

Inferência Topológica Computacional

Nuno Amaral Picado



Inferência Topológica Computacional

Nuno Amaral Picado

Dissertação para a obtenção do Grau de **Mestre em Matemática**
Área de Especialização em **Estatística, Otimização e Matemática Financeira**

Júri

Presidente: Dr.^a Maria Esmeralda Gonçalves

Orientador: Dr. Paulo Eduardo Oliveira

Vogal: Dr. João Eduardo Gouveia

Data: Julho de 2016

Resumo

Neste trabalho irão ser estudados métodos de reconstrução não apenas de um conjunto compacto mas também da sua topologia. No início serão introduzidas algumas noções e propriedades importantes, para serem de seguida estudados e simulados dois métodos de recuperação geométrica. No final, iremos libertar-nos da representação geométrica e estudaremos alguns invariantes topológicos que nos permitirão recuperar características do conjunto inicial.

Palavras Chave: Análise de dados, Reconstrução, Topologia, Função distância, Invariantes Topológicos

Abstract

In this paper we will study reconstruction methods not only for compact sets but also for its topology. We start by introducing some important notions and properties in order to study and simulate two geometric reconstruction methods. In the end, we will free ourselves from the geometric representation and present some topological invariants that will allow us to recover some characteristics of the inicial set.

Keywords: Data Analysis, Reconstruction, Topology, Distance function, Topological Invariants

Agradecimentos

Ao Professor Doutor Paulo Oliveira, pela sua ajuda na realização deste trabalho, pela sua motivação mas principalmente por tudo o que me ensinou neste último ano.

À Luísa, tenho de agradecer pela confiança que tem em mim, por sempre me ajudar em tudo o que precise mas, essencialmente, por todo o apoio prestado nas alturas mais difíceis.

A toda a minha família, por sempre ter acreditado em mim, apoiando e incentivando-me a seguir este percurso.

A todos os meus amigos, agradeço por todos os momentos de lazer, sempre importantes, que me proporcionaram nestes anos.

Conteúdo

1	Introdução	1
2	Definições e propriedades básicas	3
3	Funções de distância	9
3.1	Topologia dos offsets	10
3.1.1	Estabilidade de pontos críticos	12
3.2	Reconstrução	17
4	Distância a medidas de probabilidade	19
4.1	Distância de Wasserstein	20
4.2	Distância a uma medida de probabilidade	22
4.3	Aplicações a inferência geométrica	24
5	Simulação em R	27
6	Inferência Homológica	37
6.1	Algoritmo para calcular os números de Betti	38
6.2	Persistência topológica	40
6.2.1	Persistência de uma função	40
6.2.2	Persistência topológica de uma filtração	41
A	Códigos em R	45

Capítulo 1

Introdução

A nuvem de pontos seguinte consiste nas observações disponíveis de um conjunto do plano. Admite-se ainda que a localização de alguns dos pontos foi sujeita a algum erro. Esta nuvem de observações será, no que se segue, designada por amostra.

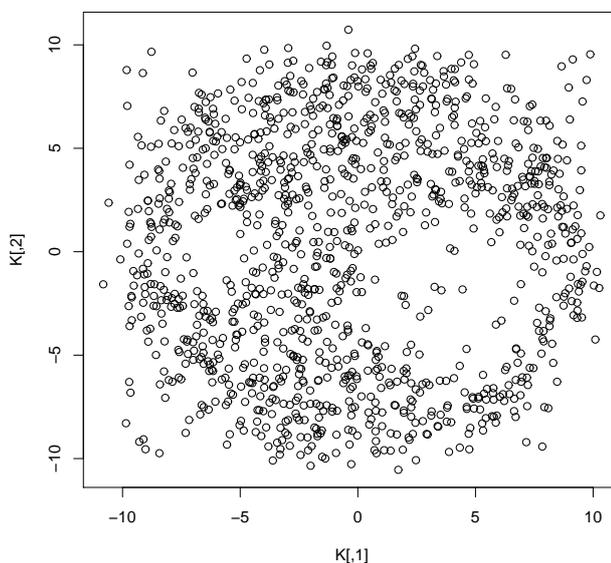


Figura 1.1 Amostra de um conjunto desconhecido

Será que a partir deste conjunto de pontos conseguimos recuperar a forma do conjunto e a sua topologia do conjunto de onde este foi retirado? O que acontece se a amostra possuir ruído e/ou outliers, isto é, pontos isolados?

A presente dissertação apresenta algumas respostas às questões anteriores, privilegiando métodos que nos libertem das representações geométricas, isto é, que permitam uma abordagem ao problema em espaços de dimensão maior do 3. O segundo capítulo contém uma descrição de estruturas que permitem a caracterização de proximidade topológica entre conjuntos. Discute-se também a caracterização de representações de conjuntos a partir de uma nuvem de pontos. No terceiro e quarto capítulos focamo-nos essencialmente na reconstrução geométrica do conjunto, apre-

sentando no final de cada um deles um método que nos irá permitir reconstruir um conjunto com a mesma topologia do inicial. De seguida, irá ser feito um estudo de simulação, onde iremos retirar uma amostra de um conjunto conhecido e usar os dois métodos referidos para tentar recuperar a topologia do conjunto inicial. No entanto, estes métodos estão limitados a conjuntos tridimensionais visto serem os únicos passíveis de representação geométrica. Para dimensões superiores, não conseguindo obter uma representação geométrica, tentaremos recuperar a topologia do conjunto através de invariantes topológicos, tais como os números de Betti e gráficos de persistência. Este assunto é aprofundado no último capítulo.

Capítulo 2

Definições e propriedades básicas

Para iniciar este trabalho, vão ser apresentadas algumas definições e propriedades relativamente básicas, mas bastante importantes para o objetivo deste trabalho. Primeiramente, visto estarmos a tentar recuperar a topologia de um conjunto, faz sentido definirmos quando é que dois conjuntos têm a mesma topologia. As próximas quatro definições, que dizem respeito à comparação entre topologias de dois conjuntos, são apresentadas nesse sentido.

Definição 2.1. *Dois espaços topológicos X e Y dizem-se **homeomorfos** se existir uma aplicação bijectiva contínua $h : X \rightarrow Y$ tal que a sua inversa h^{-1} também é contínua. A aplicação h diz-se um **homeomorfismo**.*

Definição 2.2. *X e Y dizem-se **isotópicos** se existir uma aplicação $F : X \times [0, 1] \rightarrow \mathbb{R}^n$ tal que $F(\cdot, 0)$ é a aplicação identidade em X , $F(X, 1) = Y$ e para cada t , $F(\cdot, t)$ é um homeomorfismo entre X e $F(X, t)$. À aplicação F chamamos **isotopia**.*

Definição 2.3. *Duas aplicações $f, g : X \rightarrow Y$ dizem-se **homotópicas** se existir uma aplicação contínua $H : [0, 1] \times X \rightarrow Y$ tal que para todo o $x \in X$, $H(0, x) = f(x)$ e $H(1, x) = g(x)$.*

Definição 2.4. *Dois espaços topológicos X e Y são **homotopicamente equivalentes** se existirem duas aplicações contínuas $f : X \rightarrow Y$ e $g : Y \rightarrow X$, tais que $g \circ f$ é homotópica à aplicação identidade em X e $f \circ g$ é homotópica à aplicação identidade em Y .*

Para a análise de dados estudada neste trabalho é bastante importante o estudo de invólucros convexos e polítopos para conseguir construir conjuntos a partir de pontos em \mathbb{R}^n . De facto, uma das maneiras mais simples de construir um conjunto próximo do conjunto de pontos será construir uma união de invólucros convexos a partir dos pontos da amostra. Além disso, os conjuntos convexos permitem uma manipulação geométrica e algébrica simples. No resto deste capítulo iremos aprofundar

o estudo deste mesmo assunto de forma a nos ajudar para o trabalho que pretende ser desenvolvido.

Definição 2.5. *Seja $P = \{p_1, \dots, p_k\}$ um conjunto de pontos. Define-se por **invólucro convexo** de P o menor conjunto convexo que contenha P e é dado por*

$$\text{conv}(P) = \left\{ \sum_{i=1}^k \alpha_i p_i, \alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1 \right\}.$$

*Define-se ainda por **invólucro afim** de P o conjunto*

$$\text{aff}(P) = \left\{ \sum_{i=1}^k \alpha_i p_i, \alpha_i \in \mathbb{R}, \sum_{i=1}^k \alpha_i = 1 \right\}.$$

*Além disso, um **politopo** é definido como um invólucro convexo de um número finito de pontos.*

*A **dimensão** do politopo é a dimensão do subespaço gerado pelo politopo.*

Definição 2.6. *O conjunto $\{p_1, \dots, p_k\}$ diz-se **independente de modo afim** se*

$$\forall \alpha_1, \dots, \alpha_k, \text{ não todos nulos, } \sum_{i=1}^k \alpha_i = 0 \Rightarrow \sum_{i=1}^k \alpha_i p_i \neq 0.$$

Definição 2.7. *O invólucro convexo de um conjunto de pontos independentes de modo afim diz-se um **simplex**.*

*O invólucro convexo de $k + 1$ pontos independentes de modo afim tem dimensão k e também se pode chamar **k -simplex**.*

Vendo que um politopo é a intersecção de um número finito de subespaços, estudaremos neste momento este facto e apresentamos algumas noções relacionadas com o mesmo.

Seja $a \in \mathbb{R}^n, b \in \mathbb{R}$ e h o hiperplano

$$h = \{x \in \mathbb{R}^n : h(x) = a \cdot x + b = 0\}$$

onde \cdot designa o produto escalar e consideremos ainda os dois **semi-espaços**

$$h^+ = \{x \in \mathbb{R}^n : h(x) = a \cdot x + b \geq 0\},$$

$$h^- = \{x \in \mathbb{R}^n : h(x) = a \cdot x + b \leq 0\}.$$

Definição 2.8. *Seja P um conjunto de pontos e $\mathcal{P} = \text{conv}(P)$. Se $\mathcal{P} \cap h \neq \emptyset$ e \mathcal{P} estiver contido em h^+ ou h^- dizemos que o hiperplano h **suporta** \mathcal{P} e $\mathcal{P} \cap h$ diz-se uma **face** de \mathcal{P} .*

Lema 2.9. *Uma face de um politopo é um politopo.*

Prova: Seja $f = \mathcal{P} \cap h$ uma face do politopo $\mathcal{P} = \text{conv}(P)$, $P = \{p_1, \dots, p_k\}$. Mostremos que $f = \text{conv}(P \cap h)$, e portanto, é um politopo:

Como tanto \mathcal{P} e h são convexos, f é convexo, o que implica que $\text{conv}(P \cap h) \subseteq f$.

Para a outra inclusão, seja $x \in f = \mathcal{P} \cap h$. Isto implica que $x = \sum_{i=1}^k \alpha_i p_i$, com $\alpha_i \geq 0$ para todo o i e $\sum_{i=1}^k \alpha_i = 1$.

Assim,

$$0 = h(x) = h\left(\sum_{i=1}^k \alpha_i p_i\right) = \sum_{i=1}^k \alpha_i h(p_i) \quad (2.1)$$

Assumindo, sem perda de generalidade, que $\mathcal{P} \subseteq h^+$, temos que $\alpha_i h(p_i) \geq 0$ para todo o i . O que implica, tendo em conta (2.1), que $\alpha_i h(p_i) = 0$ para todo o i . Mas assim, caso $p_i \notin h$, $\alpha_i = 0$, o que nos diz que x é combinação linear convexa dos pontos de $P \cap h$. ■

Com este lema vemos que um politopo tem um número finito de faces. À união das faces chamamos **fronteira** de \mathcal{P} . As faces de dimensão 0 são chamadas **vértices**, as de dimensão 1 **arestas**. Se o politopo tiver dimensão k , as faces de dimensão $k-1$ são chamadas **facet**s. Chamamos ainda **m-face** a uma face de dimensão m .

Lema 2.10. *Um politopo é o invólucro convexo dos seus vértices.*

De acordo com este lema, vê-se facilmente o seguinte:

- Um simplex é completamente caracterizado pelo conjunto dos seus vértices;
- Um conjunto de pontos representa um simplex.

Tal como foi dito atrás, a forma mais simples de tentar reconstruir um conjunto é a partir de uma união de invólucros convexos, pelo que faz assim sentido a seguinte definição.

Definição 2.11. *Um **complexo simplicial** é um conjunto finito de simplexes K tal que:*

1. *Qualquer face de um simplex de K é ainda um simplex de K .*
2. *A interseção de dois simplexes de K ou é vazia ou é uma face de ambos.*

A dimensão de K é a maior dimensão dos simplexes de K .

Tal como foi dito atrás, uma das maneiras mais simples de tentar reconstruir a topologia do conjunto inicial seria a partir de uma união de invólucros convexos a partir dos pontos. As três definições seguintes, dando especial destaque às duas últimas, definem alguns desses conjuntos.

Definição 2.12. *Seja $\mathcal{U} = (U_i)_{i \in I}$ uma cobertura aberta de um espaço topológico X . A esta cobertura associamos um complexo simplicial $C(\mathcal{U})$, denominado **nervo**, com conjunto de vértices \mathcal{U} e tal que*

$$\sigma = \{U_{i_1}, \dots, U_{i_k}\} \in C(\mathcal{U}) \text{ sse } \bigcap_{j=1}^k U_{i_j} \neq \emptyset$$

Definição 2.13. *Seja $P = \{p_1, \dots, p_k\}$ um conjunto de pontos em \mathbb{R}^n . Podemos definir o complexo de Čech de raio α e conjunto de pontos P dado pelo nervo de $\mathcal{U} = (\overline{B}(p_i, \alpha))_{i \in \{1, \dots, k\}}$, onde $\overline{B}(p_i, \alpha)$ diz respeito à bola fechada de centro em p_i e raio α , isto é,*

$$\{p_{i_1}, \dots, p_{i_m}\} \in \check{C}ech(P, \alpha) \Leftrightarrow \bigcap_{j=1}^m \overline{B}(p_{i_j}, \alpha) \neq \emptyset$$

Outro complexo usado, bastante mais fácil de identificar, é o complexo de Rips.

Definição 2.14. *Definimos o complexo de Rips de raio α , denotado por $Rips(P, \alpha)$, da seguinte maneira:*

$$\{p_{i_1}, \dots, p_{i_m}\} \in Rips(P, \alpha) \Leftrightarrow \|p_{i_j} - p_{i_\ell}\| \leq \alpha, \forall j, \ell \in \{1, \dots, m\}$$

Lema 2.15. $Rips(P, \alpha) \subseteq \check{C}ech(P, \alpha) \subseteq Rips(P, 2\alpha)$

Prova: Seja $\{p_{i_1}, \dots, p_{i_m}\} \in Rips(P, \alpha)$. Então, como $\|p_{i_1} - p_{i_j}\| \leq \alpha, \forall j$, temos que $p_{i_1} \in \bigcap_{j=1}^m \overline{B}(p_{i_j}, \alpha)$, o que prova que $\{p_{i_1}, \dots, p_{i_m}\} \in \check{C}ech(P, \alpha)$.

Seja agora $\{p_{i_1}, \dots, p_{i_m}\} \in \check{C}ech(P, \alpha)$. Logo, existe um ponto y tal que $\|y - p_{i_j}\| \leq \alpha, \forall j$. Vem então que

$$\|p_{i_j} - p_{i_\ell}\| \leq \|p_{i_j} - y\| + \|y - p_{i_\ell}\| \leq 2\alpha,$$

o que nos permite concluir o pretendido. ■

Com estes dois últimos conjuntos definidos podemos, escolhendo um valor de α , obter uma aproximação para o conjunto inicial. No entanto, por agora, nada nos permite concluir quanto à topologia obtida quando comparada com o conjunto

inicial. Por isso mesmo, iremos procurar outras formas de aproximar um conjunto a partir de uma sua amostra. Para finalizar este capítulo é só necessária mais uma definição que irá ser necessária mais à frente. Isto porque nalguns métodos descritos mais à frente, tendo um complexo simplicial, o que é feito é introduzir um simplex de cada vez e ir calculando o que seja necessário em cada passo. No final temos o cálculo feito para o conjunto inteiro. Surge assim a definição seguinte:

Definição 2.16. Uma *filtração* de um complexo simplicial K é uma sequência de sub-complexos tais que:

1. $\emptyset = K_0 \subset K_1 \subset \dots \subset K_m = K$.
2. $K_{i+1} = K_i \cup \sigma_{i+1}$, onde σ_{i+1} é um simplex de K .

Lema 2.17. $\emptyset = K_0 \subset K_1 \subset \dots \subset K_m = K$, com $K_{i+1} = K_i \cup \sigma_{i+1}$ é uma filtração se e só se para qualquer i , K_i contém todas as faces de σ_{i+1} .

Para finalizar iremos apenas representar os complexos de Čech e Rips para o seguinte conjunto:

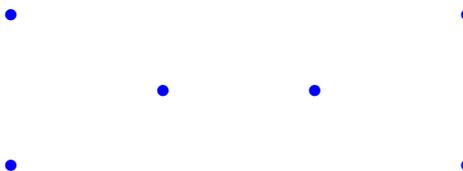


Figura 2.1 Conjunto de pontos P

Olhando para a definição, facilmente se vê que se escolhermos valores de α suficientemente pequenos, os complexos vão ser o conjunto dos pontos do conjunto:

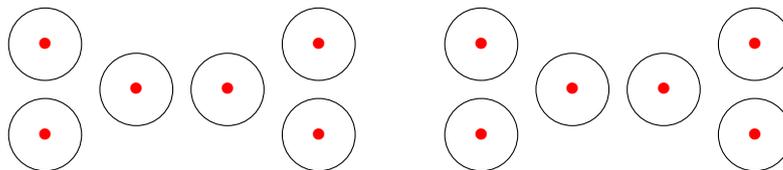


Figura 2.2 Representado a vermelho o conjunto Čech(P, α) (à esquerda) e Rips($P, 2\alpha$) (à direita) para $\alpha = 0.8$

Aumentando o valor de α iremos obtendo cada vez conjuntos maiores.

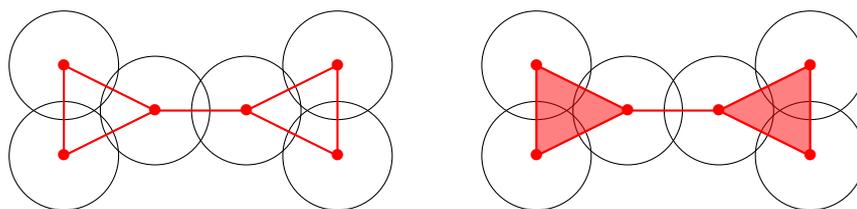


Figura 2.3 Representado a vermelho o conjunto $\check{C}ech(P, \alpha)$ (à esquerda) e $Rips(P, 2\alpha)$ (à direita) para $\alpha = 1.2$

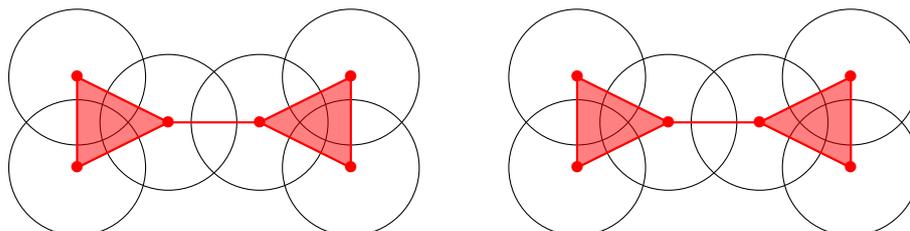


Figura 2.4 Representado a vermelho o conjunto $\check{C}ech(P, \alpha)$ (à esquerda) e $Rips(P, 2\alpha)$ (à direita) para $\alpha = 1.5$

Capítulo 3

Funções de distância

O objectivo a partir de agora será descrever relações entre características topológicas de conjuntos em função da forma de medir a distância entre eles, pelo que se começa exactamente por definir noções de distância a um conjunto e distância entre conjuntos.

Definição 3.1. Dado um conjunto $K \subset \mathbb{R}^n$, define-se a função de distância d_K a K por

$$d_K(x) = \inf_{y \in K} d(x, y), \quad \forall x \in \mathbb{R}^n,$$

com $d(x, y) = \|x - y\|$, a distância euclidiana entre pontos.

Para qualquer r não negativo, define-se o r -offset por

$$K^r = d_K^{-1}([0, r]) = \{x \in \mathbb{R}^n : d_K(x) \leq r\}$$

Além destas noções, para permitir a comparação entre conjuntos introduzimos uma função que quantifique o quão próximo dois conjuntos estão:

Definição 3.2. Sejam $K_1, K_2 \subset \mathbb{R}^n$ dois conjuntos compactos. A **distância de Hausdorff** $d_H(K_1, K_2)$ entre K_1 e K_2 é definida por qualquer uma das seguintes expressões equivalentes:

- $d_H(K_1, K_2) = \max(\sup_{y \in K_2} (\inf_{x \in K_1} \|x - y\|), \sup_{x \in K_1} (\inf_{y \in K_2} \|x - y\|))$
- $d_H(K_1, K_2)$ é o menor valor de r tal que $K_1 \subseteq K_2^r$ e $K_2 \subseteq K_1^r$
- $d_H(K_1, K_2) = \max(\sup_{y \in K_2} d_{K_1}(y), \sup_{x \in K_1} d_{K_2}(x))$
- $d_H(K_1, K_2) = \|d_{K_1} - d_{K_2}\| := \sup_{x \in \mathbb{R}^n} |d_{K_1}(x) - d_{K_2}(x)|$.

Facilmente se vê que as 3 primeiras expressões são, de facto, equivalentes. A prova da última pode encontrar-se em [4].

Esta definição de distância é puramente geométrica e bastante sensível à existência de pontos isolados que, neste trabalho, iremos tratar por outliers. Isto é, não

só não obtemos qualquer informação sobre aspetos topológicos dos conjuntos como estamos sujeitos a grande sensibilidade no comportamento da distância de Hausdorff, tal como iremos verificar mais adiante no início do próximo capítulo. Estes aspetos são críticos quando se dispõe apenas da observação de alguns pontos do conjunto e se pretende reconstruir geometricamente e topologicamente esse conjunto. Em particular se a observação dos pontos disponíveis estiver sujeita a erros de natureza aleatória ou a possíveis contaminações.

3.1. Topologia dos offsets

A ideia fundamental da reconstrução de um conjunto K_1 a partir de um outro conjunto (possivelmente um conjunto de pontos) K_2 será construir um conjunto dilatado, isto é, considerar o conjunto de pontos que estão a uma distância de K_2 controlada por um certo parâmetro. No entanto, a topologia destes conjuntos dilatados altera-se ao atingir certos pontos, ditos **pontos críticos**, especificamente pontos onde se atingem mínimos ou máximos locais para a função d_K . Precisamente por essa razão, o estudo destes pontos críticos irá ser bastante importante. No entanto, como a função d_K em geral não é diferenciável (facilmente se vê, no caso de K ser uma circunferência, que d_K não é diferenciável no centro), não podemos definir um ponto crítico da maneira usual. Antes, é primeiro necessário definir uma função que vai coincidir com o gradiente usual nos pontos onde d_K é diferenciável:

Definição 3.3. Para qualquer ponto $x \in \mathbb{R}^n$ e $K \subset \mathbb{R}^n$ um conjunto compacto seja

$$\Gamma_K(x) = \{y \in K : d(x, y) = d_K(x)\},$$

$\sigma_K(x)$ a mais pequena bola a conter $\Gamma_K(x)$, de centro $\theta_K(x)$ e raio $F_K(x)$.

Definimos assim o **gradiente generalizado** de d_K por

$$\nabla_K(x) = \frac{x - \theta_K(x)}{d_K(x)}$$

em que a norma é dada por

$$\|\nabla_K(x)\|^2 = 1 - \frac{F_K(x)^2}{d_K(x)^2} \quad (3.1)$$

Com esta função já podemos definir a noção de ponto crítico da maneira usual, isto é:

Definição 3.4. Um ponto x diz-se um **ponto crítico** de d_K se $\nabla_K(x) = 0$. Um real $c \geq 0$ diz-se **valor crítico** de d_K se existir um ponto crítico x tal que $d_K(x) = c$.

Definição 3.5. Chamamos *eixo médio* de K , denotado por $ax(K)$ ao conjunto dos pontos $x \in \mathbb{R}^n$ tais que $\Gamma_K(x)$ contém mais do que um elemento.

Pela expressão (3.1) vemos que $ax(K) = \{x \in \mathbb{R}^n : \|\nabla_K(x)\| < 1\}$.

A caracterização da topologia dos offsets, variando apenas quando se atravessam pontos críticos, é descrita no resultado seguinte e consequente corolário.

Teorema 3.6 (ver Grove, [3]). *Seja $K \subset \mathbb{R}^n$ um conjunto compacto e sejam $r_1 < r_2$ números reais tais que $[r_1, r_2]$ não contém nenhum valor crítico de d_K . Então, para $r, r' \in [r_1, r_2]$, $d_K^{-1}(r)$ é homeomorfo a $d_K^{-1}(r')$. Além disso, o conjunto $d_K^{-1}([r_1, r_2])$ é homeomorfo a $d_K^{-1}(r_1) \times [r_1, r_2]$.*

A prova deste resultado utiliza métodos de geometria diferencial pelo que será omitida neste trabalho. No entanto, a versão mais importante para proceder à reconstrução de um conjunto é o seguinte corolário:

Corolário 3.7. *Seja K um conjunto compacto e $[r_1, r_2]$ um intervalo que não contenha valores críticos de K . Se $r, r' \in]r_1, r_2[$, então K^r é isotópico a $K^{r'}$.*

Prova: Pelo teorema anterior, sabemos que para qualquer $\bar{r} \in [r_1, r_2]$ existe um homeomorfismo $h_{\bar{r}}$ de $d_K^{-1}([r_1, \bar{r}])$ para $d_K^{-1}(r_1) \times [r_1, \bar{r}]$, sendo fácil de verificar que é possível arranjar homeomorfismos tais que $h_r(x) = (y, d_K(x))$ para um certo $y \in d_K^{-1}(r_1)$ e $h_{r'}|_{[r_1, r]} = h_r$, para $r < r'$.

Tomando $a = \frac{r'-r_1}{r-r_1}$, $b = r_1$ e $f((y, s), t) = (y, s + (a(s - r_1) + b - s)t)$, facilmente se verifica que a aplicação $F : K^r \rightarrow \mathbb{R}^n$ abaixo definida é uma isotopia entre K^r e $K^{r'}$:

$$F(x, t) = \begin{cases} x & \text{se } d_K(x) \leq r_1 \\ h_{r'}^{-1}(f(h_r(x), t)) & \text{se } r_1 < d_K(x) \leq r \end{cases}$$

■

Isto significa, tomando valores de r pequenos, que a topologia dos offsets não muda enquanto r for inferior ao menor valor crítico de d_K . Faz assim sentido introduzir assim a noção de **distância topologicamente crítica**:

Definição 3.8. *Seja K um conjunto compacto de \mathbb{R}^n . A sua **distância topologicamente crítica**, denotada $wfs(K)$, é o ínfimo dos valores críticos de d_K . Se d_K não tiver valores críticos, toma-se $wfs(K) = +\infty$.*

Estamos neste momento prontos para obter um primeiro teorema que nos permita a reconstrução de um conjunto a partir de uma sua amostra.

Teorema 3.9. *Sejam K_1, K_2 dois conjuntos compactos de \mathbb{R}^n com $d_H(K_1, K_2) < \varepsilon$, $\text{wfs}(K_1) > 2\varepsilon$ e $\text{wfs}(K_2) > 2\varepsilon$. Então*

- $\mathbb{R}^n \setminus K_1$ e $\mathbb{R}^n \setminus K_2$ são homotopicamente equivalentes
- Para qualquer $0 < \alpha \leq 2\varepsilon$, K_1^α e K_2^α são homotopicamente equivalentes.

Assim, tendo K_1 um conjunto e K_2 um conjunto de pontos que satisfaçam as condições do teorema, para obtermos um conjunto com a mesma topologia de K_1 , basta considerarmos o conjunto K_2^α . No entanto, as condições do teorema geralmente não são satisfeitas pelo que teremos de aprofundar ainda mais este estudo para conseguirmos algum resultado útil.

3.1.1. Estabilidade de pontos críticos

Como a topologia dos offsets apenas se altera nos pontos críticos, é importante estudar a estabilidade destes quando perturbamos ligeiramente o conjunto K , isto é, é importante que, caso perturbemos o conjunto ligeiramente, os pontos críticos não se alterem demasiado. No entanto, esta estabilidade não existe, como se ilustra na figura que se segue.

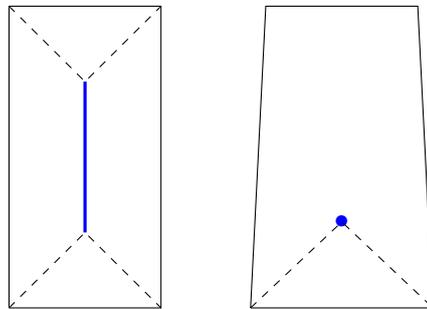


Figura 3.1 A preto o conjunto K e a azul os seus pontos críticos.

Com a imagem acima vemos que ao perturbarmos ligeiramente o conjunto do lado esquerdo, os pontos críticos alteram significativamente. Para ultrapassar este problema, é criada uma noção parametrizada de ponto crítico.

Definição 3.10. *Seja K um conjunto compacto e $0 \leq \mu \leq 1$. Um ponto $x \in \mathbb{R}^n$ diz-se μ -crítico de d_K se $\|\nabla_K(x)\| \leq \mu$.*

Com esta noção já é possível provar um teorema de estabilidade para pontos μ -críticos.

Teorema 3.11. *Sejam K_1 e K_2 dois conjuntos compactos com $d_H(K_1, K_2) \leq \varepsilon$. Para qualquer ponto μ -crítico x de K_1 , existe um ponto $(2\sqrt{\varepsilon/d_{K_1}(x)} + \mu)$ -crítico de K_2 , y , tal que $\|x - y\| \leq 2\sqrt{\varepsilon d_{K_1}(x)}$.*

Para a prova deste teorema são necessários dois lemas, apresentados de seguida.

Lema 3.12. *Seja K um conjunto compacto e x um dos seus pontos μ -críticos. Para qualquer $y \in \mathbb{R}^n$ temos*

$$d_K(y)^2 \leq d_K(x)^2 + 2\mu d_K(x)\|x - y\| + \|x - y\|^2$$

Prova: Para qualquer $x' \in \Gamma_K(x)$ temos

$$\begin{aligned} d_K(y)^2 &\leq \|y - x'\|^2 = \langle (y - x) + (x - x'), (y - x) + (x - x') \rangle \\ &= \|y - x\|^2 + \|x - x'\|^2 + 2 \langle (y - x), (x - x') \rangle \\ &= d_K(x)^2 + 2d_K(x)\|x - y\|\cos(y - x, x - x') + \|x - y\|^2 \end{aligned}$$

Para terminar basta provar que existe um ponto $x' \in \Gamma_K(x)$ tal que $\cos(y - x, x - x') \leq \mu$, isto é, o ângulo entre $y - x$ e $x - x'$ é maior ou igual a $\alpha = \cos^{-1}(\mu)$.

Caso 1: $\mu \neq 0$

Suponhamos, por redução ao absurdo, que o ângulo entre $y - x$ e $x - x'$ é inferior a α , para qualquer $x' \in \Gamma_K(x)$. Como $\Gamma_K(x)$ é compacto, existe um $\alpha' < \alpha$ tal que $\Gamma_K(x)$ está contido no cone $\{x' : \angle(y - x, x - x') \leq \alpha'\}$. Mas a intersecção deste cone com a esfera de centro x e raio $d_K(x)$ está contida numa bola de centro c' e raio $r' = d_K(x) \sin \alpha'$. Como $\Gamma_K(x)$ está contido nesta intersecção, estaria contido numa bola de raio inferior a $d_K(x) \sin \alpha = F_K(x)$, que é uma contradição!

Caso 2: $\mu = 0 \Leftrightarrow \alpha = \frac{\pi}{2}$

Se o ângulo entre $y - x$ e $x' - x$ fosse superior a $\pi - \alpha = \frac{\pi}{2}$, $\Gamma_K(x)$ teria de estar contido numa das metades da circunferência de centro x e raio $d_K(x)$, que neste caso, é igual à menor circunferência que contém $\Gamma_K(x)$. Mas isso iria implicar (devido à compacidade de $\Gamma_K(x)$), que $\Gamma_K(x)$ estaria contido numa bola de raio inferior a $d_K(x)$, que é uma contradição! ■

Lema 3.13. *Sejam K_1 e K_2 dois conjuntos compactos, com $d_H(K_1, K_2) \leq \varepsilon$. Para qualquer ponto μ -crítico de K_1 e $\rho > 0$, existe um ponto μ' -crítico de K_2 , y , tal que $\|x - y\| \leq \rho$, com*

$$\mu' \leq \frac{\rho}{2d_{K_1}(x)} + \frac{2\varepsilon}{\rho} + \mu.$$

Prova: Para qualquer $x \in \mathbb{R}^n$ mostra-se que é possível arranjar uma curva C parametrizada por comprimento de arco tal que

$$d_{K_2}(C(t)) = d_{K_2}(x) + \int_0^t \|\nabla_{K_2}(C(s))\| ds. \quad (3.2)$$

Sendo $y = C(\rho)$, temos que

$$d_{K_2}(y) = d_{K_2}(x) + \int_0^\rho \|\nabla_{K_2}(C(s))\| ds, \quad (3.3)$$

o que, pelo teorema do valor médio, implica que existe um ponto p na curva C tal que $\|p - x\| \leq \rho$ (devido à curva ser parametrizada por comprimento de arco) e

$$\|\nabla_{K_2}(p)\| \leq \frac{d_{K_2}(y) - d_{K_2}(x)}{\rho}. \quad (3.4)$$

Aplicando o Lema 3.12 e o facto de que para qualquer z , $|d_{K_1}(z) - d_{K_2}(z)| \leq \varepsilon$ (pois $d_H(K_1, K_2) \leq \varepsilon$) vem que

$$\begin{aligned} d_{K_2}(y) - d_{K_2}(x) &\leq d_{K_1}(y) - d_{K_1}(x) + 2\varepsilon \\ &\leq \sqrt{d_{K_1}(x)^2 + 2\mu d_{K_1}(x)\|x - y\| + \|x - y\|^2} - d_{K_1}(x) + 2\varepsilon \\ &\leq d_{K_1}(x) \left[\sqrt{1 + \frac{2\mu\|x - y\|}{d_{K_1}(x)} + \frac{\|x - y\|^2}{d_{K_1}(x)^2}} - 1 \right] + 2\varepsilon \\ &\leq \mu\|x - y\| + \frac{\|x - y\|^2}{2d_{K_1}(x)} + 2\varepsilon \\ &\leq \mu\rho + \frac{\rho^2}{2d_{K_1}(x)} + 2\varepsilon, \end{aligned}$$

usando que $\sqrt{1 + u} \leq 1 + \frac{u}{2}$ para $u \geq 0$.

Inserindo esta desigualdade em (3.4) obtemos o resultado pretendido. ■

Com estes dois lemas já é possível provar o Teorema 3.11, consistindo na otimização do limite do lema anterior, obtendo $\rho = 2\sqrt{\varepsilon d_{K_1}(x)}$.

Nota: Pela equação (3.2), vemos que o ponto p pode ser escolhido de modo a que $d_{K_2}(p) \geq d_{K_2}(x)$.

Definição 3.14. *Sendo K um conjunto compacto, a função crítica de K é definida por*

$$\begin{aligned} \chi_K : \mathbb{R}^+ &\rightarrow \mathbb{R}_0^+ \\ r &\rightsquigarrow \inf_{x \in d_K^{-1}(r)} \|\nabla_K(x)\|. \end{aligned}$$

É possível provar a semi-continuidade à esquerda da função anterior pelo que, juntando à compacidade de K , mostra-se que o ínfimo é de facto atingido num ponto $x \in d_K^{-1}(r)$.

Teorema 3.15. *Sejam K_1 e K_2 dois conjuntos compactos tal que $d_H(K_1, K_2) \leq \varepsilon$. Para qualquer $r > 0$, tem-se:*

$$\inf_{u \in I(r, \varepsilon)} \chi_{K_2}(u) \leq \chi_{K_1}(r) + 2\sqrt{\frac{\varepsilon}{r}},$$

onde $I(r, \varepsilon) = [r - \varepsilon, r + 2\chi_{K_1}(r)\sqrt{\varepsilon r} + 3\varepsilon]$.

Prova: Seja $x \in d_{K_1}^{-1}(r)$ tal que $\|\nabla_{K_1}(x)\| = \chi_{K_1}(r)$. Pelo Teorema 3.11 sabemos que existe um ponto p que é $(2\sqrt{\frac{\varepsilon}{r}} + \chi_{K_1}(r))$ -crítico com $\|x - p\| \leq 2\sqrt{\varepsilon r}$. Aplicando o Lema 3.12 obtemos:

$$\begin{aligned} d_{K_1}(p) &\leq \sqrt{d_{K_1}(x)^2 + 2\chi_{K_1}(r)d_{K_1}(x)\|x - p\| + \|x - p\|^2} \\ &\leq \sqrt{r^2 + 4\chi_{K_1}(r)r\sqrt{\varepsilon r} + 4\varepsilon r} \\ &\leq r\sqrt{1 + 4\chi_{K_1}(r)\sqrt{\varepsilon/r} + 4\varepsilon/r} \\ &\leq r + 2\chi_{K_1}(r)\sqrt{\varepsilon r} + 2\varepsilon \end{aligned}$$

Além disso, pela nota acima, sabemos ainda que $d_{K_2}(p) \geq d_{K_2}(x) \geq r - \varepsilon$ e $d_{K_2}(p) \leq d_{K_1}(p) + \varepsilon \leq r + 2\chi_{K_1}(r)\sqrt{\varepsilon r} + 3\varepsilon$. ■

Definição 3.16. *Para $0 < \mu \leq 1$, define-se o μ -reach de K e denota-se por $r_\mu(K)$ ao valor*

$$r_\mu(K) = \{r > 0 : \chi_K(r) < \mu\}.$$

Teorema 3.17. *Sejam K_1 e K_2 dois conjuntos compactos, com $d_H(K_1, K_2) \leq \varepsilon$ e ainda $\mu > 0$. Então*

- d_{K_1} não tem nenhum valor crítico no intervalo $]4\varepsilon/\mu^2, r_\mu(K_2) - 3\varepsilon[$;
- Para qualquer $\mu' < \mu$, $\chi_{K_1}(r) > \mu'$ para

$$r \in \left] \frac{4\varepsilon}{(\mu - \mu')^2}, r_\mu(K_2) - 3\sqrt{\varepsilon r_\mu(K_2)} \right[.$$

Prova: Suponhamos que r é um valor crítico de d_{K_1} , isto é, existe $x \in d_{K_1}^{-1}(r)$ tal que $\|\nabla_{K_1}(x)\| = 0$. Assim, pelo Teorema 3.11 sabemos que existe um ponto y com $\|x - y\| \leq 2\sqrt{\varepsilon d_{K_1}(x)}$ e $\|\nabla_{K_2}(y)\| \leq 2\sqrt{\varepsilon/d_{K_1}(x)}$.

- Caso $\|\nabla_{K_2}(y)\| \geq \mu$, temos que

$$\mu \leq 2\sqrt{\frac{\varepsilon}{d_{K_1}(x)}} \Leftrightarrow r = d_{K_1}(x) \leq \frac{4\varepsilon}{\mu^2}.$$

- Caso $\|\nabla_{K_2}(y)\| < \mu$, vem que

$$\begin{aligned}
 r_\mu(K_2) &\leq d_{K_2}(y) \leq d_{K_1}(y) + \varepsilon \leq \sqrt{d_{K_1}(x)^2 + 4\varepsilon d_{K_1}(x)} + \varepsilon \\
 &\leq d_{K_1}(x) \sqrt{1 + \frac{4\varepsilon}{d_{K_1}(x)}} + \varepsilon \leq d_{K_1}(x) \left(1 + \frac{2\varepsilon}{d_{K_1}(x)}\right) + \varepsilon \\
 &= d_{K_1}(x) + 3\varepsilon,
 \end{aligned}$$

donde se retira imediatamente que $r = d_{K_1}(x) \geq r_\mu(K_2) - 3\varepsilon$.

Para a segunda parte da prova, suponhamos que $\chi_{K_1}(r) \leq \mu'$ e seja $x \in d_{K_1}^{-1}(r)$ o ponto tal que $\|\nabla_{K_1}(x)\| = \chi_{K_1}(r) \leq \mu'$. Como x é um ponto μ' -crítico de d_{K_1} , sabemos que existe um ponto y tal que $\|\nabla_{K_2}(y)\| \leq 2\sqrt{\varepsilon/r} + \mu'$ e $\|x - y\| \leq 2\sqrt{\varepsilon r}$.

- Se $\|\nabla_{K_2}(y)\| \geq \mu$, temos que

$$2\sqrt{\frac{\varepsilon}{r}} \geq \mu - \mu' \Leftrightarrow r \leq \frac{4\varepsilon}{(\mu - \mu')^2}$$

- Se $\|\nabla_{K_2}(y)\| < \mu$, temos que

$$\begin{aligned}
 r_\mu(K_2) &\leq d_{K_2}(y) \leq d_{K_1}(y) + \varepsilon \\
 &\leq \sqrt{r^2 + 4\varepsilon r + 4\mu' r \sqrt{\varepsilon r}} + \varepsilon \\
 &\leq r \sqrt{1 + \frac{4\varepsilon}{r} + \frac{4\mu' \sqrt{\varepsilon r}}{r}} + \varepsilon \\
 &\leq r \left(1 + \frac{2\varepsilon}{r} + \frac{2\mu' \sqrt{\varepsilon r}}{r}\right) + \varepsilon \\
 &\leq r + 2\varepsilon + 2\mu' \sqrt{\varepsilon r} + \varepsilon \\
 &\leq r + 3\varepsilon + 2\sqrt{\varepsilon r}
 \end{aligned}$$

Segue-se então que

$$\begin{aligned}
 r_\mu(K_2) - 2\varepsilon &\leq r + \varepsilon + 2\sqrt{\varepsilon r} \\
 \Leftrightarrow r_\mu(K_2) - 2\varepsilon &\leq (\sqrt{r} + \sqrt{\varepsilon})^2 \\
 \Leftrightarrow r &\geq (\sqrt{r_\mu(K_2) - 2\varepsilon} - \sqrt{\varepsilon})^2 \\
 \Leftrightarrow r &\geq r_\mu(K_2) - \varepsilon - 2\sqrt{\varepsilon r_\mu(K_2) - 2\varepsilon^2} \\
 \Rightarrow r &\geq r_\mu(K_2) - \varepsilon - 2\sqrt{\varepsilon r_\mu(K_2)} \\
 \Rightarrow r &\geq r_\mu(K_2) - 3\sqrt{\varepsilon r_\mu(K_2)}
 \end{aligned}$$

■

3.2. Reconstrução

Dispomos já de resultados com caracterizações teóricas suficientes para que, nesta secção, nos interessemos por resultados que permitam a reconstrução efetiva de um conjunto a partir de algumas observações. Naturalmente, a reconstrução pretendida irá incluir as propriedades topológicas do conjunto. O primeiro teorema referido é o mais importante do capítulo, no entanto a prova é omitida neste trabalho, devido à sua dificuldade e tamanho, mas pode ser vista em [2].

Definição 3.18. *Dados α e μ positivos, diz-se que K_2 é uma (α, μ) -aproximação de K_1 se $d_H(K_1, K_2) < \alpha r_\mu(K_1)$.*

Teorema 3.19 (ver Chazal, Cohen-Steiner, Lieutier, [2]). *Seja $\mu > 0$ e K_1 um conjunto compacto tal que $r_\mu(K_1) > 0$. Seja ainda K_2 uma (α, μ) -aproximação de K_1 com*

$$\alpha < \min\left(\frac{\sqrt{5}}{2} - 1, \frac{\mu^2}{16 + 2\mu^2}\right).$$

Sejam ainda r, r' reais tais que

$$0 < r < \text{wfs}(K_1) \quad \text{e} \quad \frac{4\alpha r_\mu(K_1)}{\mu^2} \leq r' < r_\mu(K_1) - 3\alpha r_\mu(K_1)$$

Então temos:

1. K_1^r é isotópico a $K_2^{r'}$
2. $d_{K_1}^{-1}(r)$ é isotópico a $d_{K_2}^{-1}(r')$

Provamos neste capítulo uma versão mais simples do teorema anterior:

Teorema 3.20. *Seja K_2 uma (α, μ) -aproximação de um conjunto compacto K_1 . Se*

$$\alpha < \frac{\mu^2}{5\mu^2 + 12},$$

$$0 < r < \text{wfs}(K_1) \quad \text{e} \quad \frac{4d_H(K_1, K_2)}{\mu^2} < r' < r_\mu(K_1) - 3d_H(K_1, K_2).$$

Então:

1. K_1^r é homotopicamente equivalente a $K_2^{r'}$
2. $\mathbb{R}^n \setminus K_1$ é homotopicamente equivalente a $\mathbb{R}^n \setminus K_2^{r'}$.

Prova: Seja $\varepsilon = d_H(K_1, K_2) + \delta$ com δ suficientemente pequeno. Pelo Teorema 3.17 sabemos que d_{K_2} não tem valores críticos no intervalo $] \frac{4d_H(K_1, K_2)}{\mu^2}, r_\mu(K_1) - 3d_H(K_1, K_2) [$, pelo que, tendo em conta o Corolário 3.7 sabemos que todos os conjuntos $K_2^{r'}$ para r' nesse intervalo são isotópicos (e portanto, homotopicamente equivalentes). Basta então provar o teorema para $r' = \frac{4\varepsilon}{\mu^2}$.

Como $\chi_{K_2^{r'}}(\alpha) = \chi_{K_2}(\alpha + r')$, temos que $wfs(K_2^{r'}) \geq r_\mu(K_1) - 3d_H(K_1, K_2) - \frac{4\varepsilon}{\mu^2}$. Temos também que

$$d_H(K_1, K_2^{r'}) \leq d_H(K_2, K_1) + d_H(K_2, K_2^{r'}) < \varepsilon + \frac{4\varepsilon}{\mu^2}$$

Pelo Teorema 3.9 sabemos que, se $2d_H(K_1, K_2^{r'}) < \min(wfs(K_2^{r'}), wfs(K_1))$, tem-se que $(K_2^{r'})^\beta$ e K_1^β são homotopicamente equivalentes para $0 < \beta < 2d_H(K_1, K_2^{r'}) < wfs(K_1)$.

Como $(K_2^{r'})^\beta = K_2^{r'+\beta}$, para β suficientemente pequeno $K_2^{r'+\beta}$ é homotopicamente equivalente a $K_2^{r'}$. Assim, basta provar que

$$2d_H(K_1, K_2^{r'}) < \min(wfs(K_2^{r'}), wfs(K_1)).$$

Verifiquemos as duas desigualdades.

- $2d_H(K_1, K_2^{r'}) < wfs(K_2^{r'})$

$$\text{Como } wfs(K_2^{r'}) > r_\mu(K_1) - 3\varepsilon - \frac{4\varepsilon}{\mu^2} > \frac{\varepsilon}{\alpha} - 3\varepsilon - \frac{4\varepsilon}{\mu^2} \quad ,$$

escolhendo α tal que $\frac{\varepsilon}{\alpha} - 3\varepsilon - \frac{4\varepsilon}{\mu^2} > 2\varepsilon + \frac{8\varepsilon}{\mu^2}$ obtemos o pretendido. Mas isto acontece exactamente quando

$$\alpha < \frac{\mu^2}{5\mu^2 + 12}.$$

- $2d_H(K_1, K_2^{r'}) < wfs(K_1)$

Já sabemos que $2d_H(K_1, K_2^{r'}) < 2\varepsilon + \frac{8\varepsilon}{\mu^2}$. Como K_2 é uma (α, μ) -aproximação de K_1 , temos que $\varepsilon < \alpha r_\mu(K_1) \leq \alpha wfs(K_1)$. Assim, se $2\varepsilon + \frac{8\varepsilon}{\mu^2} \leq \frac{\varepsilon}{\alpha}$ temos o pretendido. Sendo esta última condição equivalente a

$$\alpha < \frac{\mu^2}{2\mu^2 + 8},$$

basta ter em conta que

$$\frac{\mu^2}{5\mu^2 + 12} < \frac{\mu^2}{2\mu^2 + 8}.$$

O Teorema 3.9 diz-nos também neste caso que $\mathbb{R}^n \setminus K_1$ é homotopicamente equivalente a $\mathbb{R}^n \setminus K_2^{r'}$. ■

Com estes dois teoremas encontramos forma de, a partir de um conjunto de dados, aproximar a topologia do conjunto de onde provêm os dados.

Capítulo 4

Distância a medidas de probabilidade

No capítulo anterior, encontramos forma de quantificar a proximidade entre dois conjuntos de forma a permitir a reconstrução das características topológicas de um conjunto.

No entanto, como normalmente a amostra de pontos que temos não é perfeita, isto é, contém ruído e/ou outliers, a distância de Hausdorff não é a melhor maneira de quantificar a proximidade entre conjuntos. De facto, basta acrescentarmos um outlier a um conjunto K para que a distância de Hausdorff entre os dois conjuntos possa ser demasiado elevada. Consequentemente, os resultados do capítulo anterior, que dependem desta distância deixam de ser úteis neste caso. Para ultrapassar este problema iremos tratar os conjuntos através de medidas de probabilidade. Intuitivamente, percebe-se que, com a medida adequada, ao acrescentar um outlier ao conjunto não irá alterar muito a medida, visto que este ponto irá ter uma medida bastante perto de zero.

Basicamente o que iremos fazer neste capítulo será alterar a função que retrata a distância de um ponto de \mathbb{R}^n ao conjunto, que neste caso está retratado através de uma medida μ . Com isto, tendo noções equivalentes às do capítulo anterior, iremos ter teoremas também eles equivalentes aos que já foram apresentados, mas neste caso para uma função distância diferente. No entanto, esta função de distância necessita de satisfazer algumas condições que a função anterior verificava, mais precisamente necessita de ser uma **função de tipo distância**:

Definição 4.1. *Uma função $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^+$ diz-se uma **função de tipo distância** se:*

1. *Para qualquer $x, y \in \mathbb{R}^n$, $|\phi(x) - \phi(y)| \leq \|x - y\|$;*
2. *A função $x \rightsquigarrow \|x\|^2 - \phi(x)^2$ for convexa;*
3. *Para qualquer conjunto compacto $A \subset \mathbb{R}$, $\phi^{-1}(A)$ é compacto.*

As duas primeiras condições são impostas para se poder definir o gradiente generalizado com $\|\nabla_x\| \leq 1$. A última condição permite garantir a compacidade das aproximações para compactos ao adaptar os argumentos discutidos no capítulo anterior. De facto, o compacto K será aproximado por $\phi^r = \phi^{-1}([0, r])$.

Definição 4.2. *Seja \mathcal{B}_n a tribo boreliana de \mathbb{R}^n , uma **medida** em \mathbb{R}^n é uma função $\mu : \mathcal{B}_n \rightarrow \mathbb{R}^+$ tal que, se $(B_i)_{i \in I}$ for uma família de subconjuntos $B_i \in \mathcal{B}_n$ disjuntos, então $\mu(\cup_{i \in I} B_i) = \sum_{i \in I} \mu(B_i)$. Caso $\mu(\mathbb{R}^n) = 1$, μ diz-se uma **medida de probabilidade**.*

*Ao menor conjunto K tal que $\mu(\mathbb{R}^n \setminus K) = 0$ chama-se **suporte** de μ .*

Dado um conjunto de pontos ou um conjunto compacto existem maneiras naturais de definir uma medida de probabilidade com suporte nesse conjunto.

- Seja P um conjunto de N pontos. A este conjunto podemos associar a medida μ_P , definida por $\mu_P(B) = \frac{1}{N}|B \cap P|$.
- Dado um conjunto m -dimensional $K \subset \mathbb{R}^n$ compacto, seja $vol_K(K)$ o seu volume m -dimensional (por exemplo, se K for uma curva em \mathbb{R}^3 , $vol_K(K)$ é o comprimento total da curva). Podemos definir a medida μ_K definida por $\mu_K(B) = vol_K(B \cap K)/vol_K(K)$.

4.1. Distância de Wasserstein

Tal como foi dito no início do capítulo, a distância de Hausdorff não é a melhor forma de quantificar a distância entre conjuntos com outliers. Nesta secção iremos apresentar uma noção que iremos ver ser melhor para retratar a distância entre tais conjuntos, neste caso tratados como medidas de probabilidade.

Definição 4.3. *Um **plano de transporte** entre duas medidas de probabilidade μ e ν em \mathbb{R}^n é uma medida de probabilidade π em $\mathbb{R}^n \times \mathbb{R}^n$ tal que $\pi(A \times \mathbb{R}^n) = \mu(A)$ e $\pi(\mathbb{R}^n \times B) = \nu(B)$.*

*O **custo** de tal transporte é dado por*

$$\mathcal{C}(\pi) = \left(\int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 d\pi(x, y) \right)^{\frac{1}{2}}.$$

Intuitivamente, pode pensar-se em $\pi(A \times B)$ como a quantidade de medida que é transportada de A para B .

Exemplo 4.4. *Sejam μ e ν as medidas com suporte finito*

$$\mu = \sum_{j=1}^{N_1} b_j \delta_{x_j} \quad \text{e} \quad \nu = \sum_{i=1}^{N_2} c_i \delta_{y_i}$$

com $\sum b_j = 1$ e $\sum c_i = 1$ Um plano de transporte é, neste caso, uma matriz não-negativa $N_1 \times N_2$, $\Pi = (\pi_{i,j})$ tal que

$$\sum_{i=1}^{N_2} \pi_{i,j} = b_j \quad \text{e} \quad \sum_{j=1}^{N_1} \pi_{i,j} = c_i.$$

O custo deste plano é

$$\mathcal{C}(\Pi) = \left(\sum_{i=1}^{N_2} \sum_{j=1}^{N_1} \pi_{i,j} \|x_j - y_i\|^2 \right)^{\frac{1}{2}}. \quad (4.1)$$

Definição 4.5. *A distância de Wasserstein entre duas medidas de probabilidade μ e ν é dada pelo custo mínimo $\mathcal{C}(\pi)$ de um plano de transporte π entre μ e ν . Esta distância denota-se por $W_2(\mu, \nu)$.*

Apesar desta distância ser praticamente impossível de calcular, é uma melhor maneira de quantificar a distância entre conjuntos com outliers. De facto, consideremos um conjunto $P = \{x_1, \dots, x_N\}$ de N pontos e um outro conjunto P' obtido substituindo os primeiros k pontos de P por pontos y_i com $d_P(y_i) = R$. Denotemos por $\mu = \frac{1}{N} \sum_{p \in P} \delta_p$ e $\nu = \frac{1}{N} \sum_{q \in P'} \delta_q$ e consideremos o plano de transporte definido pela matriz diagonal com elementos diagonais iguais a $\frac{1}{N}$. Pela expressão (4.1) temos que

$$\begin{aligned} W_2(\mu, \nu) &\leq \left(\sum_{i=1}^k \frac{1}{N} \|x_i - y_i\|^2 \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{i=1}^k \frac{1}{N} (R + \text{diam}(C))^2 \right)^{\frac{1}{2}} \\ &= \left(\frac{k}{N} (R + \text{diam}(C))^2 \right)^{\frac{1}{2}} \\ &= \left(\frac{k}{N} \right)^{\frac{1}{2}} (R + \text{diam}(C)) \end{aligned}$$

Temos então

$$W_2(\mu, \nu) \leq \left(\frac{k}{N} \right)^{\frac{1}{2}} (R + \text{diam}(C)),$$

o que significa que, se $k \ll N$, a distância de Wasserstein é praticamente zero.

4.2. Distância a uma medida de probabilidade

O próximo passo será definir uma função que quantifique a distância a uma medida, tal como foi feito no capítulo anterior com conjuntos.

Definição 4.6. *Seja μ uma medida de probabilidade e $0 \leq m < 1$. Denotamos por $\delta_{\mu,m}$ a função*

$$\begin{aligned} \delta_{\mu,m} : \mathbb{R}^n &\rightarrow \mathbb{R}_0^+ \\ x &\rightsquigarrow \inf\{r > 0 : \mu(\bar{B}(x,r)) > m\} \end{aligned}$$

Notemos que, para $m = 0$, esta função coincide com a função de distância do capítulo anterior. No entanto, esta função não é robusta relativamente a perturbações da medida μ , isto é, uma ligeira perturbação na medida μ pode alterar significativamente a função $\delta_{\mu,m}$. Considere-se, por exemplo, a medida $\mu_\varepsilon = (\frac{1}{2} - \varepsilon)\delta_0 + (\frac{1}{2} + \varepsilon)\delta_1$ e seja $m = 1/2$. Para $\varepsilon \geq 0$ temos $\delta_{\mu_\varepsilon,1/2}(t) = |1 - t|$ se $t < 0$. Caso $\varepsilon < 0$, já temos $\delta_{\mu_\varepsilon,1/2}(t) = |t|$.

Para ultrapassar este problema considera-se como função de distância a seguinte:

Definição 4.7. *Seja μ uma medida de probabilidade em \mathbb{R}^n e $0 < m_0 \leq 1$. A função de distância a μ é dada pela função*

$$\begin{aligned} d_{\mu,m_0} : \mathbb{R}^n &\rightarrow \mathbb{R}_0^+ \\ x &\rightsquigarrow \sqrt{\frac{1}{m_0} \int_0^{m_0} \delta_{\mu,m}(x)^2 dm} \end{aligned}$$

Caso a medida μ seja uma medida discreta, a função acima pode ser descrita através dos k vizinhos mais próximos:

Definição 4.8. *Seja P um conjunto de pontos em \mathbb{R}^n e $q \in \mathbb{R}^n$. Define-se por **vizinho mais próximo**, e denota-se por $nn(q, P)$ (ou simplesmente $nn(q)$), o ponto que minimiza a distância entre q e os pontos de P , isto é,*

$$d(q, P) := d(q, nn(q, P)) = \min_{p \in P} d(q, p).$$

Exemplo 4.9. *Seja P um conjunto de N pontos e seja $\mu_P = \frac{1}{N} \sum_{p \in P} \delta_p$. Para $0 < m \leq 1$, $\delta_{\mu,m}(x) = \|x - p\|$, onde p é o k vizinho mais próximo de x , com k o menor inteiro maior que mN . Assim, para $m \in \left] \frac{k}{N}, \frac{k+1}{N} \right]$, $\delta_{\mu,m}(x)$ é constante.*

Temos então, assumindo $m_0 = \frac{k_0}{N}$:

$$\begin{aligned} d_{\mu, m_0}(x) &= \sqrt{\frac{1}{m_0} \int_0^{m_0} \delta_{\mu, m}(x)^2 dm} \\ &= \sqrt{\frac{1}{m_0} \sum_{k=1}^{k_0} \frac{1}{N} \delta_{\mu, k/N}(x)^2} \\ &= \sqrt{\frac{1}{k_0} \sum_{p \in NN_P^{k_0}(x)} \|p - x\|^2}, \end{aligned}$$

onde $NN_P^{k_0}(x)$ é o conjunto dos k_0 vizinhos mais próximos de x em P .

Definição 4.10. Uma medida ν é uma **submedida** de μ se para qualquer $B \in \mathcal{B}_n$, $\nu(B) \leq \mu(B)$.

O conjunto das submedidas de μ é denotado por $\text{Sub}(\mu)$ e o conjunto das submedidas ν com $\nu(\mathbb{R}^n) = m_0$ é denotado por $\text{Sub}_{m_0}(\mu)$.

Através das submedidas de μ obtemos uma outra formulação para a função d_{μ, m_0} que nos permitirá retirar alguns resultados interessantes.

Proposição 4.11 (ver Chazal, Cohen-Steiner, Mérigot, [5]). Para qualquer medida μ , temos:

$$d_{\mu, m_0} = \min_{\nu \in \text{Sub}_{m_0}(\mu)} m_0^{-1/2} W_2(m_0 \delta_x, \nu).$$

Além disso, para qualquer medida ν_x que seja minimizante da expressão acima, temos

$$d_{\mu, m_0} = \left(\frac{1}{m_0} \int_{\mathbb{R}^n} \|x - h\|^2 d\nu_x(h) \right)^{\frac{1}{2}}$$

Lema 4.12. Sejam μ e μ' duas medidas em \mathbb{R}^n . Então

$$W_2(\mu, \mu') \geq d_H(\text{Sub}_{m_0}(\mu), \text{Sub}_{m_0}(\mu')),$$

onde

$$\begin{aligned} d_H(\text{Sub}_{m_0}(\mu), \text{Sub}_{m_0}(\mu')) &= \\ &= \max \left(\sup_{\nu' \in \text{Sub}_{m_0}(\mu')} \inf_{\nu \in \text{Sub}_{m_0}(\mu)} W_2(\nu, \nu'), \sup_{\nu \in \text{Sub}_{m_0}(\mu)} \inf_{\nu' \in \text{Sub}_{m_0}(\mu')} W_2(\nu, \nu') \right) \end{aligned}$$

Prova: Seja π o plano de transporte ótimo entre μ e μ' , isto é,

$$W_2(\mu, \mu') = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 d\pi(x, y).$$

Dada uma submedida ν de μ , prova-se, a partir do Teorema de Radon-Nykodim que é possível arranjar um plano de transporte, π' , de ν para uma submedida ν' de μ' que seja uma submedida de π . Então,

$$\begin{aligned} W_2(\nu, \nu') &\leq \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 d\pi'(x, y) \\ &\leq \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^2 d\pi(x, y) \\ &= W_2(\mu, \mu'). \end{aligned}$$

Provamos então que, para qualquer $\nu \in Sub_{m_0}(\mu)$, $\inf_{\nu' \in Sub_{m_0}(\mu')} W_2(\nu, \nu') \leq W_2(\mu, \mu')$, pelo que o seu supremo também é menor ou igual a $W_2(\mu, \mu')$. Trocando os papéis de ν e ν' , obtemos o pretendido. ■

Teorema 4.13. *Sejam μ e μ' duas medidas de probabilidade. Então*

$$\|d_{\mu, m_0} - d_{\mu', m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} W_2(\mu, \mu').$$

Prova:

É consequência da desigualdade triangular que

$$\min_{\nu \in Sub_{m_0}(\mu)} W_2(m_0 \delta_x, \nu) \leq d_H(Sub_{m_0}(\mu), Sub_{m_0}(\mu')) + \min_{\nu' \in Sub_{m_0}(\mu')} W_2(m_0 \delta_x, \nu').$$

Assim, para provar o resultado basta ver que

$$\begin{aligned} d_{\mu, m_0}(x) &= \frac{1}{\sqrt{m_0}} \min_{\nu \in Sub_{m_0}(\mu)} W_2(m_0 \delta_x, \nu) \\ &\leq \frac{1}{\sqrt{m_0}} \left(d_H(Sub_{m_0}(\mu), Sub_{m_0}(\mu')) + \min_{\nu' \in Sub_{m_0}(\mu')} W_2(m_0 \delta_x, \nu') \right) \\ &\leq \frac{1}{\sqrt{m_0}} W_2(\mu, \mu') + d_{\mu', m_0}(x). \end{aligned}$$

■

Com este último teorema, vemos que a função de distância considerada já é robusta em relação à medida, isto é, caso duas medidas sejam relativamente próximas ($W_2(\nu, \nu')$ pequeno) as suas funções de distância também o vão ser. Além disso, pode verificar-se (por exemplo, em [1]) que a função d_{μ, m_0} é uma função de tipo distância.

4.3. Aplicações a inferência geométrica

Nesta última secção do capítulo, tudo o que foi estudado no capítulo anterior irá ser traduzido para o caso de uma função de tipo distância qualquer. De facto, na prova dos resultados anteriores a distância euclidiana que servia de base apenas era utilizada através das propriedades que foram incluídas na definição de função de tipo

distância. Podemos portanto, efetuando adaptações simples, traduzir os resultados precedentes para este novo quadro. Em seguida apresentam-se, sem demonstração, estas traduções, concluindo-se tal como atrás, com um resultado que permite a reconstrução da topologia de um conjunto.

Definição 4.14. *Seja ϕ um função de tipo distância. Denotamos por $\phi^r = \phi^{-1}([0, r])$.*

*Um ponto x diz-se α -crítico se $\phi(x+h)^2 \leq \phi(x)^2 + 2\alpha\|h\|\phi(x) + \|h\|^2$ para qualquer h . Se $\alpha = 0$, x diz-se apenas um **ponto crítico**. Um real r diz-se **valor crítico** de ϕ se existir um ponto crítico x com $\phi(x) = r$.*

*A **distância topologicamente crítica em r** de ϕ , $\text{wfs}_\phi(r)$, é o maior valor de r' tal que ϕ não tem nenhum valor crítico no intervalo $]r, r+r'[\$.*

Para $0 < \alpha < 1$, o α -reach de ϕ , $\text{reach}_\alpha(\phi)$, é o maior valor de r tal que $\phi^{-1}([0, r])$ não contém nenhum ponto α -crítico.

Teorema 4.15. *Seja ϕ uma função de tipo distância e $r_1 < r_2$ tal que ϕ não tem valores críticos no intervalo $[r_1, r_2]$. Então todos os conjunto ϕ^r são isotópicos para $r \in [r_1, r_2]$.*

Teorema 4.16. *Sejam ϕ e φ duas funções de tipo distância tal que $\|\phi - \varphi\|_\infty = \sup_{x \in \mathbb{R}^n} |\phi(x) - \varphi(x)| \leq \varepsilon$, $\text{wfs}_\phi(r) > 2\varepsilon$ e $\text{wfs}_\varphi(r) > 2\varepsilon$. Então, para qualquer $0 < \eta \leq \varepsilon$, $\phi^{r+\eta}$ e $\varphi^{r+\eta}$ são homotopicamente equivalentes.*

Teorema 4.17. *Sejam ϕ e φ duas funções de tipo distância tal que $\|\phi - \varphi\|_\infty \leq \varepsilon$. Para qualquer ponto α -crítico x de ϕ , existe um ponto α' -crítico y de φ com $\|x - y\| \leq 2\sqrt{\varepsilon\phi(x)}$ e $\alpha' \leq \alpha + 2\sqrt{\varepsilon/\phi(x)}$.*

Teorema 4.18. *Sejam ϕ e φ duas funções de tipo distância tal que $\|\phi - \varphi\|_\infty \leq \varepsilon$ e $\text{reach}_\alpha(\phi) \geq R$. Então φ não tem nenhum valor crítico no intervalo $] \frac{4\varepsilon}{\alpha^2}, R - 3\varepsilon[$.*

Teorema 4.19. *Sejam ϕ e φ duas funções de tipo distância tal que $\|\phi - \varphi\|_\infty \leq \varepsilon$ e $\text{reach}_\alpha(\phi) \geq R$. Então para*

$$\frac{4\varepsilon}{\alpha^2} < r < R - 3\varepsilon \text{ e } 0 < \eta < R,$$

os conjuntos ϕ^η e φ^r são homotopicamente equivalentes desde que

$$\varepsilon \leq \frac{R}{5 + \frac{4}{\alpha^2}}.$$

Capítulo 5

Simulação em R

Para este estudo iremos considerar o conjunto K_1 dado pela figura seguinte.

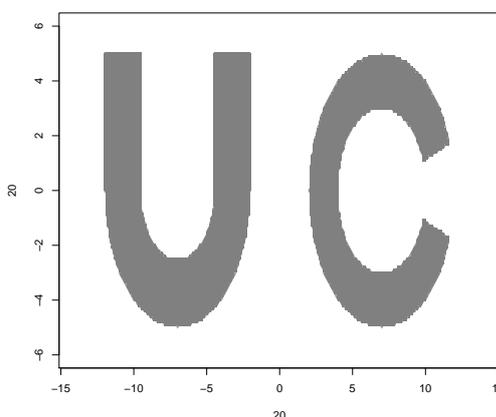


Figura 5.1 Conjunto K_1 representado a cinzento

O objectivo será, a partir de uma amostra retirada deste conjunto, possivelmente com outliers e ruído, tentar construir o conjunto de partida. Já existem diversos métodos para o fazer usando, por exemplo, o complexo de Čech. No entanto, neste trabalho, iremos usar os teoremas 3.19 e 4.19 para tentar arranjar um conjunto próximo do inicial.

Teorema. *Seja $\mu > 0$ e K_1 um conjunto compacto tal que $r_\mu(K_1) > 0$. Seja ainda K_2 uma (α, μ) -aproximação de K_1 com*

$$\alpha < \min \left(\frac{\sqrt{5}}{2} - 1, \frac{\mu^2}{16 + 2\mu^2} \right).$$

Sejam ainda r, r' reais tais que

$$0 < r < \text{wfs}(K_1) \quad \text{e} \quad \frac{4\alpha r_\mu(K_1)}{\mu^2} \leq r' < r_\mu(K_1) - 3\alpha r_\mu(K_1)$$

Então K_1^r é isotópico a $K_2^{r'}$.

Usemos este teorema com $\mu = 1$, $\alpha = \frac{1}{19} < \min \left(\frac{\sqrt{5}}{2} - 1, \frac{\mu^2}{16 + 2\mu^2} \right) = \frac{1}{18}$, supondo que a amostra que temos é uma (α, μ) -aproximação de K .

Normalmente, não se conhece o conjunto inicial, pelo que o valor de $r_1(K_1)$ teria de ser estimado. No entanto, ainda não se conhece nenhuma maneira eficaz de estimar este valor. Por isso, para este estudo, iremos calcular o valor verdadeiro sabendo que conhecemos o conjunto inicial. Facilmente se vê que o $\text{ax}(K_1)$ é o conjunto representado a preto abaixo:

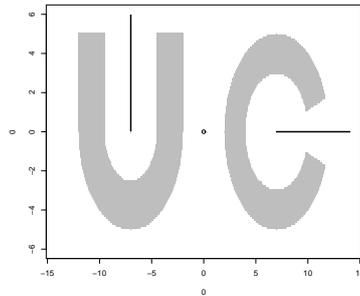


Figura 5.2 Conjunto K_1 representado a cinzento e $\text{ax}(K_1)$ a preto

Com a figura é possível verificar que $r_1(K_1) = 3 \sin(\frac{\pi}{9})$.

Pelo teorema, escolhendo $r' = \frac{4\alpha r_\mu(K_1)}{\mu^2} = \frac{12 \sin(\frac{\pi}{9})}{19}$, calculando $K_2^{r'}$ temos uma aproximação para o conjunto inicial.

Teorema. *Sejam ϕ e φ duas funções distância tal que $\|\phi - \varphi\|_\infty \leq \varepsilon$ e $\text{reach}_\alpha(\phi) \geq R$.*

Então para

$$\frac{4\varepsilon}{\alpha^2} \leq r \leq R - 3\varepsilon \text{ e } 0 < \eta < R,$$

os conjuntos ϕ^η e φ^r são homotopicamente equivalentes desde que

$$\varepsilon \leq \frac{R}{5 + \frac{4}{\alpha^2}}.$$

Consideremos ϕ e φ as funções de distância apresentadas na definição 4.7 com $m_0 = \frac{1}{100}$ e as medidas apresentadas no início do capítulo 4 com suporte no conjunto inicial e na amostra retirada, respectivamente.

Usando este teorema com α suficientemente pequeno, por exemplo, $\alpha = 0.1$ e supondo que $\|\phi - \varphi\|_\infty \leq \varepsilon$ com $\varepsilon = \frac{R}{600} < \frac{R}{405}$, basta apenas calcular o valor de R . Mais uma vez, não conseguindo estimá-lo, é necessário usar o verdadeiro valor que é possível verificar que é aproximadamente 1.493376.

Usando o valor de $r = \frac{4\varepsilon}{\alpha^2} = \frac{6R}{9}$, calculando o conjunto φ^r teremos uma aproximação para o conjunto inicial.

Nas figuras que se seguem são apresentadas duas amostras retiradas do conjunto. A da esquerda com 2000 pontos e a da direita com 6000.

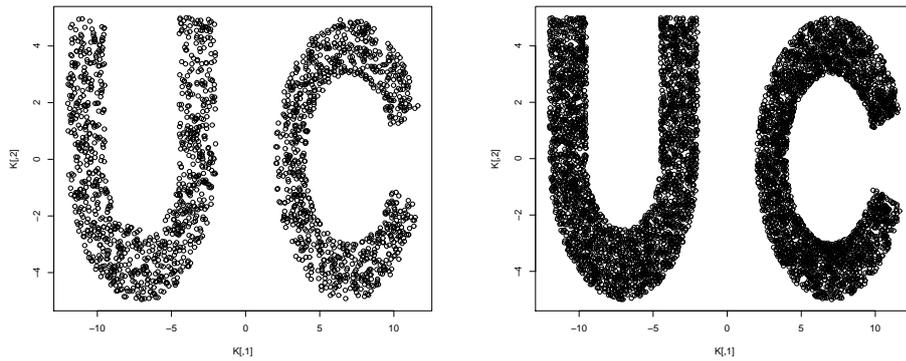


Figura 5.3 À esquerda a amostra com 2000 pontos (1000 em cada componente conexa) e à direita a amostra com 6000 pontos (3000 em cada componente)

Usando os dois métodos acima descritos para a amostra mais pequena, nota-se, observando a figura 6.4, que no primeiro caso ainda não consegue recuperar a topologia do conjunto inicial. De facto, ao observar a figura reparamos que o conjunto a cinzento do lado esquerdo ainda contém alguns buracos. Isto porque com esta quantidade de pontos a distância entre pontos ainda é demasiado elevada para termos uma (α, μ) -aproximação. Quanto ao segundo, mesmo com esta quantidade de pontos, o método já consegue recuperar a topologia do conjunto. Apesar disso, o conjunto recuperado por este método é um pouco maior que o conjunto pretendido.

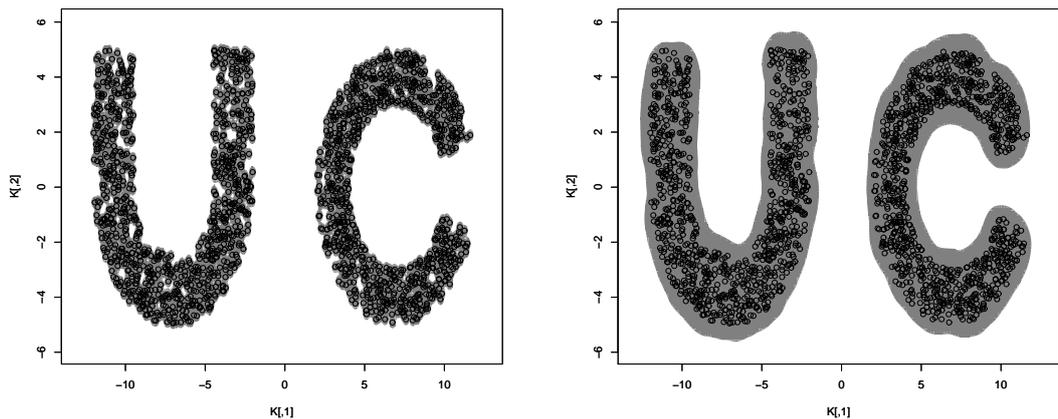


Figura 5.4 À esquerda a cinzento representado o conjunto $K_2^{r'}$ e à direita o conjunto φ^r para a amostra com 2000 pontos

Quanto aos conjuntos de 6000 pontos, vê-se que ambos conseguem recuperar a topologia do conjunto, com o primeiro método a recuperar um conjunto bastante mais perto do conjunto inicial, tal como podemos ver na figura seguinte:

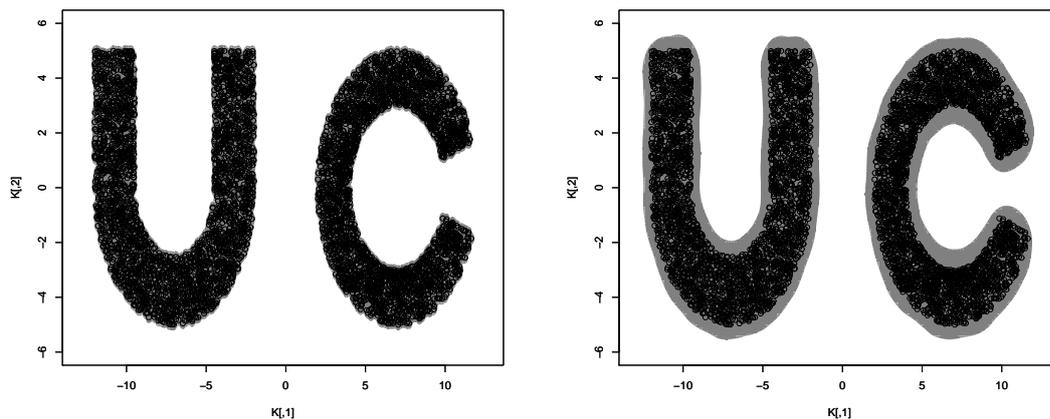


Figura 5.5 À esquerda a cinzento representado o conjunto K_2^r e à direita o conjunto φ^r para a amostra com 6000 pontos

Suponhamos agora que além das amostras retiradas de cada conjunto temos ainda 500 outliers, gerados uniformemente na janela $[-14, 14] \times [-6, 6]$. Claro que, ao gerar os outliers nesta janela, alguns deles irão cair dentro do conjunto inicial e, por isso, não são outliers. Mesmo assim, com 500 pontos gerados, ainda iremos ter uma quantidade significativa de outliers.

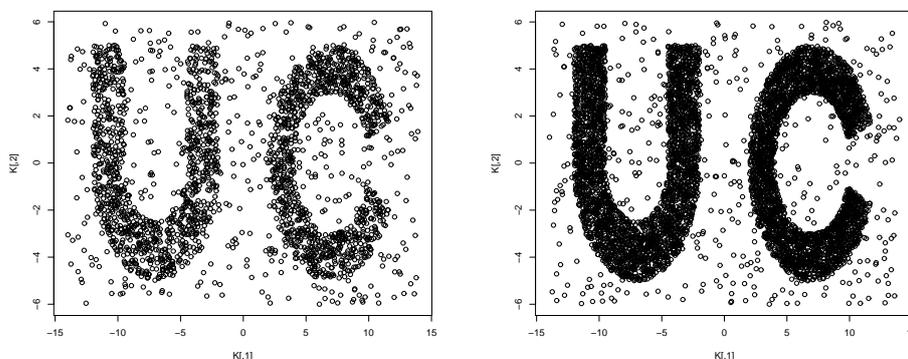


Figura 5.6 À esquerda a amostra de 2000 pontos e 500 outliers e à direita uma amostra de 6000 pontos e 500 outliers

Tal como referido no início do Capítulo 4, o primeiro método não consegue distinguir os outliers dos pontos retirados do conjunto, resultando num conjunto com uma topologia diferente do esperado (demasiadas componentes conexas). Pelo contrário, o segundo método, mesmo com uma amostra de 2000 pontos consegue fazê-lo, resultando portanto num conjunto bastante parecido com o inicial.

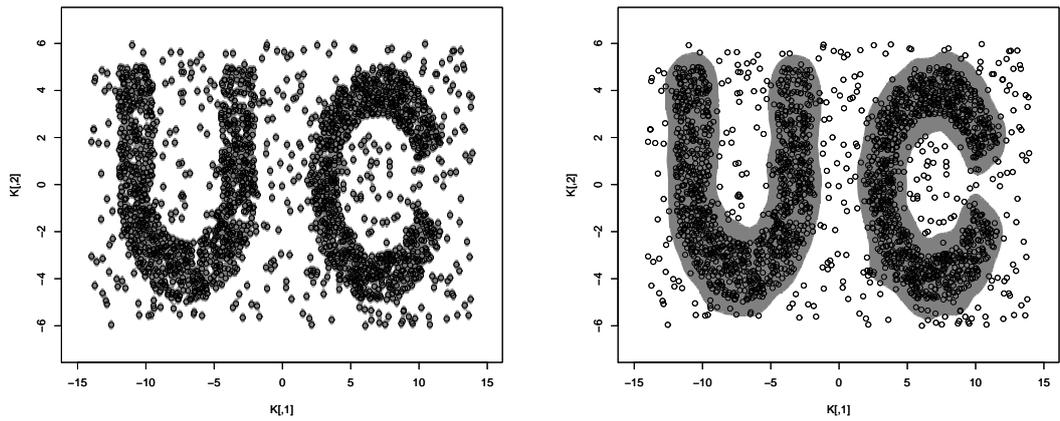


Figura 5.7 À esquerda a cinzento representado o conjunto $K_2^{r'}$ e à direita representado o conjunto φ^r para a amostra com 2000 pontos com outliers

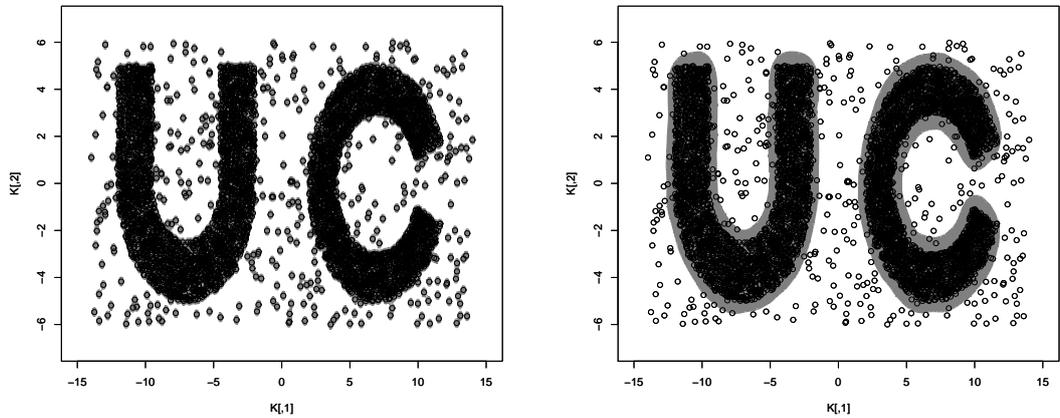


Figura 5.8 À esquerda a cinzento representado o conjunto $K_2^{r'}$ e à direita representado o conjunto φ^r para a amostra com 6000 pontos com outliers

De seguida, estudaremos o efeito de ruído nos métodos utilizados. Para isso, foram retiradas amostras do conjunto inicial somando em cada uma das componentes uma distribuição normal de média 0 e desvio padrão 0.5.

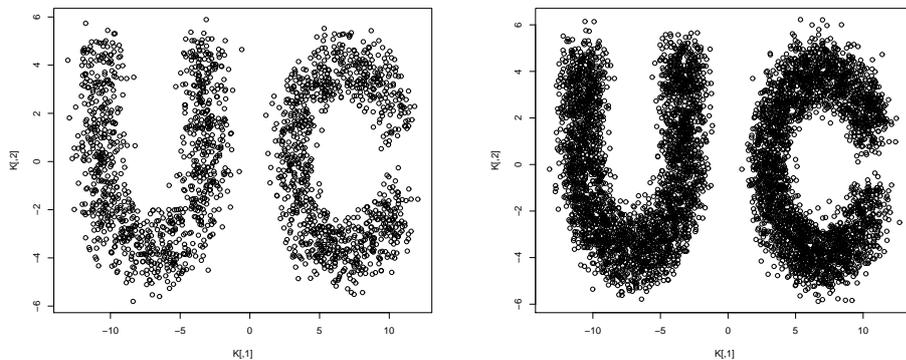


Figura 5.9 À esquerda a amostra de 2000 pontos com ruído e à direita a amostra de 6000 pontos com ruído

Olhando para as figuras seguintes, reparamos que, mais uma vez, o primeiro método não consegue recuperar a topologia do conjunto, ao contrário do segundo.

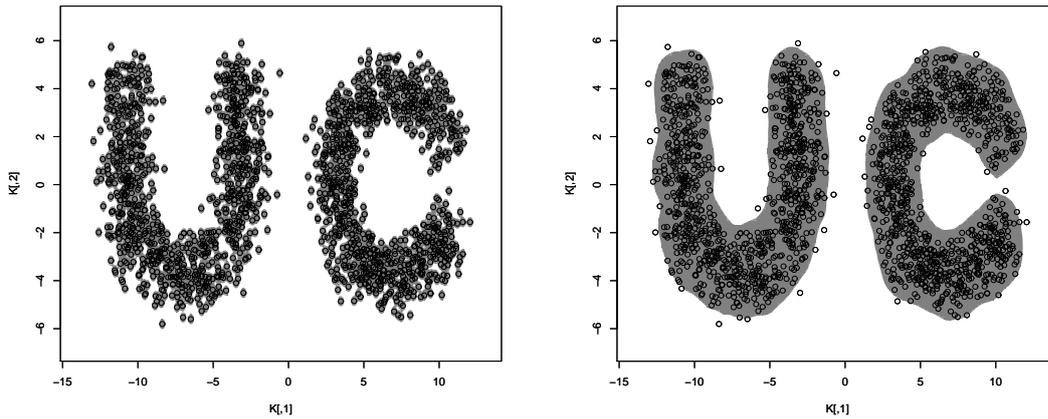


Figura 5.10 À esquerda a cinzento representado o conjunto $K_2^{r'}$ e à direita representado o conjunto φ^r para a amostra com 2000 pontos com ruído

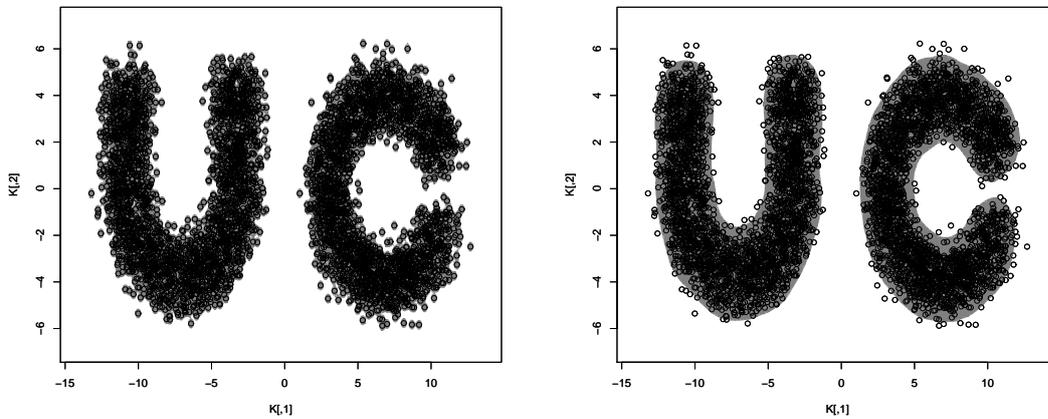


Figura 5.11 À esquerda a cinzento representado o conjunto $K_2^{r'}$ e à direita representado o conjunto φ^r para a amostra com 6000 pontos com ruído

Para finalizar, iremos verificar se os métodos funcionam se juntarmos os dois casos anteriores, isto é, se para além de ruído na amostra retirada, ainda juntarmos 500 outliers à amostra. Mais uma vez, muitos destes outliers irão cair dentro do conjunto considerado.

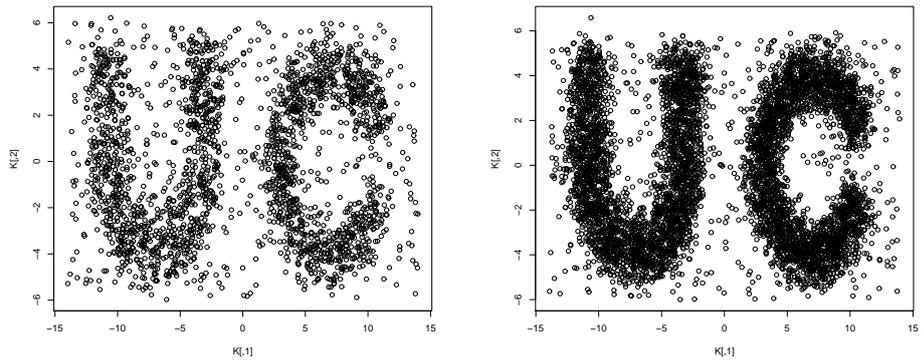


Figura 5.12 À esquerda a amostra de 2000 pontos com ruído e outliers e à direita a amostra de 6000 pontos com ruído e outliers

Mais uma vez, verificamos que apenas o segundo método consegue recuperar um conjunto parecido com o inicial.

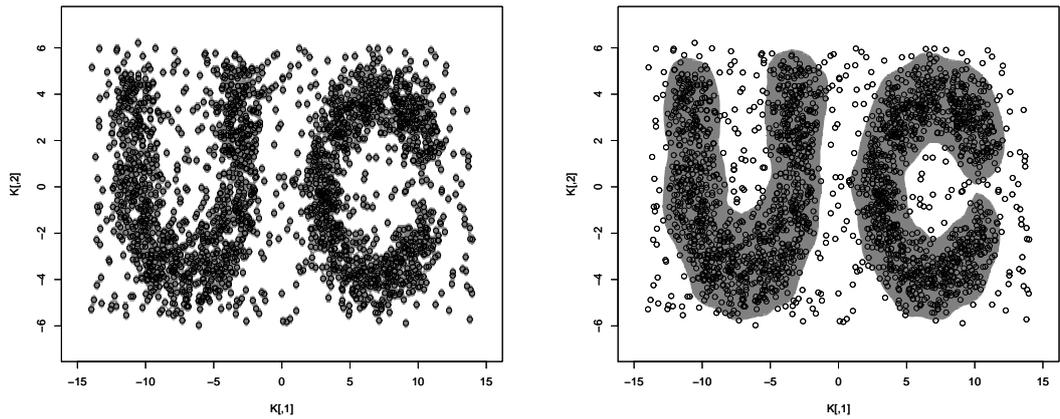


Figura 5.13 À esquerda a cinzento representado o conjunto K_2^r e à direita representado o conjunto φ^r para a amostra com 2000 pontos com ruído e outliers

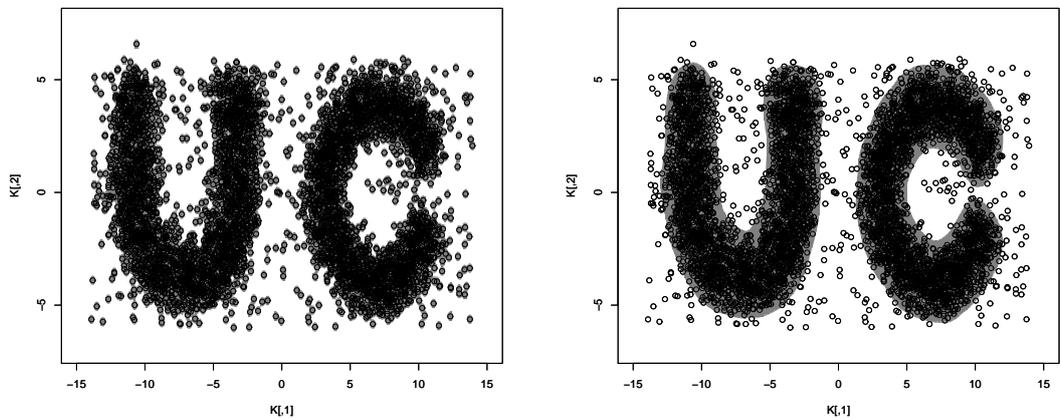


Figura 5.14 À esquerda a cinzento representado o conjunto K_2^r e à direita representado o conjunto φ^r para a amostra com 6000 pontos com ruído e outliers

Experimentemos agora para um outro conjunto, neste caso uma bola com dois buracos de diferentes dimensões, o que corresponde a um conjunto com uma topologia menos trivial. Vejamos o comportamento do segundo método, mesmo com uma amostra de 1000 pontos, para tentar reconstruir o conjunto e a sua topologia:

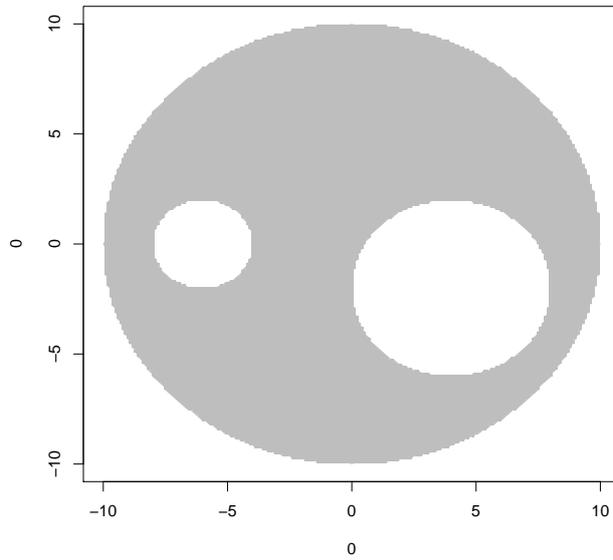


Figura 5.15 Conjunto K_1 representado a cinzento

Será que com uma amostra o segundo método consegue recuperar a topologia e resultar num conjunto com dois buracos no meio? A Figura 5.16 abaixo apresenta uma amostra de 1000 pontos do conjunto da Figura 5.15 e a sua reconstrução. A escolha de parâmetros baseados nos resultados anteriores permite uma boa reconstrução.

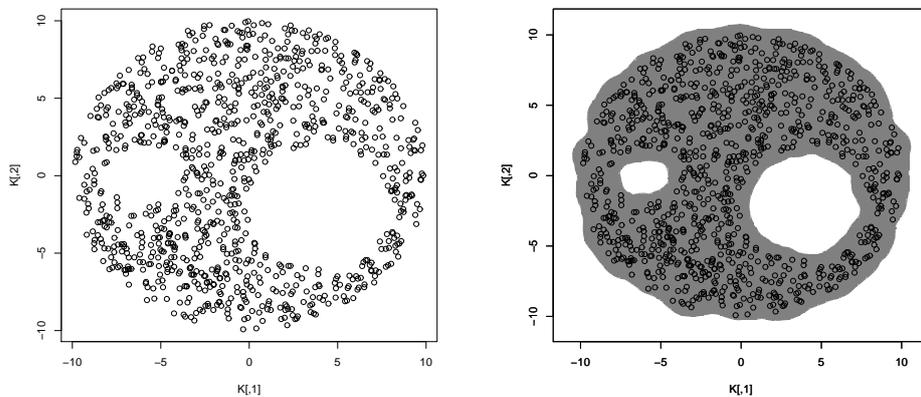


Figura 5.16 À esquerda a amostra com 1000 pontos e à direita o conjunto φ^r representado a cinzento

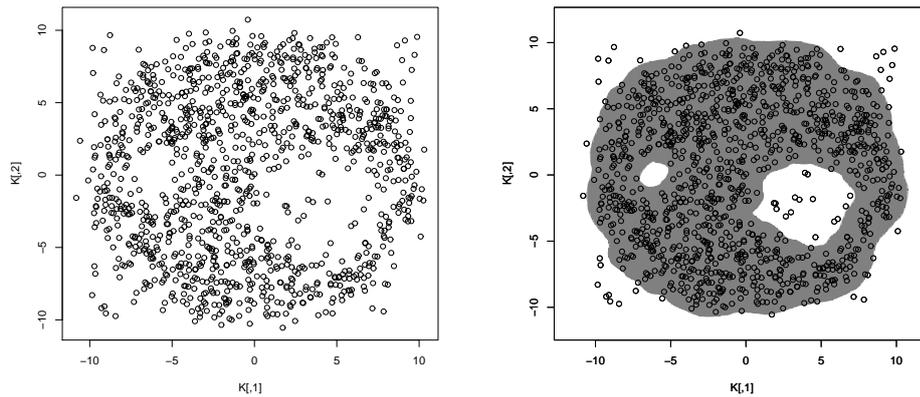


Figura 5.17 À esquerda a amostra com 1000 pontos com ruído e 200 “outliers” e à direita o conjunto φ^r representado a cinzento

A Figura 5.17 apresenta uma amostra que foi contaminada com observações perturbadas. Ainda assim, o método descrito permite identificar razoavelmente o conjunto e a existência de buracos no seu interior.

Com este estudo, verificamos realmente que, caso a amostra não contenha ruído nem outliers, o primeiro método até pode dar melhores resultados além de demorar bastante menos tempo. No entanto, ao acrescentar ruído e/ou outliers à amostra, o primeiro método torna-se ineficaz na tentativa de recuperar a topologia do conjunto. Pelo contrário o segundo método, apesar de demorar bastante mais tempo, é o único dos dois que é útil em amostras como as que temos normalmente.

Olhando novamente para a amostra anterior, será que não há nenhum método que nos indique algumas características topológicas do conjunto inicial, por exemplo, que tem apenas uma componente conexa e é um conjunto com dois buracos?

A introdução de alguns invariantes topológicos permite obter uma resposta às dificuldades de visualização geométrica inerentes a amostras de maior dimensão. Um dos conjuntos de invariantes que se tem revelado de grande utilidade, ligado às relações de homologia já referidas atrás, são os números de Betti. A sua apresentação no contexto e algoritmos para a sua determinação serão discutidos no capítulo seguinte.

Capítulo 6

Inferência Homológica

Apesar de tudo o que foi tratado anteriormente, representar geometricamente os conjuntos K^r ou ϕ^r apenas é possível até à terceira dimensão. Quando subimos de dimensão a representação de conjuntos torna-se impossível. Uma forma de contornar esta dificuldade é obtida à custa da inferência homológica, isto é, a estimação de alguns invariantes topológicos do conjunto através de uma sua amostra. Neste caso, os invariantes topológicos que nos interessarão são os números de Betti e a persistência homológica. Como se trata de um tema bastante diferente dos anteriores, iremos apresentar algumas das noções e resultados mais básicos e que fundamentam as construções a efectuar.

Definição 6.1. *Seja K um complexo simplicial e sejam $\sigma_1, \dots, \sigma_p$ os seus simplexes de dimensão k . Uma k -cadeia c é definida como a soma de simplexes de dimensão k de K , isto é,*

$$c = \sum_{i=1}^p \varepsilon_i \sigma_i \text{ com } \varepsilon_i \in \{0, 1\}.$$

Definindo a soma e produto por um escalar da seguinte maneira:

$$c + c' = \sum_{i=1}^p (\varepsilon_i + \varepsilon'_i) \sigma_i \text{ e } \lambda c = \sum_{i=1}^p (\lambda \varepsilon_i) \sigma_i,$$

onde $\varepsilon_i + \varepsilon'_i$ e $\lambda \varepsilon_i$ é calculado módulo 2, ficamos com um espaço vectorial $C_k(K)$.

Intuitivamente podemos ver a cadeia c como a união dos simplexes que a constituem, isto é, para os quais $\varepsilon_i = 1$.

Definição 6.2. *A **fronteira** de um k -simplex σ , denotada por $\partial(\sigma)$, é definida pela soma das suas $(k - 1)$ -faces.*

Com isto é possível definir o operador fronteira da maneira seguinte:

$$\begin{aligned} \partial : C_k(K) &\rightarrow C_{k-1}(K) \\ c &\rightsquigarrow \partial c = \sum_{\sigma \in c} \partial(\sigma). \end{aligned}$$

Teorema 6.3. $\partial \circ \partial = 0$.

Prova: Devido à definição do operador ∂ , apenas é necessário verificar a propriedade para um simplex. Seja $\sigma = [v_0 \cdots v_k]$ o k -simplex gerado pelos pontos v_1, \dots, v_k e $[v_0 \cdots \hat{v}_i \cdots v_k]$ o $(k-1)$ -simplex gerado pelos pontos $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k$.

$$\begin{aligned} \partial \circ \partial(\sigma) &= \partial \left(\sum_{i=0}^k [v_0 \cdots \hat{v}_i \cdots v_k] \right) \\ &= \sum_{i=0}^k \partial([v_0 \cdots \hat{v}_i \cdots v_k]) \\ &= \sum_{j < i} [v_0 \cdots \hat{v}_j \cdots \hat{v}_i \cdots v_k] + \sum_{i < j} [v_0 \cdots \hat{v}_i \cdots \hat{v}_j \cdots v_k] \\ &= 0. \end{aligned}$$

Definição 6.4. O conjunto dos k -ciclos de K , $Z_k(K)$ é dado pelo núcleo de $\partial : C_k(K) \rightarrow C_{k-1}(K)$:

$$Z_k(K) = \{c \in C_k(K) : \partial c = 0\}.$$

A imagem de $\partial : C_{k+1}(K) \rightarrow C_k(K)$, denotada por $B_k(K)$, representa o conjunto das k -fronteiras, isto é, k -cadeias que são fronteira de uma $(k+1)$ -cadeia:

$$B_k(K) = \{c \in C_k : \exists c' \in C_{k+1}, c = \partial c'\}.$$

Note-se que B_k e Z_k são subespaços de C_k e pelo teorema 6.3, temos

$$B_k \subseteq Z_k \subseteq C_k.$$

Definição 6.5. O k -ésimo grupo de homologia H_k de K é o grupo quociente

$$H_k(K) = Z_k(K)/B_k(K).$$

À sua dimensão chamamos de k -ésimo número de Betti e designamos por $\beta_k(K)$.

A **classe de homologia** de um ciclo $c \in Z_k(K)$ é o conjunto $c + B_k(K) = \{c + b : b \in B_k(K)\}$. Além disso, dizemos que dois ciclos são **homólogos** se estiverem na mesma classe de homologia.

Intuitivamente, os números de Betti β_0, β_1 e β_2 indicam o número de componentes conexas, o número de buracos bidimensionais e o número de buracos tridimensionais.

6.1. Algoritmo para calcular os números de Betti

Seja K um complexo simplicial e consideremos uma sua filtração

$$\mathfrak{F} = \{\emptyset = K^0 \subset K^1 \subset \cdots \subset K^m = K\},$$

onde, para todo o i ,

$$K^{i+1} = K^i \cup \sigma^{i+1}.$$

Algoritmo 6.6.

Dados: Uma filtração de um complexo simplicial K de dimensão d com m simplexes.

$$\beta_0 = 0; \beta_1 = 0; \dots; \beta_d = 0$$

for $i = 1$ to m **do**

$$k = \dim \sigma^i - 1$$

if σ^i estiver contido num $(k + 1)$ -ciclo de K^i **then**

$$\beta_{k+1} = \beta_{k+1} + 1$$

else

$$\beta_k = \beta_k - 1$$

Output: Os números de Betti β_0, \dots, β_d de K

Prova da correção do algoritmo: Suponhamos que já temos calculados os números de Betti de K^{i-1} .

Caso 1: σ^i está contido num $(k + 1)$ -ciclo c de K^i .

Assim, c não pode ser homólogo a nenhum $(k + 1)$ -ciclo c' em K^{i-1} . Caso contrário, iria existir uma $(k + 2)$ -cadeia em K^{i-1} , b , tal que $c + c' = c - c' = \partial b$. Mas como σ^i não pode estar contido em c' , iria estar contido em $c + c'$, o que é impossível devido à definição de filtração. Assim, c cria uma nova classe de homologia e portanto, $\beta_{k+1}(K^i) \geq \beta_{k+1}(K^{i-1}) + 1$. Para concluir o pretendido basta notar que, sendo c e c' dois $(k + 1)$ -ciclos contendo σ^i , então $c + c'$ é um $(k + 1)$ -ciclo em K^{i-1} , concluindo portanto que c' está contido no subespaço gerado por $Z_{k+1}(K^{i-1})$ e c , isto é, $\beta_{k+1}(K^i) \leq \beta_{k+1}(K^{i-1}) + 1$.

Caso 2: σ^i não está contido em nenhum $(k + 1)$ -ciclo de K^i .

Assim, o k -ciclo $\partial\sigma^i$ não é uma fronteira em K^{i-1} . Caso contrário iria existir uma cadeia c em K^{i-1} tal que $\partial\sigma^i = \partial c \Leftrightarrow \partial(\sigma^i + c) = 0$, e portanto $\sigma^i + c$ era um $(k + 1)$ -ciclo em K^i que continha σ^i . Como consequência, $\partial\sigma^i$, que não pertencia a $B_k(K^{i-1})$, já vai pertencer a $B_k(K^i)$, pelo que $\beta_k(K^i) \leq \beta_k(K^{i-1}) - 1$. Por um argumento análogo ao do caso 1, prova-se a outra desigualdade.

Definição 6.7. *Seja K um complexo simplicial e \mathfrak{F} uma sua filtração. Um simplex σ^i é dito **positivo** se estiver contido num $(k + 1)$ -ciclo em K^i . Caso contrário, é dito **negativo**.*

Com este algoritmo pode calcular-se os números de Betti de qualquer complexo simplicial. No entanto, o nosso objectivo é calcular estes números para um conjunto K e não para um complexo simplicial. Mesmo assim, usando o Teorema 3.19 e o

Teorema do Nervo apresentado no capítulo 4 de [1], é possível verificar-se que, para r suficientemente pequeno, K^r é homotopicamente equivalente a $\check{C}ech(P, \alpha)$ para um certo valor de α , o que significa que os dois têm os mesmos números de Betti. Assim, calculando os números de Betti para este último complexo simplicial, obtemos os números de Betti de K^r , uma estimaco para os números de Betti de K .

6.2. Persistência topológica

Nesta seco o objectivo ser o de calcular o grfico de persistência de um conjunto, mais um dos invariantes topolgicos estudadas. No entanto, para ser mais fcil o seu estudo iremos primeiro apresentar um caso mais simples, o grfico de persistência de uma funo.

6.2.1. Persistência de uma funo

Consideremos ento uma funo $f : \mathbb{R} \rightarrow \mathbb{R}$, por exemplo a da figura 4.1. Para j, estamos interessados em estudar a evoluo da topologia dos conjuntos do tipo $f^{-1}([-\infty, t])$. Olhando para a figura reparamos que a topologia destes conjuntos se altera nas linhas horizontais a cheio assinaladas na figura. Ao ultrapassar a primeira, $y = y_1$, é criada uma componente conexa. Quando ultrapassamos a segunda (em $y = y_2$) é criada uma segunda componente que se une à primeira quando se ultrapassa a terceira linha, em $y = y_3$. Assim, criamos o ponto (y_2, y_3) no grfico de persistência, ponto que corresponde à “criao” e “morte” da componente conexa. Ao subir ainda mais no grfico, criamos ainda o ponto (y_4, y_5) correspondentes à 4ª e 5ª linha. Por ltimo, como a primeira componente nunca desaparece, criamos o ponto $(y_1, +\infty)$. Ao grfico representamos ainda a diagonal $y = x$ por motivos que iro ser bvios mais adiante. Na figura da esquerda esto ainda representados os conjuntos $f^{-1}([-\infty, t])$ para $t = 3$ (a verde) e $t = 15$ (a vermelho).

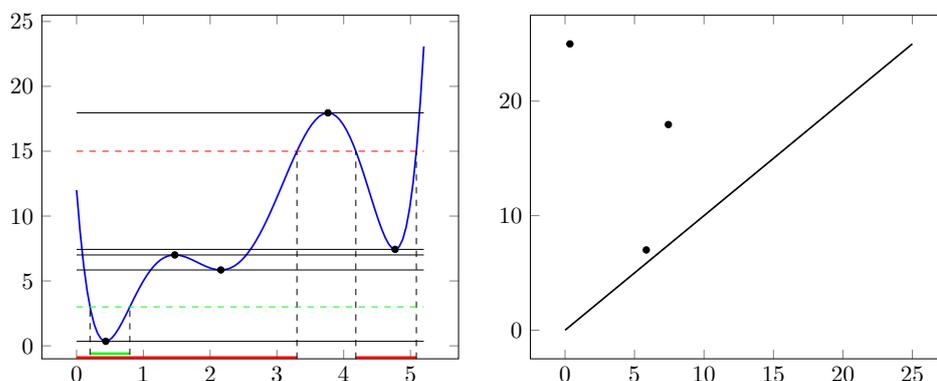


Figura 4.1 Funo f e o seu grfico de persistência

Quando subimos para dimensões superiores, ao passar valores críticos tanto pode alterar o número de componentes conexas como pode alterar outras características topológicas (ciclos, buracos, ...). Podemos assim definir um gráfico de persistência para cada dimensão.

Substituindo a função f por uma sua aproximação g , vemos que o gráfico de persistência de g tem muito mais pontos que o de f . No entanto, a maioria deles estão bastante perto da diagonal, correspondendo a ligeiras perturbações na função. Pelo contrário, os pontos mais afastados da diagonal aparecem em posições bastante próximas dos pontos do gráfico de persistência de f (a azul na figura abaixo).

Isto significa que duas funções próximas entre si têm gráficos de persistência próximos.

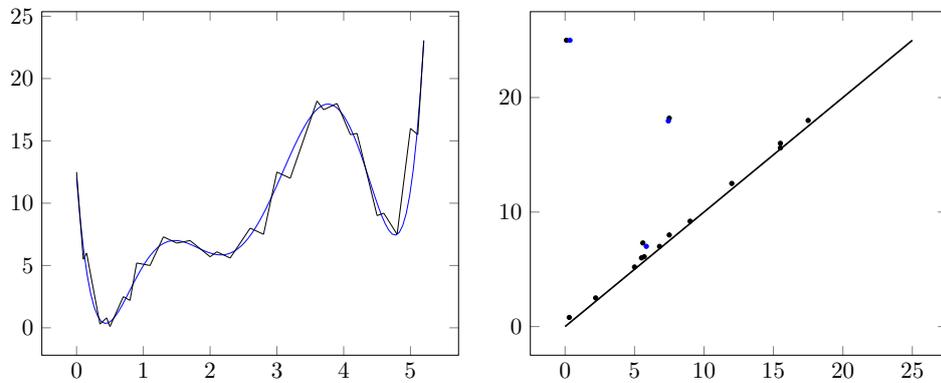


Figura 4.2 Função g e o seu gráfico de persistência

6.2.2. Persistência topológica de uma filtração

Seja K um complexo simplicial e $\mathfrak{F} = \{\emptyset = K^0 \subset K^1 \dots \subset K^m = K\}$ uma sua filtração.

Definição 6.8. Para qualquer $0 \leq i \leq m$ denotamos por C_k^i , Z_k^i e B_k^i o conjunto das k -cadeias, k -ciclos e k -fronteiras de K^i , respectivamente. Assim, o k -ésimo grupo de homologia de K^i é $H_k^i = Z_k^i/B_k^i$.

Com estas definições temos

$$Z_k^0 \subseteq Z_k^1 \subseteq \dots \subseteq Z_k^m = Z_k(K)$$

$$B_k^0 \subseteq B_k^1 \subseteq \dots \subseteq B_k^m = B_k(K)$$

Lema 6.9. Seja σ^i um k -simplex positivo de K . Então existe um k -ciclo que não é fronteira em K^i que contém σ^i e que não contém mais nenhum simplex positivo.

Prova: Suponhamos que o lema é verdade para todos os simplexes positivos adicionados à filtração antes de σ^i .

Como σ^i é positivo, existe um k -ciclo c' em K^i que contém σ^i . Sejam σ^{i_j} , $j = 1 \dots, p$, os k -simplexes positivos que estão contidos nesse ciclo, diferentes de σ^i , e c_j os k -ciclos não fronteiros que contêm σ^{i_j} e não contêm mais nenhum k -simplex positivo. Então $c = c' + c_1 + \dots + c_p$ é um k -ciclo em que σ^i é o único k -simplex positivo. Visto que não pode existir nenhum $(k + 1)$ -simplex em K^i a conter σ^i na sua fronteira, c é um ciclo que não é fronteira.

Este lema ajuda na construção de uma base de H_k^i de maneira indutiva:

- A base de H_k^0 é vazia.
- Assumamos que já conhecemos uma base de H_k^{i-1} .
 1. Caso σ^i seja um k -simplex positivo, adicionamos à base de H_k^{i-1} a classe de homologia do ciclo c_i do lema 6.9 associado a σ^i .
 2. Caso σ^i seja um $(k+1)$ -simplex negativo, sejam c_{i_1}, \dots, c_{i_p} os ciclos associados, pelo lema 6.9, a simplexes positivos $\sigma^{i_1}, \dots, \sigma^{i_p}$ tais que as suas classes de homologia formam uma base de H_k^{i-1} . A fronteira de σ^i , $d = \partial\sigma^i$ é um k -ciclo não fronteiro em K^{i-1} que passa a ser fronteira em K^i . Assim, escreve-se de forma única

$$d = \sum_{j=1}^p \varepsilon_j c_{i_j} + b,$$

com $\varepsilon_j \in \{0, 1\}$ e b uma k -fronteira de K^{i-1} . Sendo $\ell(j) = \max\{i_j : \varepsilon_j = 1\}$, obtemos uma base de H_k^i retirando a classe de homologia de $c_{\ell(j)}$ da base de H_k^{i-1} .

Definição 6.10. *Aos pares $(j, \ell(j))$ chamamos **pares de persistência**.*

O problema para calcular os pares de persistência é calcular o valor de $\ell(j)$. Para isso, é possível usar o seguinte algoritmo matricial: Seja $M = (m_{i,j})_{i,j=1,\dots,m}$ a matriz definida por:

- $m_{i,j} = 1$ se $\sigma^i \in \partial\sigma^j$,
- $m_{i,j} = 0$ caso contrário.

Para cada coluna C_j de M denotamos por $l(j)$ o índice da última linha que contém um elemento diferente de zero na coluna C_j , isto é,

$$i = l(j) \Leftrightarrow (m_{i,j} = 1 \text{ e } m_{i',j} = 0 \forall i' > i)$$

Algoritmo 6.11.

Dados: Matriz M acima definida

for $j=1$ **to** m **do**

while existir $j' < j$ tal que $l(j') == l(j)$ **do**

$$C_j = C_j + C_{j'} \pmod{2}$$

Output: Pares $(j, l(j))$

Prova da correção do algoritmo:

1. Em cada passo do algoritmo, a coluna C_j vai representar uma cadeia da forma

$$\partial \left(\sigma^j + \sum_{i < j} \varepsilon_i \sigma^i \right) \text{ com } \varepsilon_i \in \{0, 1\}$$

2. No fim do algoritmo, se j for tal que $l(j)$ está definido, então $\sigma^{l(j)}$ é um simplex positivo.

A coluna C_j , neste caso, representa uma cadeia da forma

$$\sigma^{l(j)} + \sum_{p < l(j)} \eta_p \sigma^p \text{ com } \eta_p \in \{0, 1\}$$

Mas pelo ponto 1 (e visto que $\partial \circ \partial = 0$), esta cadeia é um ciclo em $K^{l(j)}$ que contém $\sigma^{l(j)}$. Logo $\sigma^{l(j)}$ é um simplex positivo.

3. No fim do algoritmo, se j for tal que $l(j)$ não está definido, isto é, se a coluna C_j só contiver zeros, então σ^j é um simplex positivo.

Neste caso, pelo ponto 1, vemos que

$$\partial \left(\sigma^j + \sum_{i < j} \varepsilon_i \sigma^i \right) = 0,$$

pelo que σ^j está contido num ciclo de K^j .

4. No fim do algoritmo, se a coluna C_j tiver elementos não nulos, então $(j, l(j))$ é um par de persistência, isto é $\ell(j) = l(j)$.

A partir dos pontos 1,2 e 3 tiramos que

$$\partial\sigma^j = \sigma^{l(j)} + \sum_{p < l(j)} \eta_p \sigma^p + \partial \left(\sum_{i < j} \varepsilon_i \sigma^i \right)$$

Para concluir o pretendido, basta notar que $\sigma^{l(j)}$ é positivo pelo que foi adicionado à base de homologia em $i = l(j)$ e não é retirado da base antes de $i = j$.

Atrás foi referido que funções próximas iriam ter gráficos de persistência próximos. Mas como se quantifica esta proximidade entre gráficos? A resposta a esta pergunta é dada pela definição seguinte:

Definição 6.12. *Sejam D_1 e D_2 dois gráficos de persistência. A distância entre D_1 e D_2 é definida por*

$$d_B(D_1, D_2) = \inf_{\gamma} \sup_{p \in D_1} \|p - \gamma(p)\|_{\infty}$$

onde γ percorre o conjunto das bijecções entre D_1 e D_2 e $\|p - q\|_{\infty} = \max(|x_p - x_q|, |y_p - y_q|)$. Caso D_1 e D_2 não tenham o mesmo número de pontos, associamos os pontos que sobram com pontos da diagonal $x = y$.

Com esta definição de distância entre gráficos de persistência é possível provar, sob certas condições, que gráficos de funções próximas têm uma distância entre si bastante pequena, de que é exemplo a seguinte proposição.

Proposição 6.13. *Sejam K_1, K_2 compactos. Então*

$$d_B(dgm(d_{K_1}), dgm(d_{K_2})) \leq d_H(K_1, K_2),$$

onde $dgm(f)$ representa o gráfico de persistência de f .

Tendo em conta que estas duas características mencionadas no capítulo são invariantes topológicos podemos tentar aplicá-las na identificação de um conjunto através de uma sua amostra. Obtendo estas características a partir da amostra já é possível ficar com uma ideia do tipo de topologia que o conjunto poderá ter. Por exemplo, usando os números de Betti, vemos o número de componentes conexas que o conjunto poderá ter, além de muitas outras características, tal como foi dito atrás.

Apêndice A

Códigos em R

Código para gerar as amostras do conjunto da Figura 5.1:

```
1 n=3000
2 m=500
3 K= matrix(0,nrow=2*n+m,ncol=2)
4
5 for(i in 1:n)
6 {
7   r=runif(1,3,5)
8   teta=runif(1,20,340)
9   xU=runif(1,-5,5)
10  yU=runif(1,-5,5)
11  while((yU< -sqrt(5^2-xU^2)) || ((xU<2.5)&&(xU>-2.5)&&(yU>-sqrt(2.5^2-xU^2))))
12  {
13    xU=runif(1,-5,5)
14    yU=runif(1,-5,5)
15  }
16  K[i,1]=r*cos(pi*teta/180)+7#+rnorm(1,0,0.5)
17  K[i,2]=r*sin(pi*teta/180)+rnorm(1,0,0.5)
18  K[n+i,1]=xU-7#+rnorm(1,0,0.5)
19  K[n+i,2]=yU#+rnorm(1,0,0.5)
20 }
21
22 #Outliers
23 for(i in 1:m)
24 {
25   K[2*n+i,1]=runif(1,-14,14)
26   K[2*n+i,2]=runif(1,-6,6)
27 }
```

Código para gerar a Figura 5.2:

```

1 ax<-function(x,y)
  {
3   if((x>=-12)&&(x<= -9.5)&&(y>=-sqrt(25-(x+7)^2))&&(y<=5))
      {return(1)}
5   else {if((x>=-9.5)&&(x<=-4.5)&&(y>=-sqrt(25-(x+7)^2))&&(y<=-sqrt(2.5^2-(x+7)
      ^2)))
      {return(1)}
7   else {if((x>=-4.5)&&(x<=-2)&&(y>=-sqrt(25-(x+7)^2))&&(y<=5))
      {return(1)}
9   else {if(((x-7)^2+y^2<=25)&&((x-7)^2+y^2>=9)&&((x-7)/sqrt((x-7)^2+y^2)<=
      cos(20*pi/180)))
      {return(1)}
11  else {if((x==-7)&&(y>=0))
      {return(2)}
13  else {if((x>=7)&&(y==0))
      {return(2)}
15  else {if((x==0)&&(y==0))
      {return(2)}
17  else {return(0)}
      }
19  }
21  }
23  }
  }
25
x <-seq(-14,14,by=0.1)
27 y <-seq(-6,6,by=0.1)
z <-matrix(nrow=length(x),ncol=length(y))
29
for (i in 1:length(x))
31 {
33   for (j in 1:length(y))
      {
35     z[i,j] <- ax(x[i],y[j])
      z[i,j]
37   }
39 }
plot(0,0,xlim=c(-14,14),ylim=c(-6,6))
41 .filled.contour(x,y,z,levels=c(0.5,1.5,2.5),col=c("grey","black","red"))

```

Código para gerar a Figura 5.15:

```

1 conjunto<-function(x,y)
  {
3   if((x^2+y^2>100)||((x+6)^2+y^2<4)||((x-4)^2+(y+2)^2<16))
      {return(1)}
5   else {return(0)}
  }
7
9 x <-seq(-10,10,by=0.1)
10 y <-seq(-10,10,by=0.1)
11 z <-matrix(nrow=length(x),ncol=length(y))
12 for (i in 1:length(x))
  {
13   for (j in 1:length(y))
    {
15     z[i,j] <- conjunto(x[i],y[j])
16     z[i,j]
17   }
18 }
19 plot(0,0,xlim=c(-10,10),ylim=c(-10,10))
21 .filled.contour(x,y,z,levels=c(-0.5,0.5),col=c("grey","black"))

```

Código para gerar as amostras do conjunto da Figura 5.15:

```

1 m=200
2 n=1000
3 K= matrix(0,nrow=n+m,ncol=2)
4
5 for(i in 1:n)
  {
7   x=runif(1,-10,10)
8   y=runif(1,-10,10)
9   while((x^2+y^2>100)||((x+6)^2+y^2<4)||((x-4)^2+(y+2)^2<16))
    {
11    x=runif(1,-10,10)
12    y=runif(1,-10,10)
13  }
14  K[i,1]=x#+rnorm(1,0,0.5)
15  K[i,2]=y#+rnorm(1,0,0.5)
16 }
17
18 #Outliers
19 for(i in 1:m)
  {
21  K[n+i,1]=runif(1,-10,10)
22  K[n+i,2]=runif(1,-10,10)
23 }

```

Implementação do primeiro método de reconstrução:

```
1 ti=Sys.time()
3 library(SpatialTools)
5 dK <- function(x,y)
  {
7   aux1<-c(x,y)
   aux=matrix(aux1,1,2)
9   min(dist2(aux,K))
  }
11
12 R=3*sin(pi/9)
13 #R=2
15 r=4*R/19
17 x <-seq(min(K[,1]-1),max(K[,1]+1),by=0.1)
   y <-seq(min(K[,2]-1),max(K[,2]+1),by=0.1)
19 z <-matrix(nrow=length(x),ncol=length(y))
   for (i in 1:length(x))
21 {
23   for (j in 1:length(y))
   {
25     z[i,j] <- dK(x[i],y[j])
     z[i,j]
27   }
29 }
31 plot(K,xlim=c(min(K[,1]-1),max(K[,1]+1)), ylim=c(min(K[,2]-1),max(K[,2]+1)))
   .filled.contour(x,y,z,levels=c(0,r),col=grey(seq(0.5,0.5,length=length(levels)
   )))
33 par(new=TRUE)
   plot(K,xlim=c(min(K[,1]-1),max(K[,1]+1)), ylim=c(min(K[,2]-1),max(K[,2]+1)))
35
36 tf=Sys.time()
37 tf-ti
```

Implementação do segundo método de reconstrução:

```

1 ti=Sys.time()
3 library(SpatialTools)
5 dK <- function(x,y)
6 {
7   soma=0
8   nk=round(nrow(K)/100)
9   aux1<-c(x,y)
10  aux=matrix(aux1,1,2)
11  D<-dist2(aux,K)
12  for(i in 1:nk)
13  {
14    soma=soma+min(D)^2
15    nn=which.min(D)
16    D<-D[-nn]
17  }
18  sqrt(soma/nk)
19 }
21 #R=2.0007
22 R=1.493376
23
24 r=6*R/9
25
26 x <-seq(min(K[,1]-1),max(K[,1]+1),by=0.1)
27 y <-seq(min(K[,2]-1),max(K[,2]+1),by=0.1)
28 z <-matrix(nrow=length(x),ncol=length(y))
29
30 for (i in 1:length(x))
31 {
32
33   for (j in 1:length(y))
34   {
35     z[i,j] <- dK(x[i],y[j])
36     z[i,j]
37   }
38 }
39
40
41 plot(K,xlim=c(min(K[,1]-1),max(K[,1]+1)), ylim=c(min(K[,2]-1),max(K[,2]+1)))
42 .filled.contour(x,y,z,levels=c(0,r),col=grey(seq(0.5,0.5,length=length(levels)
43 )))
44 par(new=TRUE)
45 plot(K,xlim=c(min(K[,1]-1),max(K[,1]+1)), ylim=c(min(K[,2]-1),max(K[,2]+1)))
46 tf=Sys.time()
47 tf-ti

```


Bibliografia

- [1] J.-D. Boissonnat, F. Chazal, M. Yvinec. *Computational Topology Inference*.
- [2] F. Chazal, D. Cohen-Steiner, A. Lieutier. *Normal cone approximation and offset shape isotopy*. Computational Geometry Theory, 2009.
- [3] K. Grove. *Critical point theory for distance functions*. Proc. of Symposia in Pure Mathematics, volume 54, 1993. Part 3.
- [4] Daniel Kraft. *Computing the Hausdorff Distance of Two Sets from Their Signed Distance Functions*
- [5] F. Chazal, D. Cohen-Steiner, Q. Mérigot. *Geometric Inference for Probability Measures*. Foundations of Computational Mathematics, 2011.