parte II



Atenção: Os dados foram expostos em $\boldsymbol{R^2}$ para melhor visualização. Na verdade, os pontos estão definidos em $\boldsymbol{R^3}$

Descrição da Parte II

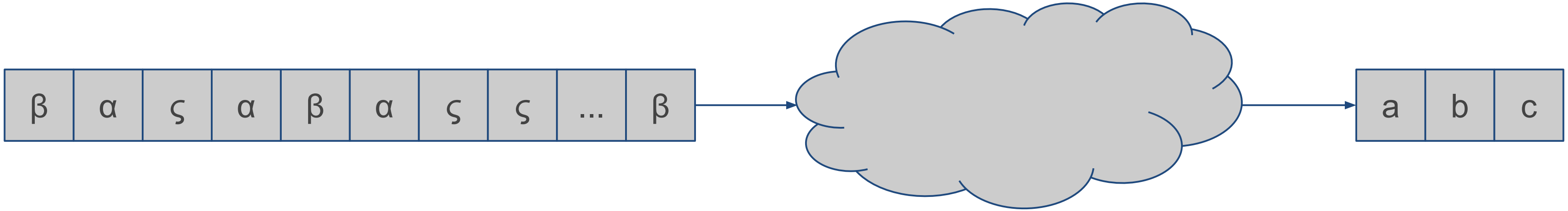
- ***a***, ***b*** e ***c*** são três pontos (ou classes) definidos em R^3
- **N = 60** pontos uniformemente distribuídos entre ***a***, ***b*** e ***c***
- Cada um sujeito a ruído gaussiano (média nula), com fator multiplicativo ξ
- Denomina-se os N pontos por ***α***, ***β*** e ***ς***, caso tenham origem em ***a***, ***b*** e ***c***, respectivamente
- **Objectivo:** Estimar a posição de ***a***, ***b*** e ***c***



Descrição da Parte II

a_k

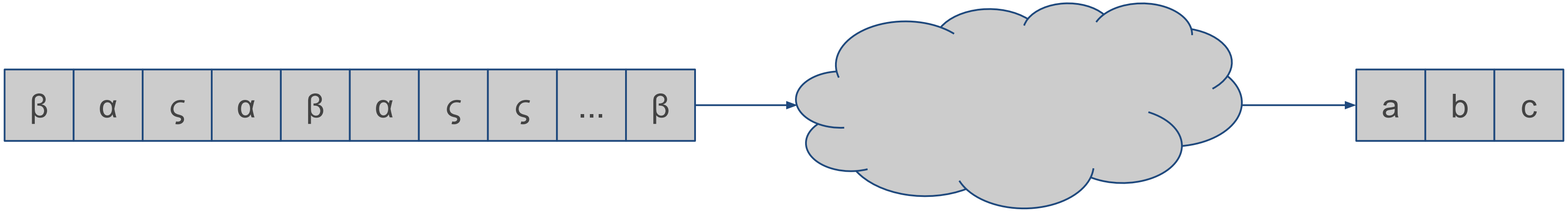
- ***a***, ***b*** e ***c*** são três pontos (ou classes) definidos em R^3
- **N = 60** pontos uniformemente distribuídos entre ***a***, ***b*** e ***c***
- Cada um sujeito a ruído gaussiano (média nula), com fator multiplicativo ξ
- Denomina-se os N pontos por ***α***, ***β*** e ***ς***, caso tenham origem em ***a***, ***b*** e ***c***, respectivamente
- **Objectivo:** Estimar a posição de ***a***, ***b*** e ***c***



Descrição da Parte II

a_k

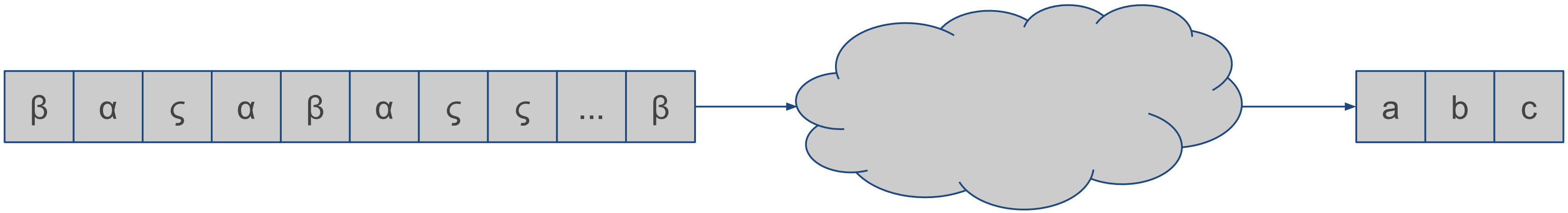
- ***a***, ***b*** e ***c*** são três pontos (ou classes) definidos em R^3
- **N = 60** pontos uniformemente distribuídos entre ***a***, ***b*** e ***c*** $y_n = a_k + ruído$
- Cada um sujeito a ruído gaussiano (média nula), com fator multiplicativo ξ
- Denomina-se os N pontos por ***α***, ***β*** e *ς*, caso tenham origem em ***a***, ***b*** e ***c***, respectivamente
- **Objectivo:** Estimar a posição de ***a***, ***b*** e ***c***



Descrição da Parte II

a_k

- ***a***, ***b*** e ***c*** são três pontos (ou classes) definidos em R^3
- **N = 60** pontos uniformemente distribuídos entre ***a***, ***b*** e ***c*** $y_n = a_k + ruído$
- Cada um sujeito a ruído gaussiano (média nula), com fator multiplicativo ξ
- Denomina-se os N pontos por ***α***, ***β*** e ***ς***, caso tenham origem em ***a***, ***b*** e ***c***, respectivamente
- **Objectivo:** Estimar a posição de ***a***, ***b*** e ***c*** a_n



Descrição da Parte II

Geração do ruído:

Ruído gaussiano de média nula

$$\boxed{r = \text{randn}()}$$

ξ - Factor multiplicativo que permite controlar a dispersão dos pontos

Desenvolvimento

Formulação do Problema:

$$\underset{a_1,\ldots,a_N}{\text{minimize}} \quad \sum_{n=1}^{N} \|y_n - a_n\|^2 + \rho \sum_{1\leq n\leq N} \sum_{n+1\leq m\leq N, m\in\Theta_n} \phi(a_n, a_m)$$

Define-se:

Θ_n - Conjunto de **K** pontos na vizinhança do ponto n.

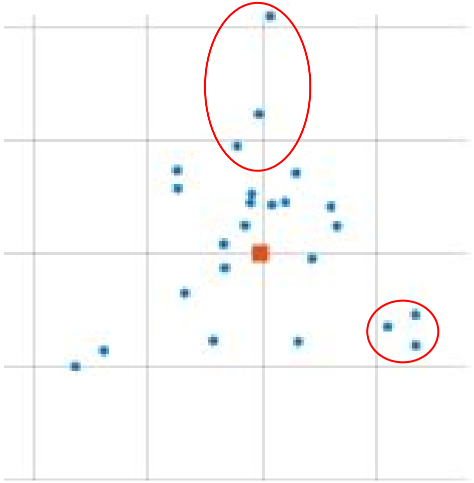
Desenvolvimento

Formulação do Problema:

$$\underset{a_1,\ldots,a_N}{\text{minimize}} \quad \sum_{n=1}^{N} \left\| y_n - a_n \right\|^2 + \rho \sum_{1\leq n\leq N} \sum_{n+1\leq m\leq N, m\in\Theta_n} \phi\left(a_n, a_m\right)$$

Define-se:

Θ_n - Conjunto de **K** pontos na vizinhança do ponto n.



Ex: K = **2** vizinhos

Desenvolvimento - 1º Passo

$$\phi\left(a_n, a_m\right) = \left\|a_n - a_m\right\|$$

Desenvolvimento - 2º Passo

$$\phi\,(a_n\,,\,a_m) = log||a_n\ -\ a_m||$$

Desenvolvimento - 2º Passo

Processo iterativo:

$$(a_n^{r+1}, a_m^{r+1})_{n=1}^N =$$

$$\underset{x_1,\ldots,x_N,y_1,\ldots,y_N}{\operatorname{argmin}} \sum_{n=1}^N \|y_n - a_n\|^2 + \rho \sum_{1\leq n\leq N} \sum_{n+1\leq m\leq N, m\in\Theta_n} \frac{\|a_n - a_m\|}{\|a_n^{(r)} - a_m^{(r)}\| + \epsilon}$$

Desenvolvimento - 3º Passo

Aperfeiçoamento dos resultados:

- Seleciona-se 3 pontos, por exemplo: $\bar{a}_1$, $\bar{a}_2$ e $\bar{a}_3$
- Divide-se as estimativas em 3 conjuntos: os pontos mais próximos de $\bar{a}_1$, $\bar{a}_2$ e $\bar{a}_3$
- Utiliza-se cada conjunto para determinar um ponto, a partir de:

$$\hat{a}_t := \underset{a}{\operatorname{argmin}} \quad \sum_{n=1, n\in\Omega_t}^{N} \|y_n - a\|^2, t = 1, \ldots, T$$

- Determinar o erro das 3 estimativas obtidas:

$$e := \sum_{t=1}^{T} \sum_{n=1, n\in\Omega_t}^{N} \|y_n - \hat{a}_t\|^2.$$

- Repete-se o processo para todas as outras combinações de $\bar{a}_t$
- Escolhe-se as 3 estimativas que garantam menor erro

Desenvolvimento - 3º Passo

Resumindo:

1. Na primeira iteração aproximamos as estimativas os K vizinhos com

$$\phi\left(a_n, a_m\right) = \left\|a_n - a_m\right\|$$

2. Nas restantes iterações beneficiamos ainda mais a aproximação com o processo iterativo usando

$$\phi^{(r)}\left(a_n, a_m\right) = \frac{\left\|a_n - a_m\right\|}{\left\|a_n^{(r)} - a_m^{(r)}\right\| + \epsilon}$$

3. Por fim, determinamos os conjuntos Ω de cada 3 pontos e escolhemos as melhores estimativas

Visão geral:

- **Análise da Parte I**
  - Descrição do Problema
  - Desenvolvimento
  - Resultados Experimentais
- **Análise da Parte II**
  - Descrição do Problema
  - Desenvolvimento
  - **Resultados Experimentais**

Resultados Experimentais

Dependência nos parâmetros *(ρ, # vizinhos)*

- Conjunto com pontos originais equidistantes:

$$a_1 = (x, 0, 0)$$
$$a_2 = (0, x, 0) \qquad x = \{ 2, 5 \}$$
$$a_3 = (0, 0, x)$$

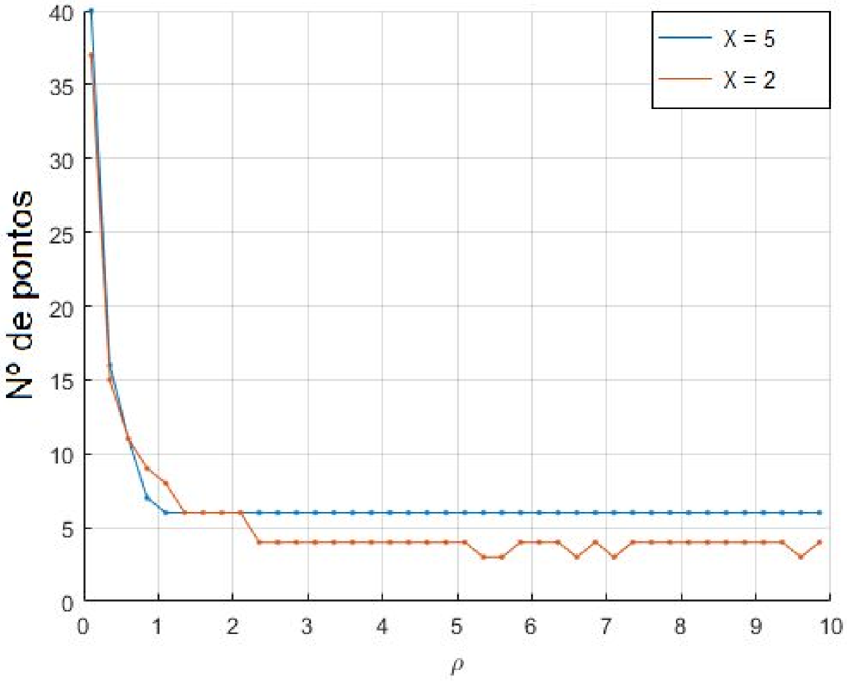
- $\xi = 0.75$

- Análise do número de diferentes a_n encontrados (1^a fase do método).

- Em 5 iterações.

Resultados Experimentais

Variação com ρ para K=5 vizinhos:



Resultados Experimentais

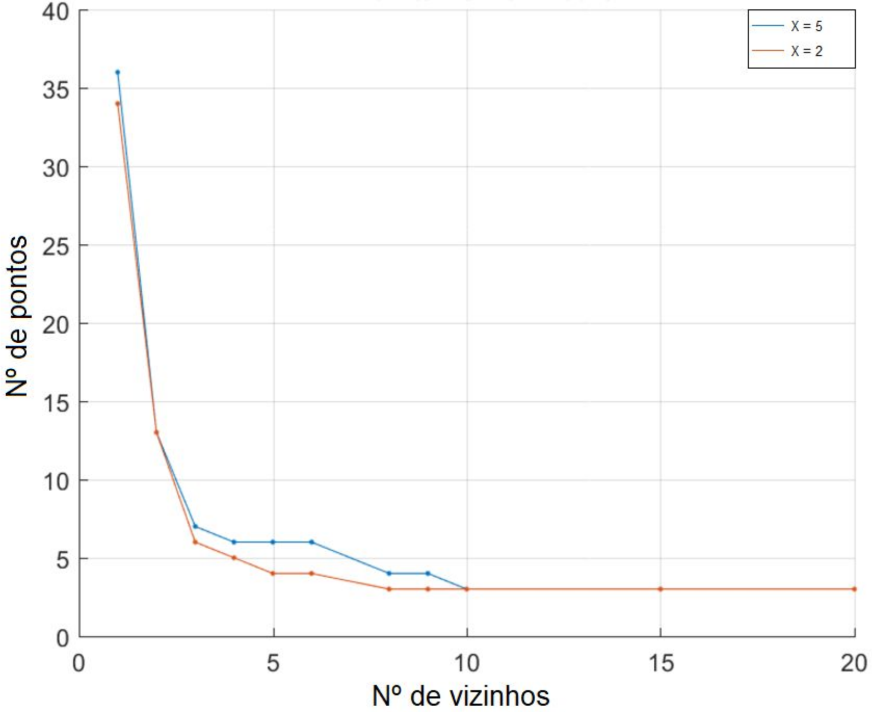
Desempenho do método

- Variação do erro com a configuração dos pontos originais
- Comparação com algoritmo de clustering, ***K-means***
- Conjunto com pontos originais equidistantes:

$$\begin{aligned} a_1 &= (x, 0, 0) \\ a_2 &= (0, x, 0) \\ a_3 &= (0, 0, x) \end{aligned}$$

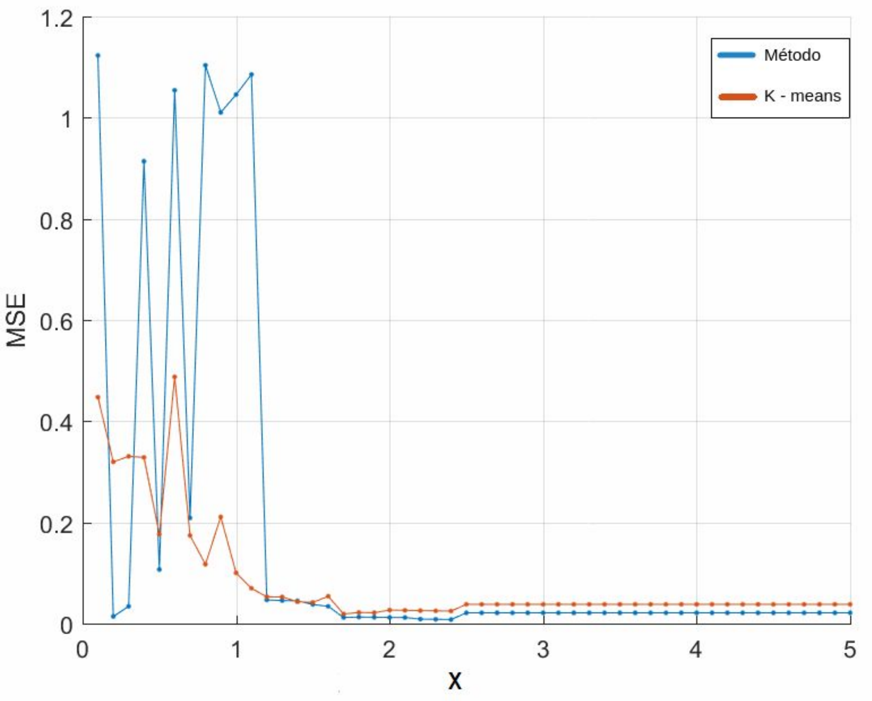
Resultados Experimentais

Variação com K para ρ =5:

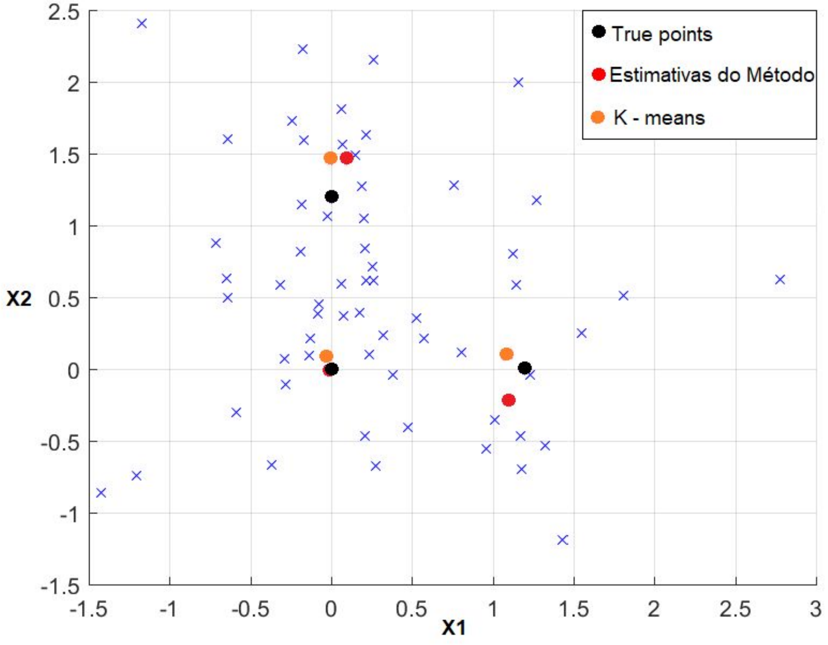


Resultados Experimentais

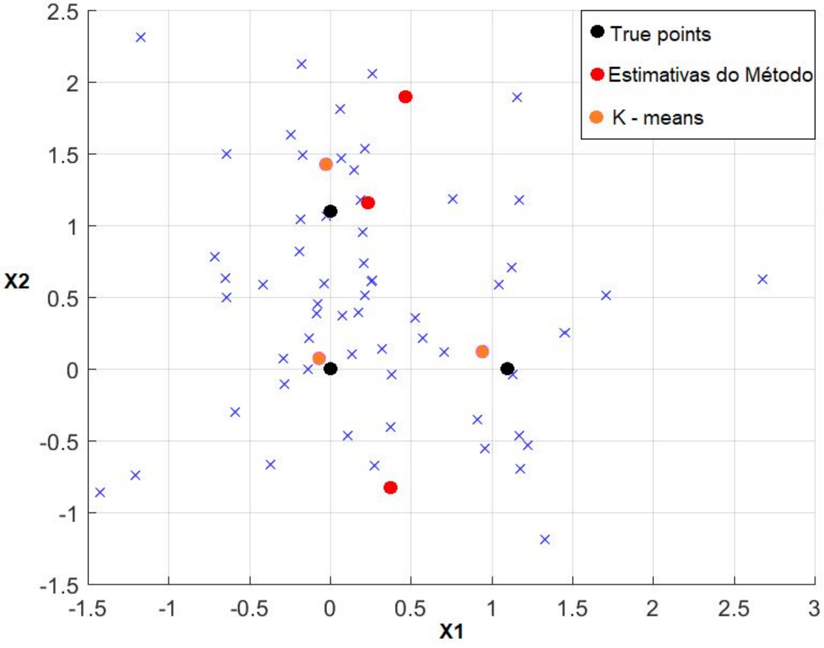
Variação com a distância x usando $\rho = 5, K = 5, \xi = 0.65$ **:**



Resultados Experimentais



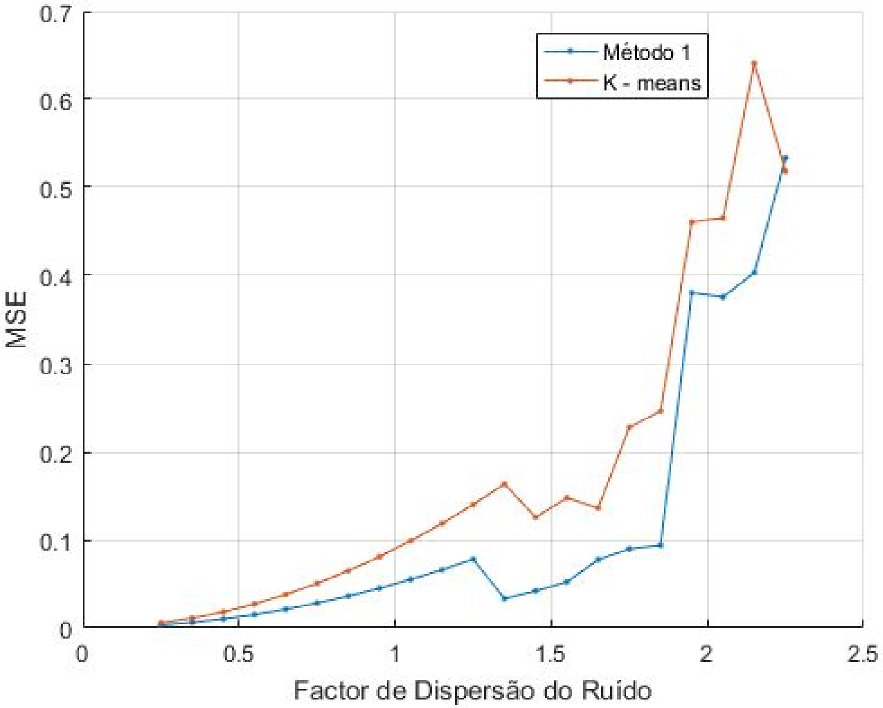
x = 1.2



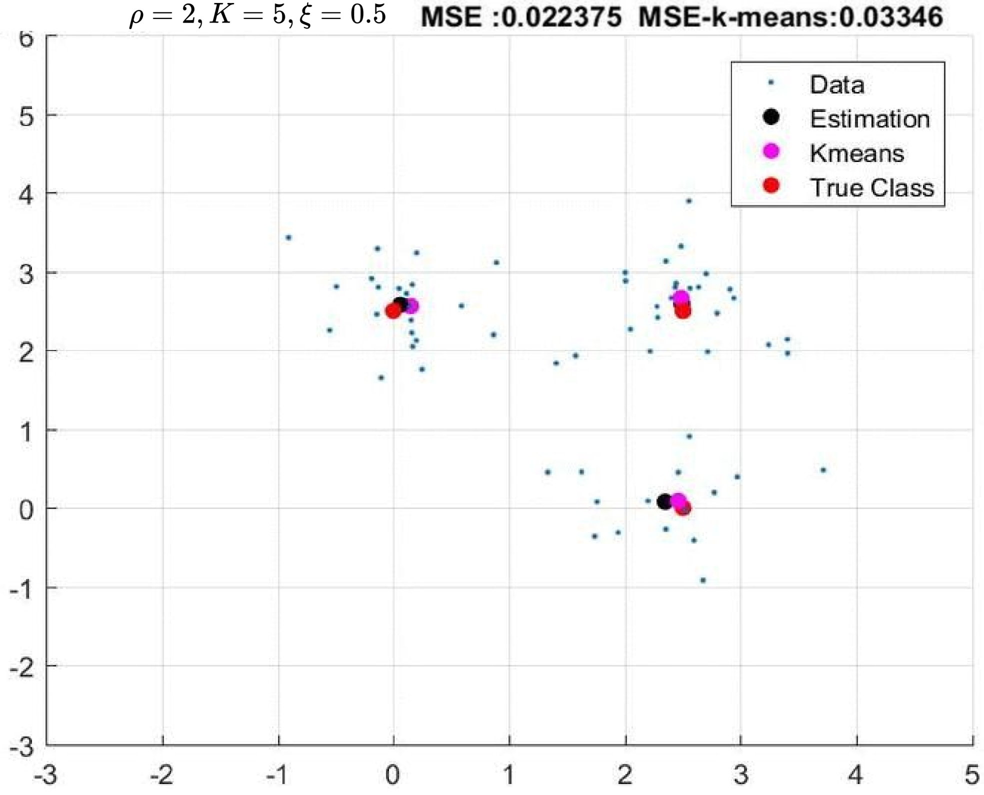
x = 1.1

Resultados Experimentais

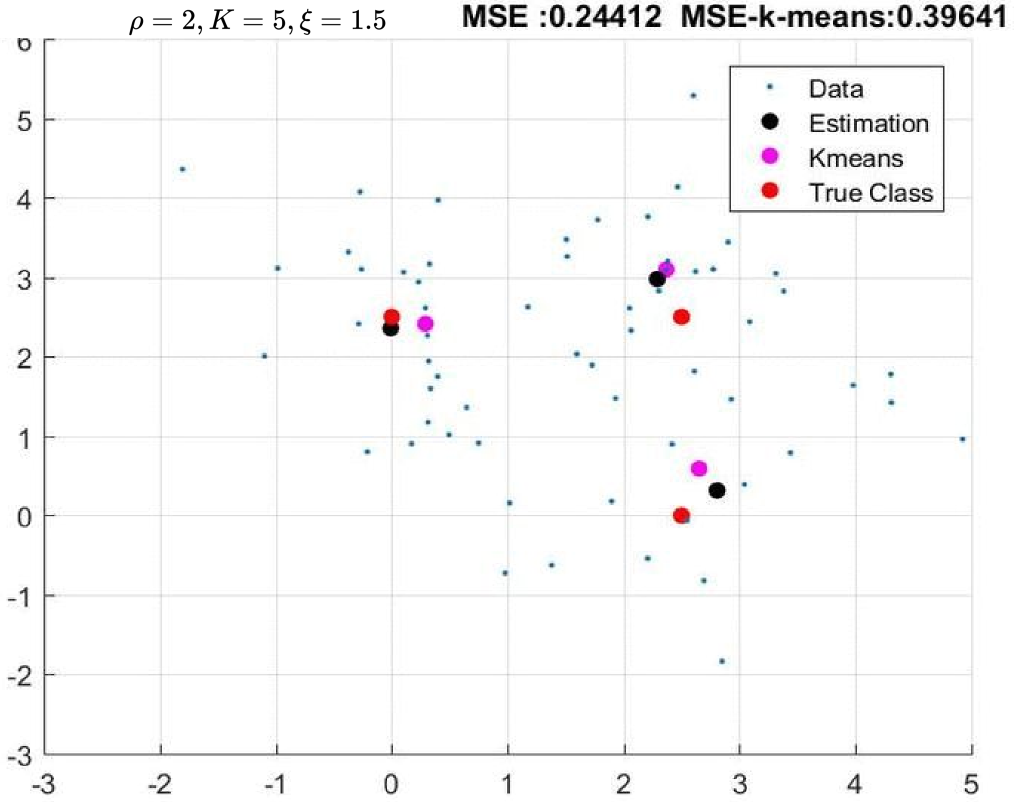
Variação com o factor ξ utilizando $\rho = 5, K = 5$ e x = 5:



Resultados Experimentais



Resultados Experimentais



Trabalho Realizado por:

- Manuel Reis, nº 81074 - manuel.b.reis@tecnico.ulisboa.pt
- Mariana Galrinho, nº 81669 - mariana.galrinho@tecnico.ulisboa.pt
- Renato Henriques, nº 81588 - renato.henriques@tecnico.ulisboa.pt

Grupo 15

"The world's most valuable resource is no longer oil, but data"

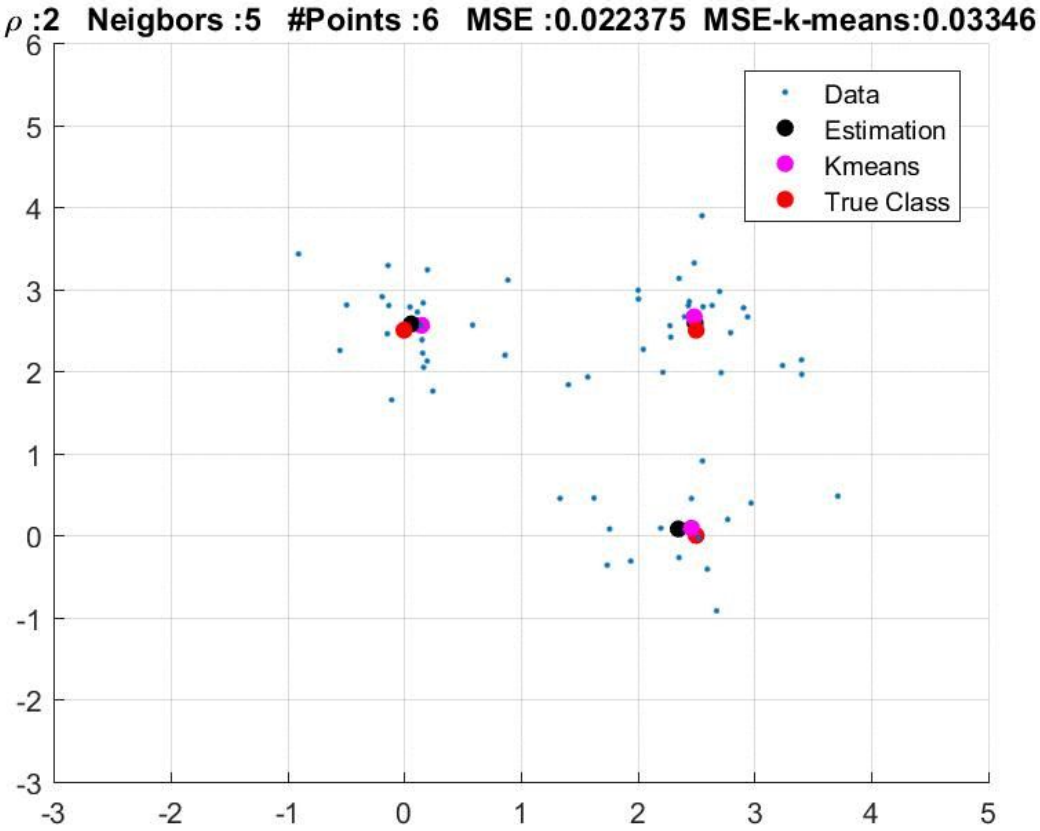
- The Economist, 2017

Passo 4?

Resultados Experimentais

Influência da configuração
0.5 ruído

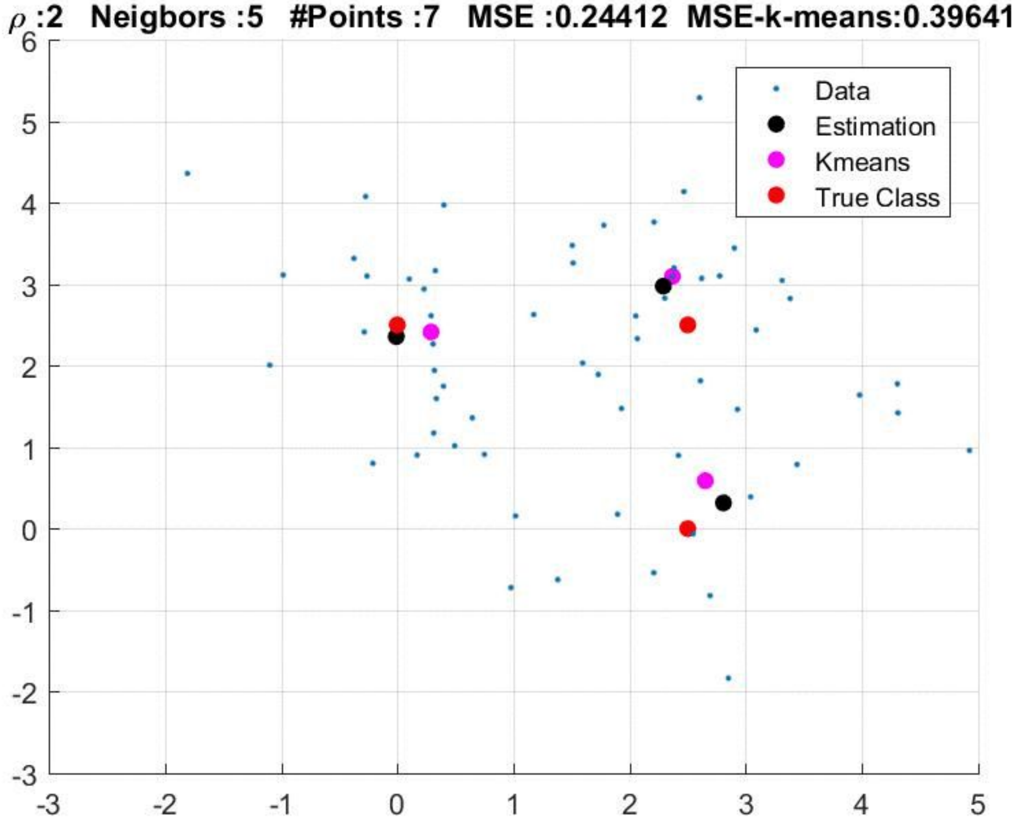
erro máximo 0.094



Resultados Experimentais

Influência da configuração
1.5 ruído

erro maximo 0.3



$$f(x_1,\ldots,x_n,y_1,\ldots,y_n) \;=\sum_{n=1}^{N}(\alpha_n^T\begin{bmatrix}x_n\\ y_n\end{bmatrix}-\beta_n)^2+\;\rho\sum_{1\leq n\leq N}\;\sum_{n+1\leq m\leq N,m\sim n}\phi\left(\begin{bmatrix}x_n\\ y_n\end{bmatrix},\begin{bmatrix}x_m\\ y_m\end{bmatrix}\right)$$

quadrática co