



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de São José do Rio Preto

Fernando Gasparotto da Silva

Estrutura e estabilidade de módulos de persistência

São José do Rio Preto
2017

Fernando Gasparotto da Silva

Estrutura e estabilidade de módulos de persistência

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Orientadora: Profa. Dra. Thaís Fernanda Mendes Monis

São José do Rio Preto
2017

Gasparotto da Silva, Fernando.

Estrutura e estabilidade de módulos de persistência/ / Fernando Gasparotto da Silva . -- São José do Rio Preto, 2017
123 f. : il.

Orientador: Thaís Fernanda Mendes Monis

Dissertação (mestrado) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Letras e Ciências Exatas

1. Matemática. 2. Topologia algébrica. 3. Teoria de homologia.
4. Estabilidade. I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 513.831

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Fernando Gasparotto da Silva

Estrutura e estabilidade de módulos de persistência

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Comissão Examinadora

Profa. Dra. Thaís Fernanda Mendes Monis
Orientadora

Prof. Dr. Anderson Paião dos Santos
Departamento de Matemática - Universidade Tecnológica Federal do Paraná

Profa. Dra. Tatiana Miguel Rodrigues de Souza
Departamento de Matemática - UNESP

São José do Rio Preto
14 de Agosto de 2017

RESUMO

O intuito deste trabalho é de integrar os aspectos aplicado e teórico da Homologia Persistente, uma ferramenta popular da Topological Data Analysis (TDA). Para isso, são apresentados e demonstrados os resultados fundamentais da teoria embasada na topologia algébrica que permitem o desenvolvimento de algoritmos e paradigmas computacionais para obter diagramas de persistência. Dessa forma, iniciaremos explorando como decodificar as informações contidas em um módulo de persistência, entendendo os conceitos de multiconjuntos, módulos de persistência e cálculos Quiver. Em seguida, o caminho contrário será explorado, onde os dados são codificados em diagramas de persistência a fim de extrair suas características topológicas, aprofundando os conceitos de funções de Morse, Homologia Persistente, diagramas de persistência, dualidade e simetria, bem como estabilidade. Por último, encerramos demonstrando duas possíveis aplicações da teoria no âmbito computacional no campo da Biologia.

Palavras-chave: Módulos de Persistência, Homologia Persistente, Estabilidade, Estrutura de Módulos de Persistência, Diagramas de Persistência.

ABSTRACT

The goal of this work is to integrate applied and theoretical aspects of Persistence Homology, a popular tool in Topological Data Analysis (TDA). For this, we present and prove fundamental theoretical results based on algebraic topology, which allow us to develop algorithms and computational paradigms to obtain persistence diagrams. In this way, we start exploring how to decode the information contained in a persistence module, understanding the concepts of multiset, persistence modules and Quiver calculations. Then, the opposite path will be explored, where the data are encoded in persistence diagrams in order to extract their topological characteristics, going deep into the concepts of Morse functions, persistent homology, persistence diagrams, duality and symmetry, as well as stability. Finally, we conclude with two possible applications, one from computational theory, and the second one in the field of biology.

Keywords: Persistence Modules, Persistent Homology, Stability, Structure of Persistence Modules, Persistence Diagrams.

À minha família e todos que me apoiaram.

Agradecimentos

Agradeço, primeiramente, à Deus por todas as oportunidades que tive em minha vida e por estar sempre presente em meus passos.

Agradeço meus pais, Luis Carlos e Rosângela, por me ensinarem que a maior riqueza que se pode ter é o conhecimento e a sabedoria, e me guiarem neste caminho, além de servirem como modelo de dedicação e inspiração. Tudo o que conquistei é, em grande parte, graças a vocês. Obrigada por comemorarem minhas vitórias comigo, me ajudarem quando preciso e sempre estarem presentes para mim. Vocês me deram a coragem de sonhar e a força para seguir sempre em frente.

Agradeço minha irmã, Ana Elisa, pelo apoio, pela querida amizade e companheirismo. Principalmente por estar sempre por perto. E Tobbei Jr pela genuína inocência e amizade.

Agradeço especialmente à Julia Vessoni, minha namorada, companheira, melhor amiga e parceira, que nunca perdeu a fé em mim, mesmo quando eu havia perdido. Obrigado por ter acreditado e, com seu dom singular de transformar meus dias em dias inesquecíveis e de vitória, independente das dificuldades enfrentadas, fez todo o caminho possível.

Agradeço todos meus amigos, que estão presentes em cada um dos momentos de minha história. A presença de vocês sempre faz a diferença em minha vida. Especialmente Danilo Madalena, Guilherme Caes e Naiara Thomaz, por toda a ajuda fornecida especificamente com a tese e também por estarem comigo desde o dia zero.

Agradeço aos meus amigos e companheiros de mestrado, Alex Melges, Márcia Richtielle e Eliton Moro por todos os momentos que dividimos ao longo desse período cheio de emoções e desafios.

Agradeço as professoras Suzete Afonso, Maria do Carmo Carbinatto e ao professor Thiago de Melo, pelo exemplo inspirador de profissionais que são. Contribuindo tanto para meu crescimento acadêmico quanto pessoal, vocês são verdadeiros modelos para mim.

Agradeço principalmente minha orientadora, professora Thaís Monis que nunca mediu esforços, tampouco poupou dedicação para ensinar, pesquisar e me guiar durante esse processo. Sou muito grato por todos nossos estudos e tenho muito orgulho do que a professora me ajudou a construir.

Agradeço imensamente ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, por possibilitar a existência desse trabalho por meio do auxílio financeiro.

*"Walk on
With hope
In your heart
And you'll never walk alone"
Rodgers e Hammerstein*

Lista de Figuras

2.1	Intervalos representados através da decoração com tracinho.	25
2.2	À esquerda, o módulo de intervalo $\mathbb{I}(1, 3]$ no semi-plano estendido $\bar{\mathcal{H}}$. À direita, uma representação com cada um dos tipos de ponto decorado com tracinho.	26
2.3	O tradicional exemplo da teoria de persistência: A curva X suavemente imersa no plano e a função altura f demarcando alguns valores avaliados.	27
2.4	Pontos decorados detectados pela medida.	35
2.5	Ilustração da propriedade de aditividade.	38
2.6	Decomposição do retângulo S em outros cinco retângulos, incluindo R como um deles.	41
2.7	Decomposição do retângulo S em outros quatro retângulos, incluindo R como um deles.	42
2.8	Decomposição do retângulo S em outros três retângulos, incluindo R como um deles.	42
2.9	Decomposição do retângulo S em outros dois retângulos, incluindo R como um deles.	43
2.10	O retângulo $S = R$ é sua própria decomposição	43
3.1	O Toro de Pé.	45
3.2	Evolução dos conjuntos de subnível do Toro de Pé.	46
4.1	A Regra do Mais Velho.	60
4.2	A classe γ nasce em K_i , uma vez que não jaz na imagem de $H_p(K_{i-1})$ (sombreada). Além disso, γ morre ao entrar em K_j pois é o primeiro momento em que sua imagem se funde para dentro da imagem de $H_p(K_{i-1})$.	62
4.3	Complexo simplicial	66
4.4	A redução da matriz bordo vista como produto de matrizes. O elementos presentes na área em branco são necessariamente 0, enquanto na área sombreada pode ser 0 ou 1.	70
4.5	Ilustração das submatrizes R_i^j , R_{i+1}^j , R_{i+1}^{j-1} e R_i^{j-1} da matriz R	71
4.6	se $R[i, j]$ for lowest 1, então R_i^j possui uma coluna não-nula a mais do que as três submatrizes.	72
4.7	$R[i, j]$ não é lowest 1 e nenhuma das colunas de 1 à $j - 1$ possuem lowest 1 na linha i	73
4.8	$R[i, j]$ não é lowest 1 e uma das colunas de 1 à $j - 1$ possui lowest 1 na linha i	73
4.9	O complexo simplicial triângulo.	75

4.10	A filtração do complexo simplicial triângulo. Inicialmente temos o conjunto vazio, que representa o falso simplexo indexado em 0, após isso adicionamos os vértices, as arestas e finalmente a face, portanto conseguimos o triângulo.	75
4.11	Da esquerda para a direita temos ilustrado os diagramas de persistência Dgm_{-1} , Dgm_0 e Dgm_1	78
4.12	Complexo simplicial.	81
4.13	As filtrações do complexo simplicial e a Regra do Mais Velho.	82
4.14	Os dois casos estudados para o 0-ésimo diagrama.	83
5.1	Cada linha armazena o valor da função definida nos simplexos enquanto o parâmetro t aumenta. Para qualquer $t \in [0, 1]$ fixado, temos f_t através da interseção com correspondente reta vertical.	94
5.2	A matriz R exibindo a deficiência ressaltada no caso 1.	96
5.3	Após somar as colunas indicadas, corrigimos a matriz R para o caso 1.	96
5.4	Ao final do processo matriz R' é reduzida, para o caso 1.	97
5.5	A matriz R exibindo a deficiência ressaltada para o caso 2.	97
5.6	Após somar as colunas indicadas, corrigimos a matriz R para o caso 2	97
5.7	Apesar de R não ser reduzida, R' é, por conta da aplicação da matriz P	97
5.8	Deficiência ressaltada na matriz R no caso 3.	98
5.9	Deficiência na matriz R após multiplicação por S , caso 3.	98
5.10	Deficiência se propaga e não obtemos $PRSP$ reduzida.	99
5.11	A deficiência da matriz R no caso 4.	99
5.12	A deficiência da matriz R após aplicar a matriz S , no caso 4.	99
5.13	A matriz $PRSP$ não é reduzida para o caso 4.	99
5.14	A família de diagramas de persistencia à 1-parâmetro da homotopia linear entre $f_0 = f$ e $f_1 = g$. Alguns pontos geram vinhas que percorrem todas as camadas, enquanto outras morrem.	103
6.1	Pontos com mesmas cores possuem mesmo vetor normal à curva.	118
6.2	À esquerda, as indicações em quais pontos da curva ocorrerá a cirurgia. À direita, cirurgia realizada e os pontos extremos de mesma cor serão identificados, gerando a nova curva N	119
6.3	Exemplo de representação de proteínas e suas diferentes protrusões e concavidades.	122

Sumário

1	Introdução	17
2	Categoria e estrutura dos módulos de persistência	19
2.1	Multiconjuntos	19
2.2	Módulos de Persistência	20
2.2.1	Pré-requisitos e Definições	21
2.3	Conjuntos de índices mais gerais	22
2.4	Categorias de Módulos	22
2.5	Módulos de Intervalo	23
2.6	O diagrama de persistência de módulos decomponíveis	26
2.7	Cálculos de Quiver	29
2.8	Medidas no Retângulo	34
2.8.1	A Medida de Persistência	34
2.8.2	Continuação da Medida de Persistência	37
2.9	r -medidas abstratas	40
3	Topologia Algébrica Computacional	45
3.1	Funções de Morse	45
3.1.1	O Toro em Pé	45
3.2	Funções Suaves	46
3.3	O Lema de Morse	47
3.4	Campo de Vetor Gradiente	51
3.5	Anexar Células	52
3.6	Transversalidade	53
3.6.1	Curvas Integrais	53
3.6.2	Variedades Estáveis e Instáveis	54
3.6.3	Funções de Morse-Smale	55
3.6.4	Complexos	55
3.6.5	Desigualdade de Morse	56
3.7	Funções Lineares por Partes	57
3.7.1	Funções Simpliciais	57
3.8	Filtração de Lower Star	58
4	Topologia Persistente Computacional	59
4.1	Homologia Persistente	59
4.1.1	Persistência	59
4.1.2	Homologia Persistente	59
4.1.3	Diagramas de Persistência	63

4.1.4	Redução Matricial	65
4.2	Implementação Eficiente	78
4.2.1	Representação da Matriz Esparsa	79
4.2.2	Análise do Algoritmo	79
4.2.3	Implementação da Matriz Esparsa	80
4.3	O 0-ésimo Diagrama	82
4.4	Superfícies	84
4.5	O Primeiro Diagrama	86
4.6	Persistência Estendida	86
4.6.1	Teorema da Dualidade de Lefschetz	86
4.7	Elevação em uma superfície	87
4.8	Filtração Estendida	88
4.8.1	Dualidade e Simetria	91
5	Estabilidade	93
5.1	Famílias a 1-parâmetro	93
5.1.1	Homotopia Linear	93
5.2	Decomposição Matricial	95
5.2.1	Atualização da Decomposição	95
5.2.2	Como o pareamento se altera	100
5.3	Teorema de Estabilidade	100
5.3.1	Distância Bottleneck	100
5.4	Estabilidade Bottleneck	102
5.4.1	Funções Tame	105
6	Aplicações	107
6.1	Medidas para Dados de Expressão Genética	107
6.1.1	Pré-requisitos	107
6.1.2	Funções Lipschitzianas no Círculo	108
6.1.3	Simplificações	109
6.1.4	Medidas	110
6.1.5	Instabilidade para pequeno grau	112
6.1.6	Estabilidade para grau pelo menos 2	112
6.1.7	Notas	115
6.2	Elevação para Ancoragem de Proteína	116
6.2.1	Curvas no Plano	116
6.2.2	Função Altura e Função Elevação	117
6.2.3	Representação da Proteína	119
6.2.4	Função Pontuação	120
6.2.5	Características	120
6.2.6	Alinhamento Grosseiro	121
	Referências	123

1 Introdução

Uma ferramenta popular em Topological Data Analysis (TDA), a Homologia Persistente é um método algébrico para medir características topológicas de dados. Em outras palavras, a homologia persistente realiza o intercâmbio entre a informação contida em um conjunto discreto de pontos, munido de uma métrica, com a representação de uma estrutura algébrica, como componentes, buracos e estrutura de gráfico.

Sua função, em essência, é representada pelas respostas que fornece para perguntas como “quais características topológicas determinados dados exibem?”, além de ser capaz de detectar periodicidade em dados de série temporal.

A homologia persistente tem sido amplamente estudada, seja quanto seus conceitos mais teóricos, como estrutura e estabilidade, bem como seus aspectos aplicados, desde algoritmos utilizados para que a extração da informação de dados seja cada vez mais precisa e eficiente, até suas implementações em campos diversos, como na Biologia.

O objetivo deste trabalho é apresentar a Homologia Persistente em sua amplitude e detalhes, exemplificando e afinando com suas aplicações possíveis que a utilizam como ferramenta. Serão desenvolvidos teoremas e resultados fundamentais, os quais edificam os alicerces da Homologia Persistente, bem como apresentaremos algoritmos computacionais, estudando sua ordem, complexidade e eficiência. Por fim, algumas aplicações que fazem uso de suas ferramentas serão demonstradas.

Para isso, a teoria será embasada inicialmente no artigo [2], de Frédéric Chazal, Vin de Silva, Marc Glisse e Steve Oudot, onde desenvolvemos os primeiros conceitos fundamentais para a teoria da Homologia Persistente: os multiconjuntos, que são o ambiente natural de representação de um diagrama de persistência; módulos de persistência; e cálculos Quiver. Nessa abordagem, a atuação se dá no sentido de como decodificar as informações contidas em um dado módulo de persistência.

Ao migrar para o livro “Computational Topology, an introduction”, de Herbert Edelsbrunner e John Harer (2009), [4], a abordagem estudada é diferente: a partir de uma função com determinadas características definida em um ambiente específico, exploramos a construção dos diagramas de persistência. Para isso, são descritas hipóteses assumidas sobre a função para que ela tenha um bom comportamento, além de aproximar os dados por um complexo simplicial, aonde será aplicado o processo de filtração a fim de obter a persistência. Para isso, serão abordados os seguintes tópicos: funções de Morse, Homologia Persistente, diagramas de persistência, dualidade e simetria e estabilidade.

É apresentado ainda, por meio desse livro, o algoritmo da matriz reduzida e sua implementação eficiente que visa demonstrar a execução, partindo de um conjunto de dados até a etapa final que é a obtenção do diagrama de persistência. Também é estudado o grau de complexidade e a quantidade de memória requerida.

O trabalho é encerrado com duas aplicações. A primeira aplicação é majoritariamente teórica, utilizando resultados para sua estruturação. A segunda é totalmente aplicada, ilustrando um algoritmo baseado na Homologia Persistente criado para obter as informações desejadas.

2 Categoria e estrutura dos módulos de persistência

2.1 Multiconjuntos

Diagramas de Persistência são mais do que conjuntos, eles são o que chamamos multiconjuntos.

Definição 2.1. Um multiconjunto $\mathcal{A} = (S, m)$ é um par que consiste em S um conjunto e $m : S \rightarrow \{1, 2, \dots\} \cup \{+\infty\}$ uma função multiplicidade que informa quantas vezes cada elemento de S ocorre em \mathcal{A} .

Definição 2.2. O cardinal de um multiconjunto $\mathcal{A} = (S, m)$ é definido como

$$\#\mathcal{A} = \sum_{s \in S} m(s),$$

que assume valores em $\{1, 2, \dots\} \cup \{+\infty\}$.

É importante ressaltar que não faremos distinção entre cardinais infinitos.

A interseção de dois multiconjuntos não é definida, no entanto, podemos restringir um multiconjunto \mathcal{A} a um conjunto B ,

$$\mathcal{A}|_B = (S \cap B, m|_{S \cap B}),$$

donde podemos abusar da notação e escrever $\mathcal{A} \cap B$.

Observação 2.1. Não permitimos que a função multiplicidade alcance o valor zero, pois esperamos que ela conte a ocorrência de cada elemento no conjunto.

Um par (B, m) no qual B é um conjunto e $m : B \rightarrow \{0, 1, 2, \dots\} \cup \{\infty\}$ é uma função define implicitamente um multiconjunto, a saber, $\mathcal{A} = (S, m|_S)$, tal que $S = B - m^{-1}(0)$.

Exemplo 2.1. Seja o par $(\{0, 1, 2, 3, 4\}, m)$, onde

$$m(x) = \begin{cases} 2, & \text{se } x \text{ for par} \\ 0, & \text{se } x \text{ for ímpar} \end{cases},$$

para $x \in \{0, 1, 2, 3, 4\}$. Assim, $m^{-1}(0) = \{1, 3\}$. Por conseguinte, $S = \{0, 1, 2, 3, 4\} - \{1, 3\} = \{0, 2, 4\}$. Temos o multiconjunto constituído pelo conjunto $S = \{0, 2, 4\}$ e pela função m que agora admite valores no conjunto $\{1, 2, \dots\} \cup \{\infty\}$.

Sejam $\mathcal{A} = (S, m)$ um multiconjunto, B um conjunto e $f : S \rightarrow B$ uma função. Tomamos o subconjunto $f(S) = \{f(a) \mid a \in S\}$ de B e construímos um multiconjunto \mathcal{B} a partir do multiconjunto dado definindo a função multiplicidade $m' : f(S) \rightarrow \{1, 2, \dots\} \cup \{\infty\}$ como

$$m'(b) = \sum_{s \in f^{-1}(b)} m(s).$$

Observação 2.2. De fato, temos uma função multiplicidade já que estamos transportando os elementos de $f(S)$ para o conjunto S através da imagem inversa de f , onde por hipótese temos um multiconjunto, ou seja, em S temos a função multiplicidade m . Logo, contar a ocorrência de um elemento b em $f(S)$ corresponde a contar quantos elementos s do domínio de f , S , são associados a ele, s em S , tal que $f(s) = b$, juntamente com a quantidade de vezes em que aparecem em S , $m(s)$. Portanto, $\mathcal{B} = (f(S), m')$ é um multiconjunto.

2.2 Módulos de Persistência

Em todo esse capítulo, \mathbb{K} é um corpo fixado.

Definição 2.3. Um módulo de persistência \mathbb{V} sobre os números reais \mathbb{R} é definido como sendo uma família a 1-parâmetro de espaços vetoriais

$$\mathbb{V} := (V_t \mid t \in \mathbb{R}),$$

juntamente com uma família a 2-parâmetros de transformações lineares

$$(v_s^t : V_s \rightarrow V_t \mid s \leq t),$$

que satisfazem a lei de decomposição

$$v_s^t \circ v_r^s = v_s^t v_r^s = v_r^t$$

sempre que $r \leq s \leq t$ e onde $v_s^s = Id$ a transformação identidade de V_s .

Exemplo 2.2. Sejam X um espaço topológico e $f : X \rightarrow \mathbb{R}$ uma função, não necessariamente contínua. Considere o conjunto de subníveis

$$X_t = (X, f)_t = \{x \in X \mid f(x) \leq t\}.$$

Então,

$$X_s \subseteq X_t \text{ sempre que } s \leq t.$$

Considere a família a 2-parâmetros de inclusões $(i_s^t : X_s \rightarrow X_t \mid s \leq t)$.

As aplicações inclusão trivialmente satisfazem a lei de decomposição

$$i_s^t \circ i_r^s = i_r^t \text{ sempre que } r \leq s \leq t \text{ e ainda } i_t^t = Id : V_t \rightarrow V_t.$$

Seja $H = H_p(\cdot, \mathbb{K})$ o funtor de homologia singular em dimensão p , com coeficientes no corpo \mathbb{K} . Defina, para cada $t \in \mathbb{R}$,

$$V_t := H(X_t),$$

e para cada $s \leq t$, o homomorfismo

$$H(i_s^t) : V_s \rightarrow V_t$$

induzido pela inclusão. Esse é o primeiro exemplo canônico de módulo de persistência.

Em Topologia Aplicada há diversos exemplos de (X, f) do qual a homologia persistente é de interesse em se estudar e entender. Muito frequentemente X é um complexo simplicial finito e cada X_t forma um subcomplexo. Assim, os espaços vetoriais $H(X_t)$ são de dimensão finita e, ao passo em que t cresce, existem apenas finitos valores críticos, isto é, valores onde o complexo modifica seu comportamento, aumentando em uma ou mais novas células.

2.2.1 Pré-requisitos e Definições

Definimos $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$, a reta real estendida. Utilizamos a seguinte regra de operação na reta estendida: $\forall x \in \mathbb{R}$, temos $x + \infty = +\infty$ e $x - \infty = -\infty$.

O plano estendido $\bar{\mathbb{R}}^2 = \mathbb{R} \times \mathbb{R}$ é dotado da norma l_∞ ,

$$\|\cdot\|_\infty: \bar{\mathbb{R}}^2 \longrightarrow \bar{\mathbb{R}}$$

definida por

$$x = (x_1, x_2) \mapsto \sup_{n=1,2} |x_n|.$$

A topologia induzida por $\|\cdot\|_\infty$ em $\bar{\mathbb{R}}^2$ é tal que os pontos de \mathbb{R}^2 , de $\{\pm\infty\} \times \mathbb{R}$, de $\mathbb{R} \times \{\pm\infty\}$ e de $\{\pm\infty\} \times \{\pm\infty\}$ formam componentes conexas distintas.

Seja $\Delta = \{(x, x) \mid x \in \bar{\mathbb{R}}\}$ a diagonal e $\Delta_+ = \{(x, y) \mid y \geq x\}$ o semi-plano fechado, acima de Δ . Mais geralmente, para qualquer $\delta > 0$, definimos $\Delta_+^\delta = \{(x, y) \in \bar{\mathbb{R}}^2 \mid y \geq x + 2\delta\}$ o semi-plano fechado, acima de Δ , distando δ de Δ segundo a distância l_∞ .

Os objetos matemáticos principais na Persistência Topológica são os denominados diagramas de persistência (ou barcodes), os quais são multiconjuntos em $\bar{\mathbb{R}}^2$.

Seja \mathcal{D} um multiconjunto munido da função multiplicidade m associada. O suporte de \mathcal{D} é entendido como sendo o multiconjunto \mathcal{D} , considerado sem as multiplicidades e é denotado por $|\mathcal{D}|$. Nessa linguagem, \mathcal{D} pode ser representado como a união disjunta

$$\mathcal{D} = \bigcup_{p \in |\mathcal{D}|} \bigsqcup_{i=1}^{m(p)} p.$$

Definição 2.4. *Dados dois multiconjuntos \mathcal{D} e \mathcal{D}' , a distância de Hausdorff entre seus suportes é definida por*

$$d_{\mathcal{H}}^\infty(|\mathcal{D}|, |\mathcal{D}'|) = \max\left\{ \sup_{p \in |\mathcal{D}|} \inf_{p' \in |\mathcal{D}'|} \|p - p'\|_\infty ; \sup_{p' \in |\mathcal{D}'|} \inf_{p \in |\mathcal{D}|} \|p - p'\|_\infty \right\}$$

Por simplicidade, abusamos da notação e escrevemos $d_{\mathcal{H}}^\infty(\mathcal{D}, \mathcal{D}')$ para a distância de Hausdorff $d_{\mathcal{H}}^\infty(|\mathcal{D}|, |\mathcal{D}'|)$. É importante observar que a distância de Hausdorff não é uma distância entre multiconjuntos e sim de seus suportes.

Definição 2.5. *Dado um espaço topológico \mathbb{X} e uma função $f : \mathbb{X} \rightarrow \mathbb{R}$, a filtração em subníveis de f é definida por*

1. A família de subespaços $\{\mathbb{X}_\alpha^f\}_{\alpha \in \mathbb{R}}$ onde, para todo $\alpha \in \mathbb{R}$,

$$\mathbb{X}_\alpha^f = f^{-1}(-\infty, \alpha) \subset \mathbb{X}.$$

2. A família canônica de funções inclusão $i_\alpha^{\alpha'} : \mathbb{X}_\alpha^f \rightarrow \mathbb{X}_{\alpha'}^f$, desde que $\alpha \leq \alpha'$.

Definição 2.6. *Sejam \mathbb{X} um espaço topológico e $f : \mathbb{X} \rightarrow \mathbb{R}$ uma função. Um valor crítico homológico de f é um elemento $\alpha \in \mathbb{R}$ para o qual existe um $k \in \mathbb{Z}$ tal que, para todo $\epsilon > 0$ suficientemente pequeno, o homomorfismo*

$$H(i_{\alpha-\epsilon}^{\alpha+\epsilon}) : H_k(\mathbb{X}_{\alpha-\epsilon}^f) \rightarrow H_k(\mathbb{X}_{\alpha+\epsilon}^f)$$

não é um isomorfismo.

Definição 2.7. *Dados um espaço topológico \mathbb{X} e uma função $f : \mathbb{X} \rightarrow \mathbb{R}$, dizemos que f é tame se ela possuir apenas uma quantidade finita de valores críticos homológicos e, ainda, se os grupos de homologia $H_k(\mathbb{X}_\alpha^f)$ forem de dimensão finita, para todo $k \in \mathbb{Z}$ e qualquer $\alpha \in \mathbb{R}$.*

2.3 Conjuntos de índices mais gerais

Podemos formalmente definir módulos de persistência sobre qualquer conjunto parcialmente ordenado \mathbf{T} do mesmo modo como fizemos para \mathbb{R} , especificando a família de espaços vetoriais indexados,

$$(V_t \mid t \in \mathbf{T})$$

e as transformações lineares, indexadas a 2-parâmetros,

$$(v_s^t \mid s \leq t; s, t \in \mathbf{T}),$$

satisfazendo $v_s^t = v_r^t v_s^r$, sempre que $s \leq r \leq t$ e $v_t^t = Id_{V_t}$, para todo t .

A coleção resultante de dados é chamada um \mathbf{T} -módulo de persistência ou um módulo de persistência sobre \mathbf{T} .

Se \mathbb{V} for um \mathbf{T} -módulo de persistência e $\mathbf{S} \subset \mathbf{T}$ um subconjunto, então conseguimos um \mathbf{S} -módulo de persistência considerando apenas aqueles espaços e transformações lineares cujos índices pertencem a \mathbf{S} . Isso é chamado de restrição de \mathbb{V} em \mathbf{S} , denotado por $\mathbb{V}_{\mathbf{S}}$. Mais comumente, trabalhamos com subconjuntos localmente finitos $\mathbf{T} \subset \mathbb{R}$, ou seja, aqueles subconjuntos sem pontos de acumulação em \mathbb{R} . Coletamos a informação sobre um \mathbb{R} -módulo de persistência, considerando sua restrição à subconjuntos localmente finitos. Isso funciona bem pois módulos de persistência sobre $\{1, 2, \dots, n\}$ ou sobre os inteiros \mathbb{Z} são bem entendidos e conhecidos.

2.4 Categorias de Módulos

Um homomorfismo Φ entre dois \mathbf{T} -módulos de persistência \mathbb{U} e \mathbb{V} é uma coleção de transformações lineares

$$(\Phi_t : U_t \rightarrow V_t \mid t \in \mathbf{T})$$

tal que o diagrama

$$\begin{array}{ccc} U_s & \xrightarrow{\Phi_s} & V_s \\ u_s^t \downarrow & & \downarrow v_s^t \\ U_t & \xrightarrow{\Phi_t} & V_t \end{array}$$

comuta para todo $s \leq t$.

Definimos

- $Hom(\mathbb{U}, \mathbb{V}) = \{\phi \mid \mathbb{U} \xrightarrow{\phi} \mathbb{V} \text{ é homomorfismo}\}$.
- $End(\mathbb{V}) = \{\phi \mid \mathbb{V} \xrightarrow{\phi} \mathbb{V} \text{ é homomorfismo}\}$.

Dados $\phi^1, \phi^2 \in Hom(\mathbb{U}, \mathbb{V})$, a soma $\phi^1 + \phi^2$ é definida por: para cada $t \in \mathbf{T}$, $(\phi^1 + \phi^2)_t = \phi_t^1 + \phi_t^2$. Também, dado α no corpo \mathbb{K} , $\alpha\phi^1$ é definido por: para cada $t \in \mathbf{T}$, $(\alpha\phi^1)_t = \alpha\phi_t^1$. Temos ainda definida a composição de homomorfismos de \mathbf{T} -módulos de persistência: dados $\phi \in Hom(\mathbb{U}, \mathbb{V})$ e $\psi \in Hom(\mathbb{V}, \mathbb{W})$, $\psi \circ \phi$ é definido por: para cada $t \in \mathbf{T}$, $(\psi \circ \phi)_t = \psi_t \circ \phi_t$. É fácil ver que:

- Se $\phi^1, \phi^2 \in Hom(\mathbb{U}, \mathbb{V})$, então $\phi^1 + \phi^2 \in Hom(\mathbb{U}, \mathbb{V})$;
- Se $\phi \in Hom(\mathbb{U}, \mathbb{V})$ e $\alpha \in \mathbb{K}$, então $\alpha\phi \in Hom(\mathbb{U}, \mathbb{V})$;
- Se $\phi \in Hom(\mathbb{U}, \mathbb{V})$ e $\psi \in Hom(\mathbb{V}, \mathbb{W})$, então $\psi \circ \phi \in Hom(\mathbb{U}, \mathbb{W})$.

Com as definições acima, temos a seguinte estrutura.

Proposição 2.1. *$End(\mathbb{V})$ munido da operação de composição de endomorfismos é uma \mathbb{K} -álgebra.*

Demonstração. Sejam $f, g, h \in End(\mathbb{V})$ e $\alpha, \beta \in \mathbb{K}$ quaisquer. Devemos provar que:

- $f \circ (h + g) = f \circ h + f \circ g$;
- $(h + g) \circ f = h \circ f + g \circ f$;
- $(\alpha f) \circ (\beta g) = (\alpha\beta)(f \circ g)$.

Mas esses três itens são consequências imediatas de que, para todo $t \in \mathbf{T}$,

- $f_t \circ (h + g)_t = f_t \circ h_t + f_t \circ g_t$;
- $(h + g)_t \circ f_t = h_t \circ f_t + g_t \circ f_t$;
- $(\alpha f)_t \circ (\beta g)_t = (\alpha\beta)(f_t \circ g_t)$.

□

2.5 Módulos de Intervalo

O alicerce fundamental da persistência são os módulos de intervalo.

Uma abordagem que foi estudada visa entender módulos de persistência através de sua decomposição em intervalos, entretanto isso não é sempre possível, contudo é possível com frequência suficiente para nossos propósitos.

Primeiramente, seja $\mathbf{T} \subset \mathbb{R}$ e seja $S \subset \mathbf{T}$. Dizemos que S é um intervalo de \mathbf{T} , ou que S é um \mathbf{T} -intervalo, se

$$s, t \in S, r \in \mathbf{T}, s < r < t \implies r \in S.$$

Essa é uma noção generalizada de intervalo.

Exemplo 2.3. Seja $\mathbb{Z} \subset \mathbb{R}$ o conjunto dos inteiros e considere $S = \{-1, 0, 1, 2, 3, 4\}$. Então, S é um intervalo de \mathbb{Z} .

Exemplo 2.4. O próprio conjunto dos números inteiros é um intervalo de si mesmo.

Assim, temos as noções de \mathbb{N} -intervalo, \mathbb{Z} -intervalo, \mathbb{Q} -intervalo, etc.

Seja $\mathbf{T} \subset \mathbb{R}$ e considere $J \subset \mathbf{T}$ um \mathbf{T} -intervalo. Então \mathbb{I}^J é definido como sendo o \mathbf{T} -módulo de persistência onde os espaços vetoriais são dados por

$$I_t = \begin{cases} \mathbb{K}, & \text{se } t \in J \\ 0, & \text{caso contrário} \end{cases}$$

e as transformações lineares são

$$i_s^t = \begin{cases} \mathbf{1}, & \text{se } s, t \in J, \text{ com } s \leq t \\ 0, & \text{caso contrário} \end{cases}.$$

Ressaltamos que $\mathbf{1}$ representa a transformação linear identidade entre os espaços vetoriais.

Em linguagem informal, \mathbb{I}^J simboliza a característica que “persiste” sobre o intervalo J e se ausenta fora deste. Escreveremos $\mathbb{I}_{\mathbf{T}}^J$ quando desejamos especificar o conjunto de índices sem ambiguidades.

Nomearemos \mathbb{I}^J módulo de intervalo.

Se \mathbf{T} for um subconjunto localmente finito de \mathbb{R} , isto é, sem pontos de acumulação em \mathbb{R} , então, qualquer intervalo limitado de \mathbf{T} contém seus pontos extremos e, assim, os escreveremos como intervalos fechados, enquanto os intervalos ilimitados serão intervalos abertos ou semi-abertos. Por exemplo, no caso de um \mathbb{Z} -módulo de persistência, existem apenas quatro tipos de módulo de intervalo: $\mathbb{I}^{[m, n]}$, $\mathbb{I}^{(-\infty, n]}$, $\mathbb{I}^{[m, +\infty)}$ e $\mathbb{I}^{(-\infty, +\infty)}$. Para deixar ainda mais explícito o intervalo, escreveremos $\mathbb{I}[m, n]$, $\mathbb{I}(-\infty, n]$, $\mathbb{I}[m, +\infty)$ e $\mathbb{I}(-\infty, +\infty)$.

Para módulos de persistência sobre \mathbb{R} , torna-se importante distinguir intervalos com o mesmo ponto extremo contudo diferentes topologias (aberto, fechado, semi-aberto). Para nosso propósito, introduziremos os números reais decorados, que são escritos como números reais ordinários porém com um subscrito “+” ou “-”. Intervalos finitos, adaptamos para a seguinte notação

$$(p^-, q^-) = [p, q]$$

$$(p^-, q^+) = [p, q]$$

$$(p^+, q^-) = (p, q)$$

$$(p^+, q^+) = (p, q]$$

onde $p < q$, exceto no caso particular de (r^-, r^+) , que se refere ao intervalo degenerado $[r, r]$. Incluiremos também os símbolos $-\infty$ e $+\infty$ para intervalos infinitos. Desde que intervalos reais são sempre abertos no infinito, estes carregam implicitamente o sobrescrito $-\infty^+$ e $+\infty^-$. No entanto, usualmente os omitiremos. Assim a notação de intervalo se estende de maneira natural para $(-\infty, q^+) = (-\infty, q]$ e $(p^-, +\infty) = [p, +\infty)$, por exemplo.

Quando quisermos nos referir ao número real decorado, mas não soubermos sua decoração, utilizaremos um asterisco. Portanto, p^* pode significar p^+ ou p^- . A notação

para um intervalo arbitrário será (p^*, q^*) , onde a ordenação dos números decorados $p^* < q^*$ é igual a dos reais propriamente ditos.

Existem algumas convenções visuais para representar módulos de intervalo sobre \mathbb{R} . Trabalhamos no semi-plano

$$\mathcal{H} = \{(p, q) \in \mathbb{R}^2 \mid p \leq q\}$$

dos pontos de \mathbb{R}^2 que ficam em, ou acima da, diagonal principal.

Um módulo de intervalo finito $\mathbb{I}(p^*, q^*)$ pode ser representado de diversas formas:

- i Como um intervalo na reta real;
- ii A coleção dos ranks das aplicações i_s^t , vista como uma função de $\mathcal{H} \rightarrow \{0, 1\}$;
- iii Como um ponto (p, q) em \mathcal{H} com um tracinho que especifica sua decoração;
- iv Etc...

Com relação ao item (ii), seja $\mathbb{I}(p^*, q^*)$ um módulo de intervalo sobre \mathbb{R} e consideremos a aplicação

$$t \in \mathbb{R} \rightarrow I_t = \begin{cases} \mathbb{K}, & \text{se } t \in (p^*, q^*) \\ 0, & \text{caso contrário} \end{cases}.$$

Primeiro, estudamos o $rank(i_s^t) = \dim(Im(i_s^t))$, visto como uma função definida do semi-plano \mathcal{H} a valores em $\{0, 1\}$. Conhecidos os valores de $rank(i_s^t)$ para cada t fixado, construímos o conjunto

$$J = \{s \mid rank(i_s^t) = 1\}.$$

Afirmamos que J é um intervalo, isto é, para $s \in J$, se $s \leq r \leq t$ então $r \in J$. De fato, dado que $s \leq r \leq t$, temos

$$i_s^t = i_r^t \circ i_s^r.$$

Agora, como $s, t \in J$, temos $rank(i_s^t) = 1$. Por conseguinte, $rank(i_r^t) = 1$, então $I_r = \mathbb{K}$, o que implica que $r \in J$. Concluímos, assim, que J é um intervalo.

Já com relação ao terceiro item, a representação de um ponto (p, q) em \mathcal{H} , segue as formas de decoração com tracinho:

$$\begin{array}{cc} (p^-, q^+) = \bullet & (p^+, q^+) = \bullet \\ (p^-, q^-) = \bullet & (p^+, q^-) = \bullet \end{array}$$

Figura 2.1: Intervalos representados através da decoração com tracinho.

Temos por convenção que o ponto com decoração com tracinho sugere em qual quadrante o elemento se encontra através da posição da decoração.

Existe também uma quarta opção para representar o ponto (p, q) que seria sem fazer uso da decoração. A não utilização da decoração indicativa é a representação clássica de diagramas de persistência, além de ser adequada para a maioria dos fins. No entanto, a precisão extra provida pela decoração é importante para a correspondência entre diagramas e medidas. Para representar um intervalo infinito como um ponto decorado ou sem decoração, trabalharemos no semi-plano estendido

$$\bar{\mathcal{H}} = \mathcal{H} \cup (\{-\infty\} \times \mathbb{R}) \cup (\mathbb{R} \times \{-\infty\}) \cup \{(-\infty, +\infty)\},$$

o qual pode ser desenhado esquematicamente como um triângulo.

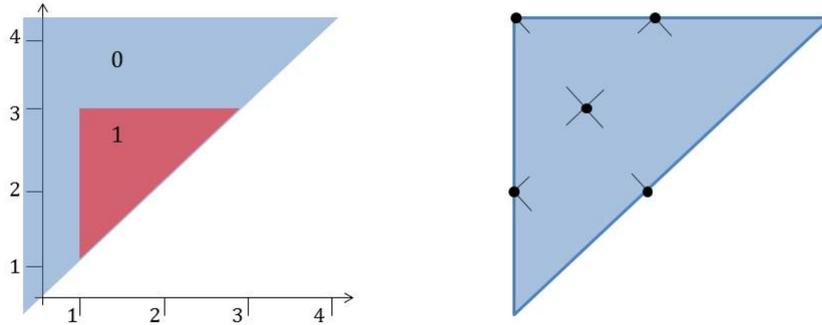


Figura 2.2: À esquerda, o módulo de intervalo $\mathbb{I}(1, 3]$ no semi-plano estendido $\bar{\mathcal{H}}$. À direita, uma representação com cada um dos tipos de ponto decorado com tracinho.

2.6 O diagrama de persistência de módulos decomponíveis

Se um módulo de persistência \mathbb{V} indexado sobre \mathbb{R} pode ser decomposto como

$$\mathbb{V} = \bigoplus_{l \in L} \mathbb{I}(p_l^*, q_l^*)$$

então definimos o diagrama de persistência decorado de \mathbb{V} como sendo o multiconjunto

$$Dgm(\mathbb{V}) = Int(\mathbb{V}) = \{(p_l^*, q_l^*) \mid l \in L\},$$

e também definimos o diagrama de persistência não-decorado como sendo o multiconjunto

$$dgm(\mathbb{V}) = int(\mathbb{V}) = \{(p_l, q_l) \mid l \in L\} - \Delta,$$

onde $\Delta = \{(r, r) \mid r \in \mathbb{R}\}$ é a diagonal no plano.

Note que $Dgm(\mathbb{V})$ e $dgm(\mathbb{V})$ não dependem da decomposição de \mathbb{V} , uma vez que o teorema de Krull-Remak-Schmidt-Azumaya (ver [1]) nos garante que caso \mathbb{V} seja descrito como duas somas distintas de módulos de intervalo, então elas são a mesma a menos de permutação.

Perceba que Dgm é um multiconjunto de pontos decorados em $\bar{\mathcal{H}}$, enquanto dgm é um multiconjunto de pontos não-decorados no interior de $\bar{\mathcal{H}}$. Por interior queremos dizer que excluimos a diagonal, porém mantemos os pontos no infinito.

A informação contida no dgm é exatamente a informação que nos interessará mais tarde, quando estudarmos a distância “Bottleneck”.

Apresentaremos dois teoremas importantes que caracterizam os módulos de persistência decomponíveis. Ambos resultados estão presentes em [2] e [3] e sua relevância é enorme já que o primeiro determina a unicidade da decomposição em módulos de intervalo enquanto o segundo nos fornece condições para decidir sobre a possibilidade de decompor determinado módulo de persistência em módulo de intervalos.

Teorema 2.1. Krull-Remak-Schmidt-Azumaya

Suponha que um módulo persistente \mathbb{V} sobre $\mathbf{T} \subset \mathbb{R}$ possa ser expressado como soma direta de módulos de intervalo de duas formas diferentes

$$\mathbb{V} = \bigoplus_{l \in L} \mathbb{I}^{J_l} = \bigoplus_{m \in M} \mathbb{I}^{K_m}$$

Então existe uma bijeção $\sigma : L \rightarrow M$ tal que $J_l = K_{\sigma(l)}$ para todo l .

Em outras palavras, podemos decompor um dado módulo persistente \mathbb{V} como soma direta de módulos de intervalo de forma única exceto por automorfismos. Mas em quais condições tal decomposição existe? Para responder a essa pergunta, temos o resultado seguinte que exhibe requisitos necessários sobre o conjunto de índices indexados do módulo de persistência e também a respeito da dimensão dos seus espaços vetoriais.

Teorema 2.2. Gabriel, Auslander, Ringel-Tachikawa, Webb

Seja \mathbb{V} módulo de persistência sobre $\mathbf{T} \subseteq \mathbb{R}$. Poderemos decompor \mathbb{V} como soma direta de módulos de intervalos.

- i Se \mathbf{T} for um conjunto finito
- ii Se \mathbf{T} for um subconjunto localmente finito de \mathbb{R} e ainda, cada V_t tiver dimensão finita

Exemplo 2.5. Um exemplo tradicional em teoria de persistência é uma curva X suavemente imersa no plano \mathbb{R}^2 , juntamente com $f : X \rightarrow \mathbb{R}$, denominada função altura, que associa a cada ponto da curva $(x, y) \in X$ sua coordenada $f(x, y) = y$. A filtração obtida através da função altura nos fornece, para cada $t \in \mathbb{R}$, os conjuntos de subníveis

$$(X, f)^t = X^t = \{x \in X \mid f(x) \leq t\} = f^{-1}(-\infty, t].$$

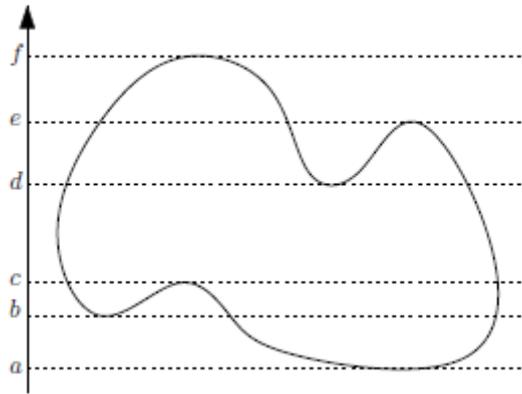


Figura 2.3: O tradicional exemplo da teoria de persistência: A curva X suavemente imersa no plano e a função altura f demarcando alguns valores avaliados.

Como visto anteriormente, temos $X^s \subseteq X^t$ desde que $s \leq t$, e assim temos as funções inclusão $i_s^t : X^s \hookrightarrow X^t$. Seja $H_0 = H_0(\cdot, \mathbb{R})$ o funtor de homologia singular de dimensão 0 com coeficientes em \mathbb{R} e, para cada $t \in \mathbb{R}$, definimos $V_t := H_0(X^t)$. Consequentemente, obtemos

$$\mathbb{V} = \{V_t \mid t \in \mathbb{R}\}.$$

Nosso objetivo agora é descrever \mathbb{V} como soma de módulos de intervalo. Estudaremos o comportamento dos espaços V_t conforme t cresce ao longo da reta \mathbb{R} .

Para $t < a$, $X^t = \emptyset$, por conseguinte $H_0(\emptyset) = 0$

Quando t alcança o valor a , o mínimo global da função altura avaliada na curva, obtemos o nascimento de uma classe de homologia de dimensão 0, que significa no nosso contexto o nascimento de uma componente conexa por caminhos, uma vez que $X^{t=a} = \{a\}$ o que implica

$$V_a = H_0(X^{t=a}) = H_0(\text{ponto}) = \mathbb{R}.$$

Enquanto tivermos $a \leq t < b$, obtemos $V_t = \mathbb{R}$ pois X^t será conexo por caminhos para esses valores.

Uma evolução nos subníveis ocorre em $t = b$, mínimo local da função altura, gerando o nascimento de outra componente conexa por caminhos, logo

$$V_b = H_0(X^{t=b}) = \mathbb{R} \oplus \mathbb{R},$$

e essa característica de duas componentes permanece para $b \leq t < c$.

Já em $t = c$, máximo local da função altura, acontece a morte de uma das componentes - em breve, teremos um aval teórico que possibilitará descrever o comportamento de nascimento e morte das classes de homologia e, assim, teremos recursos para afirmar que a componente que morre é a que nasceu por último, no nosso caso, a que surge no instante $t = b$. Logo,

$$V_c = H_0(X^{t=c}) = \mathbb{R}.$$

Ainda notamos que para $c \leq t < d$, mantêm-se esse comportamento de uma componente conexa.

Mais uma vez a função altura atinge seu mínimo local, agora em $t = d$ e, nesse momento,

$$V_t = H_0(X^{t=d}) = \mathbb{R} \oplus \mathbb{R},$$

configurando o nascimento de uma nova classe de homologia de dimensão 0, já que temos duas componentes conexas por caminhos. Essa configuração $V_t = \mathbb{R} \oplus \mathbb{R}$ se mantém para $d \leq t < e$.

Ao assumir o valor de máximo local em $t = e$, temos a evolução no subnível $X^{t=e}$ com a morte de uma componente conexa por caminhos, pois no instante $t = e$ as duas componentes conexas por caminhos distintas se fundem, formando apenas uma. Assim, $V_e = H_0(X^{t=e}) = \mathbb{R}$.

Note que para todo $t \geq e$, $V_t = \mathbb{R}$ porque não existe mais alterações nos conjuntos de subnível que sejam sensíveis ao funtor de homologia singular de dimensão 0, H_0 . Contudo, no momento em que t admite o valor máximo global da função altura, f , ocorre o nascimento de uma classe de homologia singular de dimensão 1, que é perceptível somente segundo o funtor $H_1 = H_1(\cdot, \mathbb{R})$ aplicado no subnível da filtração $X^{t=f}$.

Para escrevermos \mathbb{V} como soma de intervalos, precisamos codificar as informações acima a fim de representá-las na forma $\mathbb{V}(p^*, q^*)$.

Assim, temos o nascimento de uma componente conexa por caminhos que surge em a . Com o desenvolvimento dos subníveis, ocorre o nascimento de outra componente conexa por caminhos no mínimo local b e que persiste até o instante em que acrescentamos o máximo local c , assim a segunda componente conexa por caminhos nasce

em b e morre ao entrar em c . Em símbolos, temos que essa componente conexa por caminhos perdura no intervalo $\mathbb{I}(b^-, c^-)$, recordando que (b^-, c^-) representa $[b, c)$, nos informando que nasce em b , contudo só se mantém até antes de c .

Similarmente, ao passarmos pelo mínimo local d , acrescenta-se uma outra componente conexa por caminhos que se mantém até passarmos pelo ponto de máximo local e , logo a persistência dessa segunda componente conexa é $\mathbb{I}(d^-, e^-)$, que codifica a informação de uma classe de homologia singular de dimensão 0 nascendo no instante em que entramos em d e morrendo ao entrar em e , por isso o intervalo é $[d, e)$.

Concluimos, finalmente, que $\mathbb{V} = \mathbb{I}(a^-, +\infty^-) \oplus \mathbb{I}(b^-, c^-) \oplus \mathbb{I}(d^-, e^-)$.

Podemos visualizar ainda o comportamento de \mathbb{V} por meio dos barcodes ou da árvore fundida.

Observação 2.3. Para funções de Morse definidas em uma variedade compacta com valores críticos a_i , $0 \leq i \leq n$, os intervalos são sempre semi-abertos do tipo $[a_i, a_j) = (a_i^-, a_j^-)$.

2.7 Cálculos de Quiver

Agora, criaremos a notação e as ferramentas algébricas para manipulação de módulos de persistência sobre um conjunto finito de índices.

Um módulo de persistência \mathbb{V} indexado sobre um subconjunto finito

$$\mathbf{T} : a_1 < a_2 < \dots < a_{n-1} < a_n$$

da reta real, pode ser concebido como um diagrama de n espaços vetoriais e $(n - 1)$ transformações lineares:

$$\mathbb{V} : V_{a_1} \rightarrow V_{a_2} \rightarrow \dots \rightarrow V_{a_{n-1}} \rightarrow V_{a_n}.$$

Esse diagrama é a **representação** do seguinte **Quiver**

$$\bullet \rightarrow \bullet \rightarrow \dots \rightarrow \bullet \rightarrow \bullet$$

Pelo Teorema 2.2 de Gabriel, Auslander, Ringel-Tachikawa e Webb, temos que \mathbb{V} se decompõe como soma finita de módulos de intervalo $\mathbb{I}(a_i, a_j)$.

No caso em que tivermos n pequeno, podemos representar esses módulos de intervalos pictoricamente. O próximo exemplo servirá para elucidar como.

Exemplo 2.6. Sejam $a \leq b \leq c$. Existem seis possíveis módulos de intervalos sobre $\{a, b, c\}$, que são:

$$\mathbb{I}[a, a] = \bullet_a \text{ --- } \circ_b \text{ --- } \circ_c$$

$$\mathbb{I}[b, b] = \circ_a \text{ --- } \bullet_b \text{ --- } \circ_c$$

$$\mathbb{I}[c, c] = \circ_a \text{ --- } \circ_b \text{ --- } \bullet_c$$

$$\mathbb{I}[a, b] = \bullet_a \text{ --- } \bullet_b \text{ --- } \circ_c$$

$$\mathbb{I}[b, c] = \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c$$

$$\mathbb{I}[a, c] = \bullet_a \text{ --- } \bullet_b \text{ --- } \bullet_c$$

Nessa notação, usamos os círculos preenchidos, \bullet , para indicar onde o módulo tem rank 1, e os círculos vazios, \circ , indicam onde o módulo tem rank 0. As transformações lineares que os conectam tem rank máximo.

Esse exemplo tem o intuito de também estabelecer a notação de como indicaremos os módulos através de seus ranks enquanto fizermos uso dessa representação.

A fim de ilustrar, consideremos $\mathbb{I}[a, b] = \bullet_a \text{ --- } \bullet_b \text{ --- } \circ_c$. Sabemos que

$$I_t = \begin{cases} \mathbb{K}, & \text{se } t \in [a, b] \\ 0, & \text{caso contrário} \end{cases}$$

e as transformações lineares

$$i_s^t = \begin{cases} \mathbf{1}, & \text{se } s, t \in [a, b] \\ 0, & \text{caso contrário} \end{cases}.$$

Portanto, $I_a = I_b = \mathbb{K}$ por estarem sendo simbolizados por \bullet , que significa que os espaços vetoriais tem rank 1, enquanto $I_c = 0$ já que a indicação \circ nos diz que o espaço vetorial tem rank 0.

Temos ainda que $i_a^b = \mathbf{1}$ a função identidade e $i_b^c = 0$ a função nula. Assim, $i_a^c = i_b^c i_a^b = 0$, uma vez que $i_b^c = 0$.

Vemos que através da nova notação, somos capazes de extrair todas as informações relevantes do módulo.

Seja \mathbb{V} um módulo de persistência indexado em \mathbb{R} . Para qualquer conjunto finito de índices $\mathbf{T} : a_1 < a_2 < \dots < a_n$ e qualquer intervalo $[a_i, a_j] \subset \mathbf{T}$, definimos a multiplicidade de $[a_i, a_j]$ em $\mathbb{V}_{\mathbf{T}}$ como sendo o número de cópias de $\mathbb{I}[a_i, a_j]$ que ocorre na decomposição de intervalos de $\mathbb{V}_{\mathbf{T}}$, que assume valores no conjunto $\{0, 1, 2, \dots, +\infty\}$.

Escrevemos $\langle [b, c], \mathbb{V}_{a,b,c} \rangle$, ou ainda $\langle \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c \mid \mathbb{V} \rangle$, para representar a multiplicidade de $\circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c$ no módulo de três termos

$$\mathbb{V}_{a,b,c} = (V_a \xrightarrow{i_a^b} V_b \xrightarrow{i_b^c} V_c).$$

No geral $\langle [b, c] \mid \mathbb{V}_{b,c} \rangle$ não é necessariamente o mesmo que $\langle [b, c] \mid \mathbb{V}_{a,b,c} \rangle$. Quando \mathbb{V} estiver claro segundo o contexto, o omitiremos e vamos escrever:

$$\langle \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c \rangle.$$

Proposição 2.2. Somas Diretas

Suponha $\mathbb{V}_{\mathbf{T}}$ um módulo de persistência que possa ser escrito como soma direta

$$\mathbb{V}_{\mathbf{T}} = \bigoplus_{l \in L} \mathbb{V}_{\mathbf{T}}^l.$$

Então,

$$\langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}} \rangle = \sum_{l \in L} \langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle,$$

para quaisquer conjunto de índices $\mathbf{T} = \{a_1, a_2, \dots, a_n\}$ e intervalo $[a_i, a_j] \subset \mathbf{T}$.

Demonstração. É importante ressaltar que o conjunto de índices \mathbf{T} é finito, no entanto o conjunto onde os módulos de persistência estão indexados, L , não necessariamente é finito. Para cada $l \in L$, $\mathbb{V}_{\mathbf{T}}^l$ pode ser escrito como soma de módulos de intervalos,

por satisfazer a condição (i) do Teorema 2.2 de Gabriel, Auslander, Ringel-Tachikawa, Webb. A fim de simplificar, já tomaremos $\mathbb{V}_{\mathbf{T}}^l$ como a decomposição de módulos de intervalo, $\mathbb{V}_{\mathbf{T}}^l = \bigoplus_{i \in L_l} \mathbb{I}^l[p_i, q_i]$.

Desde que $\mathbb{V} = \bigoplus_{l \in L} \mathbb{V}^l$, pela definição de soma direta, para cada $t \in \mathbf{T}$ temos

$$V_t = \bigoplus_{l \in L} V_t^l.$$

Calcular

$$\langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}} \rangle = \langle \circ_{a_{i-1}} \text{---} \bullet_{a_i} \text{---} \bullet_{a_j} \text{---} \circ_{a_{j+1}} \mid \mathbb{V}_{\mathbf{T}} \rangle$$

é determinar a multiplicidade de $\circ_{a_{i-1}} \text{---} \bullet_{a_i} \text{---} \bullet_{a_j} \text{---} \circ_{a_{j+1}}$ no módulo de n termos

$$V_{a_1} \rightarrow \cdots \rightarrow V_{a_{i-1}} \rightarrow V_{a_i} \rightarrow \cdots \rightarrow V_{a_j} \rightarrow V_{a_{j+1}} \rightarrow \cdots \rightarrow V_{a_n}.$$

Como $V_t = \bigoplus_{l \in L} V_t^l$, a sequência de espaços vetoriais e transformações lineares fica

$$\bigoplus_{l \in L} V_{a_1}^l \rightarrow \cdots \rightarrow \bigoplus_{l \in L} V_{a_{i-1}}^l \rightarrow \bigoplus_{l \in L} V_{a_i}^l \rightarrow \cdots \rightarrow \bigoplus_{l \in L} V_{a_j}^l \rightarrow \bigoplus_{l \in L} V_{a_{j+1}}^l \rightarrow \cdots \rightarrow \bigoplus_{l \in L} V_{a_n}^l.$$

Para cada $l \in L$, temos um módulo de persistência específico

$$\mathbb{V}^l = V_{a_1}^l \rightarrow \cdots \rightarrow V_{a_{i-1}}^l \rightarrow V_{a_i}^l \rightarrow \cdots \rightarrow V_{a_j}^l \rightarrow V_{a_{j+1}}^l \rightarrow \cdots \rightarrow V_{a_n}^l,$$

e calculamos $\langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle = \langle [a_i, a_j] \mid \bigoplus_{i \in L_l} \mathbb{I}^l[p_i, q_i] \rangle$. Se percorrermos todos o elementos do conjunto L , obtemos

$$\sum_{l \in L} \langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle = \sum_{l \in L} \langle [a_i, a_j] \mid \bigoplus_{i \in L_l} \mathbb{I}^l[p_i, q_i] \rangle.$$

Como já mencionado, os módulos de intervalo desempenham um papel tanto sintético quanto analítico. Desconstruímos \mathbb{V} por meio da soma direta de módulos de persistência, que por sua vez são reescritos a nível de seus módulos de intervalo onde analisamos a incidência de dado intervalo $[a_i, a_j]$. Agora, vamos reconstruir \mathbb{V} através da soma direta de cada peça fundamental \mathbb{V}^l , para todo $l \in L$, e desde que estamos lidando com módulos de intervalo, temos que $\bigoplus_{l \in L} \mathbb{V}^l = \mathbb{V}$, garantia do Teorema 2.1 de Krull-Remak-Schmidt-Azumaya.

Então, $\mathbb{V} = \bigoplus_{l \in L} \mathbb{V}^l = \bigoplus_{l \in L} \bigoplus_{i \in L_l} \mathbb{I}^l[p_i, q_i]$.

Assim, $\langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle$, para cada $l \in L$, e tomando a soma sobre cada $l \in L$ obtemos

$$\sum_{l \in L} \langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle.$$

Uma vez que a multiplicidade de $[a_i, a_j]$ é definida como o número de cópias de $\mathbb{I}[a_i, a_j]$ que ocorre na decomposição de $\mathbb{V}_{\mathbf{T}}$, temos que

$$\langle [a_i, a_j], \mathbb{V}_{\mathbf{T}} \rangle = \sum_{l \in L} \langle [a_i, a_j] \mid \mathbb{V}_{\mathbf{T}}^l \rangle.$$

□

Proposição 2.3. *Invariantes por isomorfismos de uma transformação linear $V_a \xrightarrow{v} V_b$ são*

1. $\text{rank}(v) = \langle \bullet_a \text{---} \bullet_b \mid \mathbb{V} \rangle.$
2. $\text{nulidade}(v) = \langle \bullet_a \text{---} \circ_b \mid \mathbb{V} \rangle.$
3. $\text{conulidade}(v) = \langle \circ_a \text{---} \bullet_b \mid \mathbb{V} \rangle.$

O resultado enunciado a seguir está presente em [2] e [3] e nos permite calcular com maior eficiência o rank dos módulos de intervalo.

Proposição 2.4. Princípio da Restrição

Sejam \mathbf{T} , \mathbf{S} conjuntos finitos de índices com $\mathbf{S} \subset \mathbf{T}$. Então,

$$\langle \mathbb{I} \mid \mathbb{V}_{\mathbf{S}} \rangle = \sum_{\mathbb{J}} \langle \mathbb{J} \mid \mathbb{V}_{\mathbf{T}} \rangle$$

onde a soma é sobre os intervalos $\mathbb{J} \subset \mathbf{T}$, tais que $\mathbb{J} \cap \mathbf{S} = \mathbb{I}$.

Exemplo 2.7. Seja $a < b < c$. Considere os novos índices p e q rearranjados de forma que $a < p < b < q < c$. Denominaremos $S = \{a, b, c\}$ e $T_1 = \{a, b, q, c\}$. Note que $S \subset T_1$. Seja

$$[b, c] = \mathbb{I} = \circ_a \text{---} \bullet_b \text{---} \bullet_c.$$

Calcular a multiplicidade do intervalo $[b, c]$ no módulo de persistência \mathbb{V} indexado sobre o conjunto S ,

$$\langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \mid \mathbb{V}_S \rangle,$$

segundo o Princípio da Restrição é o mesmo que computar a multiplicidade do intervalo

$$[b, c] = \mathbb{J} = \circ_a \text{---} \bullet_b \text{---} \bullet_q \text{---} \bullet_c$$

sobre o módulo \mathbb{V} mas agora indexado sobre o conjunto T_1 . Observamos que, de fato, $\mathbb{J} \subset T_1$, porque $\{b, q, c\} \subset \{a, b, q, c\}$. E, ainda, $\mathbb{J} \cap S = \mathbb{I}$ já que se trata de $\{b, q, c\} \cap \{a, b, c\} = \{b, c\}$. Note também que \mathbb{J} é a única opção de intervalo para esses conjuntos de índices. Logo,

$$\begin{aligned} \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \mid \mathbb{V}_S \rangle &= \sum_{\mathbb{J}} \langle \mathbb{J}, \mathbb{V}_{T_1} \rangle \\ &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_q \text{---} \bullet_c \mid \mathbb{V}_{T_1} \rangle. \end{aligned}$$

Agora, no cenário que tivermos outro conjunto de índices,

$$T_2 = \{a, p, b, c\},$$

e quisermos fazer o cálculo da multiplicidade para o mesmo intervalo $\mathbb{I} = [b, c]$ sobre o módulo de persistência \mathbb{V}_S , temos de considerar o seguinte: sejam

$$\mathbb{J}_1 = [b, c] = \circ_a \text{---} \circ_p \text{---} \bullet_b \text{---} \bullet_c$$

e

$$\mathbb{J}_2 = [b, c] = \circ_a \text{---} \bullet_p \text{---} \bullet_b \text{---} \bullet_c.$$

Verifiquemos que \mathbb{J}_1 e \mathbb{J}_2 satisfazem as exigências do Princípio da Restrição: $\mathbb{J}_1 \subset T_2$ desde que $\{b, c\} \subset \{a, p, b, c\}$, bem como $\mathbb{J}_2 \subset T_2$ pois $\{p, b, c\} \subset \{a, p, b, c\}$. Além disso, $\mathbb{J}_1 \cap S = \{b, c\} \cap \{a, b, c\} = \{b, c\} = \mathbb{I}$ e, similarmemente, $\mathbb{J}_2 \cap S = \{p, b, c\} \cap \{a, b, c\} = \{b, c\} = \mathbb{I}$. Portanto, \mathbb{J}_1 e \mathbb{J}_2 satisfazem as condições requeridas pelo Princípio da Restrição. Perceba também que eles são os únicos intervalos existentes, segundo tais conjuntos de índices. Logo,

$$\begin{aligned} \langle \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c \mid \mathbb{V}_S \rangle &= \sum_{i=1}^2 \langle \mathbb{J}_i, \mathbb{V}_{T_2} \rangle \\ &= \langle \circ_a \text{ --- } \circ_p \text{ --- } \bullet_b \text{ --- } \bullet_c, \mathbb{V}_{T_2} \rangle \\ &\quad + \langle \circ_a \text{ --- } \bullet_p \text{ --- } \bullet_b \text{ --- } \bullet_c, \mathbb{V}_{T_2} \rangle . \end{aligned}$$

Existe uma variação na resolução dos casos vistos no exemplo anterior. Isso se dá por conta da localização do índice inserido. Em suma, sempre que tivermos uma configuração em que um novo índice é colocado entre dois módulos iguais, teremos que ele também será igual aos demais. Em outras palavras, para

$$\langle \bullet_a \text{ --- } \bullet_b \rangle ,$$

ao adicionarmos um elemento p , $a < p < b$, necessariamente teremos

$$\langle \bullet_a \text{ --- } \bullet_p \text{ --- } \bullet_b \rangle .$$

Basta entendermos que cada ponto, seja preenchido (\bullet) ou vazio (\circ), representa um espaço vetorial de rank 1 ou rank 0, respectivamente. E as transformações lineares entre eles satisfazem a “lei de decomposição” possuindo rank máximo. Dessa maneira, se possuíssimos um caso em que fosse permitido

$$\langle \bullet_a \text{ --- } \circ_p \text{ --- } \bullet_b \rangle ,$$

aconteceria que $i_a^b : \mathbb{K} \rightarrow \mathbb{K}$ teria rank máximo igual a 1. Por outro lado, ao considerar as transformações lineares $i_a^p : \mathbb{K} \rightarrow 0$, tal que $k \mapsto 0$, para todo $k \in \mathbb{K}$ e $i_p^b : 0 \rightarrow \mathbb{K}$ que naturalmente associa o 0 a $0 \in \mathbb{K}$, teríamos $i_a^b = i_p^b \circ i_a^p$ com rank 0, uma contradição.

No caso em que temos $\langle \circ_a \text{ --- } \bullet_b \rangle$, ao inserirmos um elemento q , $a < q < b$, admitimos ter tanto \circ_p quanto \bullet_p , pois ambos os casos caracterizam um intervalo e ainda, à luz da teoria de persistência, isso significa uma “característica”, ou seja, um instante em que saímos dos espaços vetoriais nulos e começamos nos espaços vetoriais de rank 1.

Proposição 2.5. $rank(V_b \rightarrow V_c) \geq rank(V_a \rightarrow V_d)$, quando $a \leq b \leq c \leq d$.

Demonstração. A Proposição 2.5 é uma afirmação bem conhecida. Queremos aqui demonstrá-la fazendo uso dos cálculos de quiver e tendo em vista o Princípio da Restrição. Assim,

$$rank(V_b \rightarrow V_c) = \langle \text{ --- } \bullet_b \text{ --- } \bullet_c \text{ --- } \rangle .$$

Pelo Princípio da Restrição, temos

$$\begin{aligned}
\langle \text{---} \bullet_b \text{---} \bullet_c \text{---} \rangle &= \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle \\
&+ \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle \\
&+ \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\
&+ \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\
&\geq \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle \\
&= \langle \bullet_a \text{---} \bullet_d \rangle \\
&= \text{rank}(V_a \rightarrow V_d).
\end{aligned}$$

□

2.8 Medidas no Retângulo

Para um \mathbf{T} -módulo de persistência decomponível,

$$\mathbb{V} \cong \bigoplus_{l \in L} \mathbb{I}(p_l^*, q_l^*),$$

já definimos o diagrama de persistência decorado como sendo o multiconjunto

$$Dgm(\mathbb{V}) = \{(p_l^*, q_l^*) \mid l \in L\}$$

e o diagrama de persistência não-decorado como sendo o multiconjunto

$$dgm(\mathbb{V}) = \{(p_l, q_l) \mid l \in L\}.$$

No entanto, se não tivermos que \mathbb{V} é decomponível, precisaremos proceder de forma diferente. A ideia, a grosso modo, é que se soubermos quantos pontos de Dgm estão contidos em cada retângulo contido no semi-plano superior \mathcal{H} , então conheceremos Dgm completamente. Acontece que, para módulos de persistência, contar pontos em retângulos é fácil. A linguagem da teoria da medida é bem adequada aqui. Mostraremos que cada módulo de persistência define uma medida de valor inteiro no retângulo. Se o módulo for tame, então a medida é de valor finito. A medida de persistência que construiremos não é uma medida no sentido próprio da palavra - ela será aditiva no sentido de revestimento ao invés do sentido usual de união disjunta. As discrepâncias aparecerão quando dividirmos retângulos em dois triângulos: o que acontecerá com os pontos nas arestas comuns? A qual triângulo eles pertencem? A fim de resolver isso, seremos naturalmente conduzidos a noção de pontos decorados. Que serve perfeitamente com nossa utilização de decorações para distinguir extremos de intervalos abertos e fechados.

2.8.1 A Medida de Persistência

Seja \mathbb{V} um módulo de persistência. A medida de persistência de \mathbb{V} é a função

$$\mu_{\mathbb{V}}(R) = \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle$$

definida no retângulo $R = [a, b] \times [c, d]$ no plano, com $a < b \leq c < d$.

Para módulos de persistência decomponíveis, existe uma clara relação entre $\mu_{\mathbb{V}}$ e as somas de intervalos de \mathbb{V} .

Vamos considerar primeiro o caso que temos um módulo de intervalo.

Proposição 2.6. *Seja $\mathbb{V} = \mathbb{I}^J$, onde $J = (p^*, q^*)$ é um intervalo na reta real. Seja $R = [a, b] \times [c, d]$ com $a < b \leq c < d$. Então,*

$$\mu_{\mathbb{V}}(R) = \begin{cases} 1, & \text{se } [b, c] \subseteq J \subseteq (a, d) \\ 0, & \text{caso contrário} \end{cases} .$$

Demonstração. Seja $R = [a, b] \times [c, d]$ um retângulo com $a < b \leq c < d$. O módulo $\mathbb{I}_{\{a,b,c,d\}}^J$, por definição, é a família de espaços vetoriais a 1-parâmetro $(I_t \mid t \in \{a, b, c, d\})$, com $I_t = \begin{cases} \mathbb{K}, & \text{se } t \in J \subseteq (a, d) \\ 0, & \text{caso contrário} \end{cases} .$

Precisamos analisar algumas condições sobre o conjunto J . Se $J = (p^*, q^*)$ estiver contido em (a, d) e ainda possuir $[b, c]$ como subconjunto ou se o intervalo $J = (p^*, q^*)$ contiver (a, d) , naturalmente $[b, c]$ também; Então J restrito ao conjunto (a, d) - pois exigimos $J \subseteq (a, d)$ quando estruturamos a teoria de módulos de intervalo - devemos ter J um intervalo segundo o conjunto $\{a, b, c, d\}$ e uma vez que temos a condição $a < b \leq c < d$ segue que o primeiro elemento que pode ser alcançado por J seria b e, por conseguinte, c o último, porque os extremos são abertos.

Se J for um intervalo degenerado, por exemplo,

$$\langle \circ_a \text{ --- } \circ_b \text{ --- } \bullet_c \text{ --- } \circ_d \rangle ,$$

temos que o rank da transformação linear $v_b^c : \mathbb{K} \rightarrow 0$ é zero. Senão, J é um intervalo não-degenerado, e a única opção seria

$$\langle \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c \text{ --- } \circ_d \rangle .$$

Nesse caso, o rank da transformação linear $v_b^c : \mathbb{R} \rightarrow \mathbb{R}$ é 1.

Assim, se $[b, c] \subseteq J \subseteq (a, d)$, teremos que a multiplicidade do intervalo $[b, c]$ segundo o retângulo $R = [a, b] \times [c, d]$ é 1. Caso contrário, teremos rank 0. Portanto,

$$\mu_{\mathbb{V}}(R) = \begin{cases} 1, & \text{se } [b, c] \subseteq J \subseteq (a, d) \\ 0, & \text{caso contrário} \end{cases} .$$

□

A proposição acima tem uma interpretação gráfica de que representa o intervalo $J \subseteq \mathbb{R}$ como um ponto decorado no plano estendido. A imagem abaixo indica exatamente quais pontos decorados (p^*, q^*) são detectados por $\mu_{\mathbb{V}}(R)$.

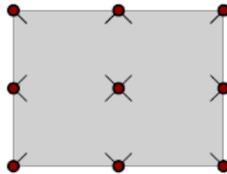


Figura 2.4: Pontos decorados detectados pela medida.

Se (p, q) estiver contido no interior de R , então (p^*, q^*) sempre será detectado, independente de sua decoração. Contudo, se (p, q) estiver na fronteira, então (p^*, q^*) será detectado desde que seu tracinho esteja indicando a direção ao interior de R .

Definição 2.8. *Seja $R = [a, b] \times [c, d]$, com $a < b \leq c < d$, e considere o ponto decorado (p^*, q^*) , com $p^* < q^*$. Diremos que $(p^*, q^*) \in R$ se*

1. *o intervalo $J = (p^*, q^*)$ satisfizer $[b, c] \subseteq J \subset (a, d)$;
ou equivalentemente*
2. *o ponto (p, q) juntamente com sua decoração estiverem contidos no retângulo fechado R .*

Usamos a notação

$$R^\blacksquare = \{(p^*, q^*) \in R\}$$

quando queremos explicitar o conjunto de pontos decorados contidos em R . Entretanto, preferimos escrever $(p^*, q^*) \in R$, ao invés de $(p^*, q^*) \in R^\blacksquare$ quando simplesmente indicamos a relação entre um ponto decorado e um retângulo.

Corolário 2.1. *Suponha \mathbb{V} um módulo de persistência decomponível sobre \mathbb{R} ,*

$$\mathbb{V} = \bigoplus_{l \in L} \mathbb{I}(p_l^*, q_l^*).$$

Então,

$$\mu_{\mathbb{V}}(R) = \text{card}(Dgm(\mathbb{V})|_R).$$

Demonstração. Recordando a definição, temos que se $\mathcal{A} = (S, m)$ é um multiconjunto, então $\text{card}(\mathcal{A}) = \sum_{s \in S} m(s)$. Sabemos que $Dgm(\mathbb{V})$ é um exemplo de multiconjunto.

Seja $R = [a, b] \times [c, d]$, com $a < b \leq c < d$, um retângulo. Logo,

$$\begin{aligned} \langle \circ_a \text{ --- } \bullet_b \text{ --- } \bullet_c \text{ --- } \circ_d \mid \mathbb{V} \rangle &= \langle [b, c] \mid \mathbb{V} \rangle \\ &= \sum_{l \in L} \langle [b, c] \mid \mathbb{I}(p_l^*, q_l^*) \rangle \end{aligned}$$

pois \mathbb{V} é decomponível em módulos de intervalo e, pela Proposição 2.2, temos a igualdade acima.

Além disso,

$$\text{card}(Dgm(\mathbb{V})|_R) = \sum_{l \in L} \langle [p_l^*, q_l^*] \mid \mathbb{V}|_R \rangle,$$

ou seja, equivale a considerar $[p_l^*, q_l^*]$ que estão contidos em R , portanto contamos os elementos da forma $p_l^* = b$ e $q_l^* = c$, para todo $l \in L$, pois os elementos $[p_l^*, q_l^*]$ do módulo de intervalo contido no retângulo $R = [a, b] \times [c, d]$ são necessariamente da forma $[b, c]$.

Assim,

$$\begin{aligned} \text{card}(Dgm(\mathbb{V})|_R) &= \sum_{l \in L} \langle [p_l^*, q_l^*] \mid \mathbb{V}|_R \rangle \\ &= \sum_{l \in L} \langle [b, c] \mid \mathbb{V}|_R \rangle \\ &= \sum_{l \in L} \langle [b, c] \mid \mathbb{V} \rangle \\ &= \mu_{\mathbb{V}}(R). \end{aligned}$$

Portanto,

$$\mu_{\mathbb{V}}(R) = \text{card}(Dgm\mathbb{V}|_R).$$

□

Agora, iremos articular nossa estratégia a fim de definir diagrama de persistência sem assumir que o módulo seja decomponível. É a seguinte:

- Construir a medida de persistência $\mu_{\mathbb{V}}$;
- Definir $Dgm(\mathbb{V})$ como sendo o multiconjunto no semi-plano superior tal que $\mu_{\mathbb{V}}(R) = \text{card}(Dgm(\mathbb{V}) \mid_R)$ seja uma igualdade válida para todos os retângulos R .

Para a realização desse trabalho, necessitamos saber se tal multiconjunto existe e, ainda, se é único.

2.8.2 Continuação da Medida de Persistência

Chamamos $\mu_{\mathbb{V}}$ de medida porque ela é aditiva com respeito a divisão de um triângulo em outros dois triângulos. Comprovaremos em breve esse fato. Agora, vamos dar uma prova de uma fórmula de soma alternada para $\mu_{\mathbb{V}}(R)$. Definimos $r_s^t = \text{rank}(v_s^t : V_s \rightarrow V_t)$.

Proposição 2.7. *Seja \mathbb{V} um módulo de persistência e sejam $a < b \leq c < d$. Se os espaços V_a, V_b, V_c e V_d são de dimensão finita, ou menos estritamente, se $r_b^c < \infty$, então*

$$\langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle = r_b^c - r_a^c - r_b^d + r_a^d.$$

Demonstração. Consideremos o conjunto $\mathbf{T} = \{a, b, c, d\}$. Vamos descrever o módulo de persistência \mathbb{V} sobre \mathbf{T} gerando o módulo de quatro termos. Note que, segundo o Proposição 2.5, temos $\text{rank}(V_a \rightarrow V_d) \leq \text{rank}(V_b \rightarrow V_c) = r_b^c < \infty$, implicando que todos os ranks são finitos. Assim, somos aptos a tomar e somar todos os ranks.

Observemos que, por um lado,

$$\begin{aligned} r_b^c &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ r_a^c &= \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ r_b^d &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle \\ r_a^d &= \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle, \end{aligned}$$

que segue da definição.

Por outro lado, o Princípio da Restrição nos permite calcular

1.

$$\begin{aligned} \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &+ \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &+ \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle \\ &+ \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle. \end{aligned}$$

2.

$$\langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle = \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle + \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle .$$

3.

$$\langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle = \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle + \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle .$$

4.

$$\langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle = \langle \bullet_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_d \rangle .$$

Reunindo as informações, conseguimos

$$r_b^c - r_a^c - r_b^d + r_a^d = \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle .$$

□

Proposição 2.8. μ é aditiva com respeito a divisão horizontal e vertical, isto é,

$$\mu_{\mathbb{V}}([a, b] \times [c, d]) = \mu_{\mathbb{V}}([a, p] \times [c, d]) + \mu_{\mathbb{V}}([p, b] \times [c, d])$$

e

$$\mu_{\mathbb{V}}([a, b] \times [c, d]) = \mu_{\mathbb{V}}([a, b] \times [c, q]) + \mu_{\mathbb{V}}([a, b] \times [q, d])$$

sempre que $a < p < b \leq c < q < d$.

A propriedade de aditividade pode ser ilustrada pela imagem

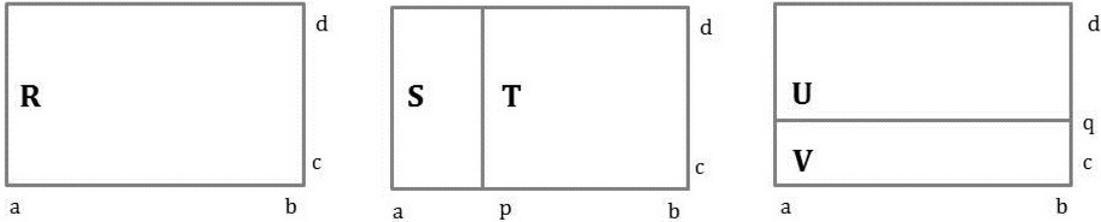


Figura 2.5: Ilustração da propriedade de aditividade.

Nossa afirmação é que $\mu_{\mathbb{V}}(R) = \mu_{\mathbb{V}}(S) + \mu_{\mathbb{V}}(T) = \mu_{\mathbb{V}}(U) + \mu_{\mathbb{V}}(V)$

Demonstração. Apresentaremos duas demonstrações para o enunciado acima. Cada uma traz consigo aspectos distintos.

Seja $a < p < b \leq c < d$. Usando o cálculo de quiver,

$$\begin{aligned} \mu_{\mathbb{V}}([a, b] \times [c, d]) &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &= \langle \circ_a \text{---} \circ_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &\quad + \langle \circ_a \text{---} \bullet_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &= \mu_{\mathbb{V}}([a, p] \times [c, d]) + \mu_{\mathbb{V}}([p, b] \times [c, d]) \end{aligned}$$

e, assim, provamos a aditividade com respeito a cortes horizontais.

Para os cortes verticais,

$$\begin{aligned} \mu_{\mathbb{V}}([a, b] \times [c, d]) &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_q \text{---} \circ_d \rangle \\ &\quad + \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_q \text{---} \circ_d \rangle \\ &= \mu_{\mathbb{V}}([a, b] \times [c, q]) + \mu_{\mathbb{V}}([a, b] \times [q, d]). \end{aligned}$$

Essa foi a primeira demonstração. A segunda demonstração só é possível assumindo $r_b^c < \infty$, o que, como vimos anteriormente, implica que os demais ranks são também todos finitos. A fórmula da soma alternada nos fornece

$$\langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle = r_b^c - r_a^c - r_b^d + r_a^d$$

através da Proposição 2.7. Agora, somando $0 = (r_p^c - r_p^c) + (r_p^d - r_p^d)$, obtemos

$$\begin{aligned} &r_b^c - r_a^c - r_b^d + r_a^d + (r_p^c - r_p^c) + (r_p^d - r_p^d) \\ &= (r_p^c - r_a^c - r_p^d + r_a^d) + (r_b^c - r_p^c - r_b^d + r_p^d) \\ &= \langle \circ_a \text{---} \bullet_p \text{---} \bullet_c \text{---} \circ_d \rangle + \langle \circ_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle \\ &= \mu_{\mathbb{V}}([a, p] \times [c, d]) + \mu_{\mathbb{V}}([p, d] \times [c, d]). \end{aligned}$$

Procedemos analogamente para concluir que

$$\begin{aligned} r_b^c - r_a^c - r_b^d + r_a^d &= (r_b^c - r_a^c - r_b^q + r_a^q) + (r_b^q - r_a^q - r_b^d + r_a^d) \\ &= \mu_{\mathbb{V}}([a, b] \times [c, q]) + \mu_{\mathbb{V}}([a, b] \times [q, d]). \end{aligned}$$

□

A segunda prova é particularmente transparente quando desenhada geometricamente no plano: os sinais + e - nos vértices dos retângulos se cancelam de uma forma desejável.

Proposição 2.9.

$$\begin{aligned} \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \rangle &= \dim \left(\frac{\text{Im}(v_b^c) \cap \ker(v_c^d)}{\text{Im}(v_a^c) \cap \ker(v_c^d)} \right) \\ &= \dim \left(\frac{\ker(v_b^d)}{\text{Im}(v_a^c) + \text{Im}(v_b^c) \cap \ker(v_b^d)} \right). \end{aligned}$$

A proposição acima expressa a medida do retângulo como a dimensão de um espaço vetorial construído de \mathbb{V} . Essa abordagem pode ser útil, no entanto, existem propriedades, tais como a aditividade, que não são óbvias de serem checadas nesses termos.

2.9 r -medidas abstratas

Consideraremos a medida de retângulos de forma mais abstrata. A medida de persistência é, de fato, nosso exemplo primário, mas a formulação mais geral é útil em diversas outras situações. Para fins de exposição, inicialmente trabalharemos no plano \mathbb{R}^2 ao invés do plano estendido $\overline{\mathbb{R}^2}$.

Definição 2.9. *Seja D subconjunto de \mathbb{R}^2 . Defina*

$$\text{Rect}(D) = \{[a, b] \times [c, d] \subset D \mid a < b \text{ e } c < d\}$$

o conjunto dos retângulos fechados contidos em D . A medida do retângulo, também chamada de r -medida, em D é uma função

$$\mu : \text{Rect}(D) \rightarrow \{0, 1, 2, \dots\} \cup \{+\infty\}$$

que é aditiva segundo as divisões horizontal e vertical.

Proposição 2.10. *Seja μ uma r -medida em $D \subset \mathbb{R}^2$. Então, μ é*

- **Finitamente aditiva:** *Se $R \in \text{Rect}(D)$ pode ser escrito como união*

$$R = \bigcup_{i=1}^k R_i,$$

com R_i retângulos com interiores disjuntos, então

$$\mu(R) = \sum_{i=1}^k \mu(R_i).$$

- **Monótona:** *Se $R \subseteq S$, então $\mu(R) \leq \mu(S)$.*

Demonstração. Seja $R = [a, b] \times [c, d]$ em $\text{Rect}(D)$ e suponha

$$R = \bigcup_{i=1}^k R_i,$$

R_i 's retângulos com interiores disjuntos. Suponha, primeiramente, os R_i 's da forma $R_i = [a_i, a_{i+1}] \times [c, d]$, com $a_i < a_{i+1}$ retângulo em D e $a = a_1 < a_2 < \dots < a_k = b$. Como μ é uma r -medida, seque que μ é aditiva segundo divisão vertical. Assim,

$$\begin{aligned} \mu(R) &= \mu([a, b] \times [c, d]) \\ &= \mu([a_1, a_2] \times [c, d]) + \mu([a_2, a_3] \times [c, d]) + \dots + \mu([a_{k-1}, a_k] \times [c, d]). \end{aligned}$$

Agora, sendo μ uma r -medida, é também aditiva segundo divisão horizontal. Seja $c = c_1 < c_2 < \dots < c_l = d$. Então,

$$\begin{aligned} \mu(R) &= \mu([a, b] \times [c, d]) \\ &= \mu([a, b] \times [c_1, c_2]) + \mu([a, b] \times [c_2, c_3]) + \dots + \mu([a, b] \times [c_{l-1}, c_l]). \end{aligned}$$

No caso geral,

$$R = \bigcup_{1 \leq i \leq k, 1 \leq j \leq l} R_{ij}$$

fazemos:

$$\begin{aligned} \mu(R) &= \mu([a, b] \times [c, d]) \\ &= \sum_{i=1}^{k-1} \mu([a_i, a_{i+1}] \times [c, d]) \\ &= \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} \mu([a_i, a_{i+1}] \times [c_j, c_{j+1}]) \\ &= \sum_{i,j} \mu(R_{ij}). \end{aligned}$$

Para demonstrar que μ é monótona, vamos decompor S em uma coleção de retângulos de interiores disjuntos constituídos do próprio R e R_1, \dots, R_{k-1} . Como $R \subseteq S$, com R e $S \in \text{Rect}(D)$, então $R = [r_a, r_b] \times [r_c, r_d]$ com $r_a < r_b$ e $r_c < r_d$, bem como $S = [s_a, s_b] \times [s_c, s_d]$, onde $s_a < s_b$ e $s_c < s_d$, e $[r_a, r_b] \times [r_c, r_d] \subseteq [s_a, s_b] \times [s_c, s_d]$ com $s_a \leq r_a < r_b \leq s_b$ e $s_c \leq r_c < r_d \leq s_d$.

Escreva

$$S = R \cup \left(\bigcup_{1 \leq i < k; 1 \leq j < l} R_{ij} \right),$$

de tal forma que $R \cap R_{ij} = \emptyset$ para quaisquer i e j , e também $R_{ij} \cap R_{pq} = \emptyset$ sempre que $i \neq p$ e $j \neq q$. Portanto,

$$\mu(S) = \mu\left(R \cup \left(\bigcup_{1 \leq i < k; 1 \leq j < l} R_{ij} \right)\right) = \mu(R) + \sum_{ij} \mu(R_{ij}) \geq \mu(R)$$

uma vez que μ assume valores no conjunto $\{0, 1, \dots\} \cup \{+\infty\}$.

Afirmamos que para tal decomposição é necessário, no máximo, 5 retângulos.

Caso 1: O “pior” caso ocorre quando $s_a < r_a < r_b < s_b$ e $s_c < r_c < r_d < s_d$.

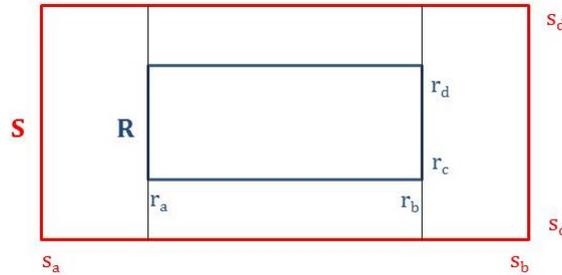


Figura 2.6: Decomposição do retângulo S em outros cinco retângulos, incluindo R como um deles.

Assim, teremos os retângulos

$$R_1 = [s_a, r_a] \times [s_c, s_d],$$

$$\begin{aligned}
 R_2 &= [r_a, r_b] \times [s_c, r_c], \\
 R_3 &= [r_b, s_b] \times [s_c, s_d], \\
 R_4 &= [r_a, r_b] \times [r_d, s_d], \\
 R_5 &= R = [r_a, r_b] \times [r_c, r_d].
 \end{aligned}$$

Caso 2:

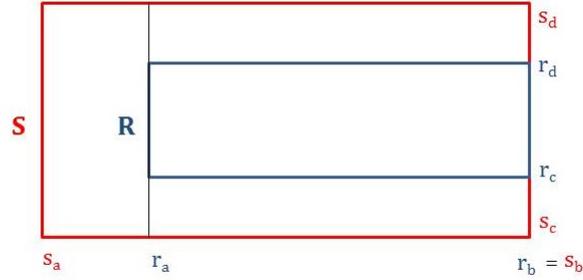


Figura 2.7: Decomposição do retângulo S em outros quatro retângulos, incluindo R como um deles.

Se $s_a < r_a < r_b = s_b$ e $s_c < r_c < r_d < s_d$, conseguimos reescrever segundo os seguintes retângulos

$$\begin{aligned}
 R_1 &= [s_a, r_a] \times [s_c, s_d], \\
 R_2 &= [r_a, s_b] \times [s_c, r_c], \\
 R_3 &= [r_a, s_b] \times [r_d, s_d], \\
 R_4 &= R = [r_a, s_b] \times [r_d, s_d].
 \end{aligned}$$

Caso 3:



Figura 2.8: Decomposição do retângulo S em outros três retângulos, incluindo R como um deles.

Se $s_a = r_a < s_b = r_b$ e $s_c < r_c < r_d < s_d$, podemos decompor nos seguintes retângulos

$$\begin{aligned}
 R_1 &= [s_a, s_b] \times [s_c, r_c], \\
 R_2 &= [s_a, s_b] \times [r_d, s_d], \\
 R_3 &= R = [s_a, s_b] \times [r_c, r_d].
 \end{aligned}$$

Caso 4:

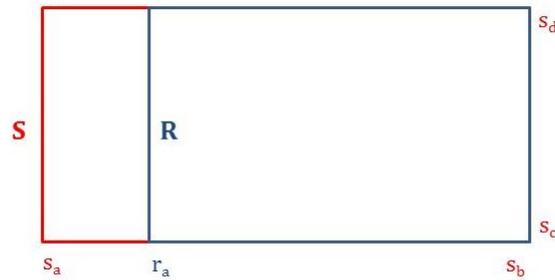


Figura 2.9: Decomposição do retângulo S em outros dois retângulos, incluindo R como um deles.

Se $s_a < r_a < r_b = s_b$ e $s_c = r_c < r_d = s_d$, obtemos

$$R_1 = [s_a, r_a] \times [s_c, s_d],$$

$$R_2 = R = [r_a, s_b] \times [s_c, s_d].$$

Caso 5: O melhor dos casos ocorre quando temos $S = R$, quando a conclusão é trivial.



Figura 2.10: O retângulo $S = R$ é sua própria decomposição

□

Observação 2.4. Notemos que em cada um dos casos há configurações completamente análogas provenientes do movimento rígido de translação do retângulo R , mas mantendo as hipóteses com relação à quantidade de contato que ocorre no retângulo maior S . Assim, a quantidade requerida de retângulos, para cada caso, se mantém, apenas modificando a forma de decompor.

3 Topologia Algébrica Computacional

3.1 Funções de Morse

A classe de funções à valores reais definidas em uma variedade é muito ampla e diversa, inclusive difícil de se caracterizar, ainda que restringi-las para funções contínuas não surte muito efeito em entender seu comportamento.

Mesmo funções suaves podem ser bastante complicadas em seu comportamento, e é melhor, portanto, adicionarmos outra hipótese, a saber, a genericidade. O que conseguimos então é a classe de funções de Morse, que se distingue ao ter apenas pontos críticos simples. A maior parte da teoria está interessada no estudo dos pontos críticos, sua estrutura e o que eles dizem sobre a variedade e a função. Apesar de raramente encontrarmos funções de Morse em aplicações reais, ou até mesmo funções suaves, para esse assunto, saber sobre a sua estrutura beneficia significativamente a nossa compreensão de funções suaves gerais e até funções lineares por partes.

3.1.1 O Toro em Pé

Iniciaremos nossa discussão com um exemplo que ilustrará perfeitamente como a informação contida em uma variedade de dimensão 2 pode ser extraída dos pontos críticos de uma função definida nela à valores reais.

Sejam \mathbb{M} o Toro de dimensão 2 e $f : \mathbb{M} \rightarrow \mathbb{R}$ a função altura.

Considere $\mathbb{M}_a = f^{-1}(-\infty, a] = \{x \in \mathbb{M} \mid f(x) \leq a\}$ o conjunto de subnível. Estamos interessados na evolução dos subníveis a medida que aumentamos o limitante a .

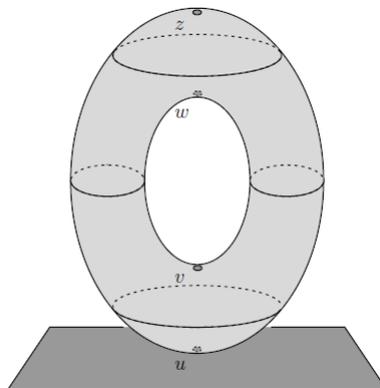


Figura 3.1: O Toro de Pé.

Vamos estudar o comportamento de \mathbb{M}_a conforme variamos a .

Sejam u, v, w e $w \in \mathbb{M}$. Para $a \in \mathbb{R}$, tal que $a < f(u)$, temos $\mathbb{M}_a = \emptyset$.

Quando tomamos $a \in \mathbb{R}$, com $f(u) \leq a < f(v)$, \mathbb{M}_a é homeomorfo ao disco fechado, por conseguinte tem o tipo de homotopia de um ponto.

Já quando $f(v) \leq a < f(w)$, ocorre outra evolução nos conjuntos de subnível e obtemos que \mathbb{M}_a é homeomorfo a um cilindro, portanto, tem o mesmo tipo de homotopia de um círculo.

Quando $f(w) \leq a < f(z)$, o conjunto de subnível \mathbb{M}_a tem o mesmo tipo de homotopia que “a figura 8”.

Finalmente, para $f(z) < a$, \mathbb{M}_a é o toro.

Ressaltamos que o comportamento dos conjuntos de subnível só se modifica quando passamos pelos pontos u, v, w e z , que são os pontos críticos de f - um ponto de mínimo global, duas selas e um ponto de máximo global da função f .

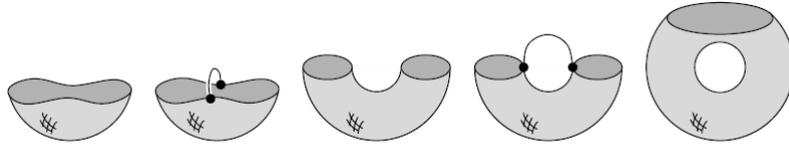


Figura 3.2: Evolução dos conjuntos de subnível do Toro de Pé.

3.2 Funções Suaves

Seja \mathbb{M} uma variedade suave de dimensão d , isto é, \mathbb{M} possui cartas locais difeomorfas a bolas abertas de \mathbb{R}^d . Relembrando que um difeomorfismo é um homeomorfismo que é suave em ambas as direções. Além disso, tecnicamente, ser suave significa que as derivadas de todas as ordens existem, já na prática, apenas precisamos que as derivadas de primeira e segunda ordem existam para a maioria dos resultados que desenvolveremos aqui.

Denotaremos o espaço tangente ao ponto $x \in \mathbb{M}$ por $T_x\mathbb{M}$ que por sua vez é um espaço vetorial de dimensão d consistindo de todos os vetores tangentes a \mathbb{M} em x .

Uma função suave definida entre variedades suaves $f : \mathbb{M} \rightarrow \mathbb{N}$ induz uma transformação linear entre os espaços tangentes, a derivada $Df_x : T_x\mathbb{M} \rightarrow T_{f(x)}\mathbb{N}$, para todo $x \in \mathbb{M}$.

Estamos principalmente interessados em funções a valores reais, ou seja, as quais $\mathbb{N} = \mathbb{R}$, conseqüentemente, obtemos as transformações lineares $Df_x : T_x\mathbb{M} \rightarrow T_{f(x)}\mathbb{R}$. O espaço tangente em um ponto da reta real, \mathbb{R} , é a própria reta real \mathbb{R} .

Definição 3.1. *Seja \mathbb{M} uma variedade suave de dimensão d e considere $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função suave.*

i Diremos que $x \in \mathbb{M}$ é ponto regular de f se Df_x for sobrejetora.

ii Chamaremos $x \in \mathbb{M}$ ponto crítico de f se Df_x for a aplicação nula.

Caso tenhamos um sistema de coordenadas locais (x_1, x_2, \dots, x_d) para uma vizinhança do ponto $x \in \mathbb{M}$, então x será ponto crítico se, e somente se, todas as derivadas parciais de primeira ordem forem nulas, isto é,

$$\frac{\partial f}{\partial x_1}(x) = \frac{\partial f}{\partial x_2}(x) = \dots = \frac{\partial f}{\partial x_d}(x) = 0$$

ou ainda podemos escrever como $\nabla f_x = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_d}(x) \right) = 0$.

Definição 3.2. *A imagem de um ponto regular x de f é denominada valor regular de f . Já a imagem de um ponto crítico x de f é nomeada valor crítico de f .*

Além disso, utilizamos a segunda derivada para sermos capazes de distinguir diferentes tipos de pontos críticos. A matriz Hessiana de f no ponto x é definida por

$$\text{Hess}_x(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_d} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1} & \frac{\partial^2 f}{\partial x_d \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix} \quad (3.1)$$

Definição 3.3. *Um ponto crítico x de f é dito não-degenerado se $\det \text{Hess}_x(f) \neq 0$, ou seja, a Hessiana é não-singular*

3.3 O Lema de Morse

Em um ponto crítico, todas as derivadas de primeira ordem se anulam, assim a Fórmula de Expansão Local de Taylor não tem termos lineares. Se o ponto crítico for não-degenerado, então o comportamento da função em uma vizinhança pequena do ponto crítico será dominado pelos termos quadráticos.

Além disso, somos capazes de encontrar coordenadas locais tais que não hajam termos de ordem superior a 2.

Lema 3.1. Lema de Morse

Sejam \mathbb{M} uma variedade suave de dimensão d e $f : \mathbb{M} \rightarrow \mathbb{R}$ função suave. Se $u \in \mathbb{M}$ é um ponto crítico não-degenerado de f , então existem coordenadas locais, com $u = (0, 0, \dots, 0)$ e

$$f(x) = f(u) - x_1^2 - x_2^2 - \dots - x_q^2 + x_{q+1}^2 + \dots + x_{d-1}^2 + x_d^2,$$

para todo $x = (x_1, \dots, x_d)$ pertencente a essa vizinhança pequena de u .

Demonstração. Seja $\phi : U \rightarrow V$ uma carta local em torno do ponto $u \in \mathbb{M}$ tal que $\phi(0) = u$. Então, $f \circ \phi$ é uma função suave do aberto U de \mathbb{R}^d em \mathbb{R} , com ponto crítico não-degenerado no elemento $0 \in \mathbb{R}^d$. Suponhamos que $f(u) = f \circ \phi(0) = 0$.

Afirmção 1. Seja $U \subset \mathbb{R}^d$ vizinhança aberta de $0 \in \mathbb{R}^d$. Toda função suave $F : U \rightarrow \mathbb{R}$, com $F(0) = 0$, pode ser escrita no forma

$$F(x) = x_1 g_1(x_1, \dots, x_d) + \dots + x_d g_d(x_1, \dots, x_d),$$

em alguma vizinhança da origem, onde g_i são funções suaves, tais que

$$\frac{\partial F}{\partial x_i}(0) = g_i(0).$$

Com efeito, desde que U é aberto, segue que existe uma bola aberta D inteiramente contida nele, com D vizinhança de $0 \in U$. Para todo $x \in D$, pelo Teorema Fundamental do Cálculo,

$$\begin{aligned} F(x) &= \int_0^1 \frac{d}{dt} F(tx) dt, \text{ uma vez que } F(0) = 0 \\ &= \int_0^1 \left(\frac{\partial F}{\partial x_1}(tx)x_1 + \dots + \frac{\partial F}{\partial x_d}(tx)x_d \right) dt \\ &= x_1 g_1(x) + \dots + x_d g_d(x), \text{ onde } g_i(x) = \int_0^1 \frac{\partial F}{\partial x_i}(tx) dt, \text{ para } i = 1, \dots, d. \end{aligned}$$

Temos, assim,

$$g_i(0) = \int_0^1 \frac{\partial F}{\partial x_i}(tx) dt \Big|_{x=0} = \frac{\partial F}{\partial x_i}(0),$$

como desejávamos.

Aplicando a Afirmação 1 a $f \circ \phi$, nos é garantido

$$f\phi(x) = x_1 g_1(x) + \dots + x_d g_d(x)$$

para x em uma vizinhança aberta de 0 , contida em U . Além disso,

$$g_i(0) = \frac{\partial(f \circ \phi)}{\partial x_i}(0) = 0, \text{ já que } 0 \text{ é ponto crítico.}$$

Logo, estamos nas condições de aplicar, novamente, a Afirmação 1, mas agora para cada função g_i , obtendo

$$g_i(x) = x_1 h_{i,1}(x) + x_2 h_{i,2}(x) + \dots + x_d h_{i,d}(x),$$

isto é,

$$g_i(x) = \sum_{k=1}^d x_k h_{i,k}(x)$$

em uma vizinhança aberta de 0 .

Portanto, conseguimos

$$\begin{aligned} f \circ \phi(x) &= \sum_{i,k} x_i x_k h_{i,k}(x), \\ f \circ \phi(x) &= x_1^2 h_{1,1}(x) + x_1 x_2 h_{1,2}(x) + \dots + x_1 x_d h_{1,d}(x) + \\ &+ x_2 x_1 h_{2,1}(x) + x_2^2 h_{2,2}(x) + \dots + x_2 x_d h_{2,d}(x) + \dots \\ &\dots + x_d x_1 h_{d,1}(x) + x_d x_2 h_{d,2}(x) + \dots + x_d^2 h_{d,d}(x). \end{aligned}$$

Em outras palavras,

$$f \circ \phi(x) = x_1^2 h_{1,1}(x) + 2x_1 x_2 (h_{1,2} + h_{2,1})(x) + \dots + 2x_i x_j (h_{i,j} + h_{j,i})(x) + \dots + x_d^2 h_{d,d}(x).$$

Defina

$$H_{i,k} = \frac{(h_{i,k} + h_{k,i})}{2}$$

e perceba que $H_{k,i} = H_{i,k}$. Então,

$$f \circ \phi(x) = \sum_{i,k} x_i x_k H_{k,i}(x),$$

$$f \circ \phi(x) = x_1^2 H_{1,1}(x) + x_1 x_2 H_{1,2}(x) + \dots + x_i x_j H_{i,j}(x) + \dots + x_d^2 H_{d,d}(x).$$

Calcularemos a derivada parcial na direção de x_i ,

$$\frac{\partial(f \circ \phi)}{\partial x_i}(x) = \frac{\partial(x_1^2 H_{1,1} + x_1 x_2 H_{1,2} + \dots + x_i x_j H_{i,j} + \dots + x_d^2 H_{d,d})}{\partial x_i}(x)$$

por conseguinte

$$\frac{\partial(f \circ \phi)}{\partial x_i}(x) = x_1 H_{i,1}(x) + x_i x_1 \frac{\partial H_{i,1}}{\partial x_i}(x) + \dots + 2x_i H_{i,i}(x) + x_i^2 \frac{\partial H_{i,i}}{\partial x_i}(x) + \dots + x_d H_{i,d}(x) + x_i x_d \frac{\partial H_{i,d}}{\partial x_i}(x).$$

Agora, se aplicarmos a segunda derivada na direção i , conseguimos

$$\begin{aligned} \frac{\partial^2(f \circ \phi)}{\partial x_i^2}(x) &= 2x_1 \frac{\partial H_{i,1}}{\partial x_i}(x) + x_i x_1 \frac{\partial^2 H_{i,1}}{\partial x_i^2}(x) + \dots + 2H_{i,i}(x) + 2x_i \frac{\partial H_{i,i}}{\partial x_i}(x) + x_i^2 \frac{\partial^2 H_{i,i}}{\partial x_i^2}(x) + \dots + \\ &\quad \dots + 2x_d \frac{\partial H_{i,d}}{\partial x_i}(x) + x_i x_d \frac{\partial^2 H_{i,d}}{\partial x_i^2}(x) \end{aligned}$$

Avaliando $\frac{\partial^2(f \circ \phi)}{\partial x_i^2}(x)$ em $x = 0$ implica que

$$\frac{\partial^2(f \circ \phi)}{\partial x_i^2}(0) = 2H_{i,i}(0),$$

porque $x = (x_1, x_2, \dots, x_d) = (0, 0, \dots, 0) = 0$.

Analogamente obtemos que

$$\frac{\partial^2(f \circ \phi)}{\partial x_j \partial x_i}(0) = 2H_{i,j}(0).$$

Pela hipótese de que 0 é ponto crítico não-degenerado, segue que $\det(H_{k,i}(0)) \neq 0$, pela continuidade de $h_{i,j}$ temos que existe uma vizinhança aberta V' de 0 do qual $h_{i,j}(x) \neq 0$, para todo $x \in V'$. Assim, em particular, $H_{1,1}(0) \neq 0$.

Seja (y_1, x_2, \dots, x_d) um novo sistemas de coordenadas locais para a vizinhança, donde $y_1 = \sqrt{|H_{1,1}|}(x_1 + \sum_{i=2}^d x_i \frac{H_{1,i}}{H_{1,1}})$.

Assim,

$$y_1^2 = |H_{1,1}|(x_1 + \sum_{i=2}^d x_i \frac{H_{1,i}}{H_{1,1}})^2.$$

$$y_1^2 = H_{1,1}x_1^2 + 2 \sum_{i=2}^d x_1 x_i H_{1,i} + \frac{(\sum_{i=2}^d x_i H_{1,i})^2}{H_{1,1}},$$

caso tivermos $H_{1,1} > 0$, e

$$y_1^2 = -H_{1,1}x_1^2 - 2 \sum_{i=2}^d x_1 x_i H_{1,i} - \frac{(\sum_{i=2}^d x_i H_{1,i})^2}{H_{1,1}},$$

caso tivermos $H_{1,1} < 0$.

Implica que teremos a função

$$f(x) = y_1^2 + \sum_{i,j=2}^d x_i x_j H_{i,j} - \frac{(\sum_{i=2}^d x_i H_{1,i})^2}{H_{1,1}},$$

para $H_{1,1} > 0$.

Senão, teremos a função

$$f(x) = -y_1^2 + \sum_{i,j=2}^d x_i x_j H_{i,j} - \frac{(\sum_{i=2}^d x_i H_{1,i})^2}{H_{1,1}},$$

caso $H_{1,1} < 0$.

Isolando os termos diferentes de $\pm y_1^2$, observamos que satisfazem as mesmas condições que f , entretanto com menos variáveis, portanto, podemos induzir o mesmo procedimento até que tenhamos

$$f(y_1, \dots, y_d) = -y_1^2 - y_2^2 - \dots - y_q^2 + y_{q+1}^2 + \dots + y_{d-1}^2 + y_d^2.$$

Note que, se $f(u) \neq 0$, então definimos $f_1 : \mathbb{M} \rightarrow \mathbb{R}$ dada por $f_1(x) = f(x) - f(u)$ e aplicamos o argumento para f_1 .

O número de sinais negativos no polinômio quadrático independe dos sistemas de coordenadas. \square

Definição 3.4. *Sejam $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função suave definida em uma variedade suave de dimensão d à valores reais, $u \in \mathbb{M}$ um ponto crítico de f . Considere a função fornecida pelo Lema de Morse*

$$f(x) = f(u) - x_1^2 - x_2^2 - \dots - x_q^2 + x_{q+1}^2 + \dots + x_{d-1}^2 + x_d^2,$$

para todo $x \in U$ em uma vizinhança do ponto crítico u .

Definimos o índice do ponto crítico u como sendo a quantidade de sinais negativos fornecidos pela função dada no Lema de Morse. Denotamos por $index(u) = q$.

O índice classifica os pontos críticos não-degenerados em $(d + 1)$ tipos.

Exemplo 3.1. Para uma variedade de dimensão 2, temos 3 tipos de pontos críticos não-degenerados, classificado pelo índice por:

- Mínimo, $index(u) = 0$.
- Sela, $index(u) = 1$.
- Máximo, $index(u) = 2$.

Já para uma variedade de dimensão 3, temos 4 tipos na classificação dos pontos críticos não-degenerados, por meio do índice:

- Mínimo, $index(u) = 0$.
- Sela, $index(u) = 1$ ou $index(u) = 2$.

- Máximo, $index(u) = 3$.

Definição 3.5. *Seja \mathbb{M} uma variedade suave de dimensão d . Uma função suave $f : \mathbb{M} \rightarrow \mathbb{R}$ é função de Morse se*

1. *Todos os pontos críticos de f forem não-degenerados;*
2. *Pontos críticos tem valores críticos distintos.*

Para definir função de Morse observamos que é possível deixar de exigir a condição 2, porém para nossos propósitos a manteremos.

O Lema de Morse garante que para cada ponto crítico não-degenerado existe uma vizinhança que o separa de outro ponto crítico não-degenerado. De fato, já mencionamos que um ponto u é crítico de f se $D(f \circ \phi)_u : \mathbb{R}^d \rightarrow \mathbb{R}$ não for sobrejetora, ou equivalentemente, se a Jacobiana $J(f \circ \phi)_u = 0$. Assim, se $J(f \circ \phi)_u(x) = (\pm 2x_1, \pm 2x_2, \dots, \pm 2x_d) = (0, 0, \dots, 0)$, em uma vizinhança do ponto crítico u , logo $x = 0$. Portanto, concluímos que existe uma vizinhança de u que o isola dos demais pontos críticos. Implicando que uma função de Morse definida numa variedade compacta possui, no máximo, uma quantidade finita de pontos críticos.

Observação 3.1. A função altura $f : \mathbb{M} \rightarrow \mathbb{R}$ definida no Toro de Pé a valores reais é uma função de Morse. Em contraste, se ao invés tivéssemos considerado o Toro Deitado juntamente com a mesma função f , obteríamos um círculo inteiro de pontos de mínimo e um círculo inteiro de pontos de máximo, onde todos esses pontos são degenerados. Seus índices, nesse caso, não estão definidos.

3.4 Campo de Vetor Gradiente

Um campo de vetor $\chi : \mathbb{M} \rightarrow T\mathbb{M}$ sobre uma variedade suave \mathbb{M} é uma função que associa, para cada $x \in \mathbb{M}$, um vetor $\chi(x) \in T_x\mathbb{M}$ no espaço tangente a \mathbb{M} em x .

Dados uma função $f : \mathbb{M} \rightarrow \mathbb{R}$ suave e um campo de vetores $\chi : \mathbb{M} \rightarrow T\mathbb{M}$, denotamos a derivada direcional de f ao longo do campo de vetores χ por $\chi[f]$; o qual associa cada $x \in \mathbb{M}$ com a derivada direcional de f na direção $\chi(x)$ avaliada no ponto x .

Um particular e muito útil campo de vetor é aquele em que os pontos na direção mais ingreme crescem. Todavia para definí-lo precisamos medir comprimentos, onde introduziremos a métrica de Riemann que é uma variação suave do produto interno definido no espaço tangente.

Se \mathbb{M} for uma variedade suave de dimensão d , imersa em algum espaço Euclidiano, então o espaço tangente $T\mathbb{M}$ será um subespaço linear do mesmo espaço Euclidiano e podemos incorporar a métrica emprestada.

Definição 3.6. *Dados uma variedade suave \mathbb{M} , uma métrica Riemanniana em \mathbb{M} e uma função suave $f : \mathbb{M} \rightarrow \mathbb{R}$, definimos o gradiente de f como o campo de vetor $\nabla f : \mathbb{M} \rightarrow T\mathbb{M}$ caracterizada por $\langle \chi(x), \nabla f(x) \rangle = \chi[f]$ para todo campo de vetor χ .*

Assumindo coordenadas locais com base de vetores ortonormais x_i , o gradiente no ponto x é

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}(x); \frac{\partial f}{\partial x_2}(x); \dots; \frac{\partial f}{\partial x_d}(x) \right]^T.$$

Utilizamos o gradiente para introduzir o grupo de difeomorfismo a 1-parâmetro $\varphi : \mathbb{R} \times \mathbb{M} \rightarrow \mathbb{M}$ que possui duas propriedades características.

Para cada $t \in \mathbb{R}$, $\varphi_t : \mathbb{M} \rightarrow \mathbb{M}$ dada por $x \mapsto \varphi_t(x) = \varphi(t, x)$ é um difeomorfismo de \mathbb{M} sobre \mathbb{M} . Também $\varphi_{t+t_0} = \varphi_t \circ \varphi_{t_0}$ para quaisquer $t, t_0 \in \mathbb{R}$.

O grupo de difeomorfismo a 1-parâmetro é, de fato, um grupo.

Sejam $\varphi_t, \varphi_s, \varphi_r \in \text{Diff}(\mathbb{M})$ com $s, t, r \in \mathbb{R}$.

1. Associatividade.

$$\begin{aligned} \varphi_s \circ (\varphi_r \circ \varphi_t)(x) &= \varphi_s \circ (\varphi_{r+t})(x) = \varphi_s \circ \varphi_{r+t}(x) = \varphi_{s+(r+t)}(x) = \\ &= \varphi_{(s+r)+t}(x) = \varphi_{s+r} \circ \varphi_t(x) = (\varphi_s \circ \varphi_r) \circ \varphi_t(x). \end{aligned}$$

2. Existência do Elemento Neutro.

$Id : \mathbb{M} \rightarrow \mathbb{M}$ difeomorfismo dado por $x \mapsto x$, para todo $x \in \mathbb{M}$.

Para cada $\varphi_t \in \text{Diff}(\mathbb{M})$, temos por um lado,

$$\varphi_t(x) = \varphi_{0+t}(x) = \varphi_0 \circ \varphi_t(x).$$

Por outro,

$$\varphi_t(x) = Id \circ \varphi_t(x).$$

Portanto, $\varphi_0(x) = Id$.

3. Existência do Elemento Simétrico.

Seja $\varphi_s \in \text{Diff}(\mathbb{M})$ para algum $s \in \mathbb{R}$, então existe $(-s) \in \mathbb{R}$, tal que $\varphi_{-s} \in \text{Diff}(\mathbb{M})$ elemento simétrico.

De fato, $\varphi_s \circ \varphi_{-s}(x) = \varphi_{s+(-s)}(x) = \varphi_0(x) = Id(x) = x = Id(x) = \varphi_0(x) = \varphi_{-s+s}(x) = \varphi_{-s} \circ \varphi_s(x)$.

Então, o grupo de difeomorfismos a 1-parâmetro é um grupo.

Esse grupo de difeomorfismos segue a evolução dos conjuntos de subníveis e pode ser utilizado para demonstrar que não existe alterações topológicas entre dois valores críticos vizinhos.

Teorema 3.1. *Sejam \mathbb{M} uma variedade suave de dimensão d , $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função suave e considere valores regulares $a < b$, tais que $f^{-1}[a, b]$ é compacto e não contém pontos críticos de f . Então, \mathbb{M}_a é difeomorfo a \mathbb{M}_b .*

3.5 Anexar Células

O nosso objetivo nesse trabalho é ser capaz de extrair e analisar as informações contidas nos módulos de persistência. Afirmamos que se os valores críticos de f são $a_1 < a_2 < \dots < a_n$, então as informações do módulo de persistência estão contidas em

$$H(\mathbb{M}_{a_1}) \rightarrow H(\mathbb{M}_{a_2}) \rightarrow \dots \rightarrow H(\mathbb{M}_{a_n}),$$

onde H é o funtor homologia singular de dimensão p .

Notemos que o comportamento dos subníveis \mathbb{M}_t evolui ao passo em que $t \in \mathbb{R}$ cresce e alcança os valores críticos da função f , como já mencionamos anteriormente.

O intuito particular desse capítulo é mostrar que, de fato, o comportamento dos subníveis se modifica ao passar pelos valores críticos e que existe uma previsão de como essa mudança ocorrerá, dada através dos índices dos pontos críticos.

A situação é bem diferente da presente no Teorema 3.1, se considerarmos valores regulares $a < b$ tais que $f^{-1}[a, b]$ é compacto mas contém ponto crítico de f .

Seja $u \in f^{-1}[a, b] \subset \mathbb{M}$ um ponto crítico de f com $index(u) = q$. Nesse caso, \mathbb{M}_b possui o mesmo tipo de homotopia de \mathbb{M}_a com uma q -célula anexada em sua fronteira. Para ilustrar o que isso significa, vamos relembrar que a bola unitária de dimensão q , \mathbb{B}^q possui fronteira $\partial\mathbb{B}^q = \mathbb{S}^{q-1}$. Seja $g : \mathbb{S}^{q-1} \rightarrow \partial\mathbb{M}_a$ uma função contínua. A fim de anexar uma célula a \mathbb{M}_a , identificamos cada ponto $x \in \mathbb{S}^{q-1}$ com sua imagem $g(x) \in \partial\mathbb{M}_a$. O único caso um pouco diferente é quando o índice q é zero. Então \mathbb{S}^{-1} é o conjunto vazio e anexar uma 0-célula significa adicionar um ponto.

Exemplo 3.2. Retomando o exemplo da variedade suave de dimensão 2, \mathbb{M} , o do Toro de Pé, e considerando novamente a função altura $f : \mathbb{M} \rightarrow \mathbb{R}$. Temos $u \in \mathbb{M}$ um ponto mínimo global, logo $index(u) = 0$, enquanto v e w são pontos de sela com $index(v) = 1$ e $index(w) = 2$, já $z \in \mathbb{M}$ máximo global com $index(z) = 3$.

Agora, se tomarmos a e $b \in \mathbb{R}$, tais que $a < f(u) < b < f(v)$, com $f^{-1}[a, b]$ compacto contendo u ponto crítico com índice 0, mínimo global. Assim, teremos $\mathbb{M}_a = \emptyset$, enquanto \mathbb{M}_b é homeomorfo ao disco D cujo tipo de homotopia é o do ponto, por ser contrátil. Logo colar uma 0-célula ao conjunto \mathbb{M}_a é anexar um ponto ao conjunto vazio, conseguindo um ponto e, portanto, ambos têm o mesmo tipo de homotopia.

Bem como, se tomarmos $c \in \mathbb{R}$ com $f(v) < c < f(w)$ e considerarmos $f^{-1}[b, c]$ compacto contendo v ponto crítico de f do qual $index(v) = 1$, uma sela. Temos que \mathbb{M}_c é homeomorfo ao cilindro, por conseguinte, tem o mesmo tipo de homotopia que o círculo \mathbb{S}^1 . Já \mathbb{M}_b , que é do mesmo tipo de homotopia de um ponto, ao colarmos uma 1-célula no conjunto, obtemos um círculo, pois anexamos os dois extremos da 1-célula, que consiste em um segmento de reta, no ponto, conseguindo assim um círculo.

Observação 3.2. No exemplo do Toro de Pé, conferimos que todas as informações do módulo de persistência está codificada em apenas 4 pontos críticos.

3.6 Transversalidade

Dada uma função de Morse, podemos percorrer o fluxo do gradiente e decompor a variedade dependendo de onde se origina e de onde termina o fluxo. Para que tal decomposição gere um complexo, requeremos que a função satisfaça uma hipótese adicional de genericidade.

3.6.1 Curvas Integrais

Seja $\varphi : \mathbb{R} \times \mathbb{M} \rightarrow \mathbb{M}$ um grupo de difeomorfismo à 1-parâmetro definido por $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função de Morse definida em uma variedade compacta \mathbb{M} a valores reais, munido da métrica Riemanniana.

A curva integral que passa por $x \in \mathbb{M}$, ponto regular de f , é

$$\gamma = \gamma_x : \mathbb{R} \rightarrow \mathbb{M},$$

dada por

$$\gamma(t) = \gamma_x(t) = \varphi(t, x), \text{ para todo } t \in \mathbb{R},$$

solução da EDO definida por

$$\begin{cases} \dot{\gamma}_t = \nabla f(\gamma_t) \\ \gamma_t(0) = x. \end{cases}.$$

Uma vez que φ e, conseqüentemente γ , são definidas para todo $t \in \mathbb{R}$, e desde que \mathbb{M} é compacto, a curva integral necessariamente se aproxima de um ponto crítico, para t se aproximando tanto de mais infinito quanto de menos infinito.

Denominamos esses pontos críticos de origem e destino da curva integral,

$$\text{org}(\gamma) = \lim_{t \rightarrow -\infty} \gamma(t),$$

$$\text{dest}(\gamma) = \lim_{t \rightarrow +\infty} \gamma(t),$$

respectivamente.

A função cresce ao longo das curvas integrais, o que implica em $\text{org}(\gamma) \neq \text{dest}(\gamma)$.

O Teorema de Existência e Unicidade de EDO nos garante que uma curva integral que passa por outro ponto regular, digamos y , ou é a mesma que passa por x ou as curvas integrais são disjuntas.

Em outras palavras, se γ_y é a curva integral que passa pelo ponto regular $y \in \mathbb{M}$, então $\text{Im}\gamma_y = \text{Im}\gamma_x$ ou senão $\text{Im}\gamma_y \cap \text{Im}\gamma_x = \emptyset$.

A propriedade sugere decompor a variedade em curvas integrais.

3.6.2 Variedades Estáveis e Instáveis

Considere $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função de Morse definida em uma variedade suave de dimensão d a valores reais e tome $u \in \mathbb{M}$ um ponto crítico de f .

Definimos a variedade estável do ponto crítico u de f por

$$S(u) = \{u\} \cup \{x \in \mathbb{M} \mid \text{dest}(\gamma_x) = u\}.$$

E definimos a variedade instável do ponto crítico u de f por

$$U(u) = \{u\} \cup \{x \in \mathbb{M} \mid \text{org}(\gamma_x) = u\}.$$

Note que para cada $x \in \mathbb{M}$ ponto regular de f , a curva integral que passa por ele é definida apenas para pontos regulares, portanto, as variedades estável e instável consistem em apenas pontos regulares, com exceção do ponto crítico do qual ela é construída. Além disso, as palavras “estável” e “instável” fazem alusão aos sistemas de equações diferenciais ordinárias em que temos o comportamento de campo de vetores em torno do ponto crítico. Curvas integrais que tendem ao ponto crítico com o passar do tempo ($t \rightarrow +\infty$) são estáveis porque localmente conseguimos prever e entender o comportamento de pontos próximos e que se mantêm próximos, tal qual na variedade estável $S(u)$ do qual consiste nas curvas integrais que convergem para o ponto crítico. Em outras palavras, $u = \text{dest}(\gamma_x) = \lim_{t \rightarrow +\infty} \gamma_x(t)$. Já a variedade instável $U(u)$ é tal que com o passar do tempo, os pontos se distanciam do ponto crítico.

A função cresce ao longo das curvas integrais, implica que na variedade estável, $S(u)$, temos $f(u) \geq f(x)$ para todos $x \in S(u)$. Por essa razão, muitas vezes referem-se

a ela como variedade descendente de u . Simetricamente, $f(u) \leq f(y)$ para todos os elementos $y \in U(u)$ e, por isso, essa é comumente chamada de variedade ascendente de u .

Suponha uma variedade suave \mathbb{M} de dimensão d e seja q o índice do ponto crítico u . Então existe uma $(q - 1)$ -esfera de direções da qual as curvas integrais se aproximam de u . Além disso, $S(u)$ é uma subvariedade de dimensão q homeomorfa ao \mathbb{R}^q , imersa em \mathbb{M} .

Exemplo 3.3. Retomando o exemplo da variedade suave \mathbb{M} , Toro de Pé, com a função altura $f : \mathbb{M} \rightarrow \mathbb{R}$. Relembrando que os pontos críticos de f são u, v, w, z . E em especial, o crescimento da função altura se dá na direção vertical com sentido para cima.

Para a variedade estável $S(u)$, $S(u) = \{u\}$, uma vez que u é ponto de mínimo global, logo seu índice é 0. Assim, percebemos que estamos compatíveis com a teoria estabelecida. Por outro lado, a variedade instável $U(u)$ consiste do Toro todo, já que todas as curvas integrais se distanciam do mínimo.

No ponto crítico v , uma sela, com $index(v) = 1$, temos a variedade estável $S(v) = \{v\} \cup C_1$, onde C_1 é o segmento de reta aberto proveniente da circunferência retirada seu ponto u , $C_1 = (\mathbb{S}^1 - \{u\})$. Já a variedade instável, $U(v)$, consiste do segmento de reta aberto C_2 obtido da retirada do ponto w da \mathbb{S}^1 .

Simetricamente, temos o caso para o ponto de sela w .

Para z , máximo global com índice 3, temos, para a variedade instável, $U(z) = \{z\}$, uma vez que com o passar do tempo nenhum ponto se distancia dele. Enquanto a variedade estável, $S(z)$, consiste do Toro de Pé retirando as duas curvas C_1 e C_2 , isto é, $S(z) = \{z\} \cup (\mathbb{M} - (C_1 \cup C_2))$.

3.6.3 Funções de Morse-Smale

Sejam $r : \mathbb{R}^p \rightarrow \mathbb{M}$ e $v : \mathbb{R}^q \rightarrow \mathbb{M}$ duas funções suaves. Considere $z \in r(\mathbb{R}^p) \cap v(\mathbb{R}^q)$. Diremos que r e v se intersectam transversalmente em z se

$$dr_x(T_x\mathbb{R}^p) + dv_y(T_y\mathbb{R}^q) = T_z\mathbb{M},$$

onde $x \in r^{-1}(z)$ e $y \in v^{-1}(z)$.

Diremos que r e v são transversais se para todo $z \in r(\mathbb{R}^p) \cap v(\mathbb{R}^q)$, tivermos que r e v se intersectam transversalmente.

Definição 3.7. *Seja \mathbb{M} uma variedade suave de dimensão d . Uma função $f : \mathbb{M} \rightarrow \mathbb{R}$ é dita função de Morse-Smale se f for função de Morse e as variedades estável e instável se intersectam transversalmente.*

3.6.4 Complexos

Assumindo transversalidade, a interseção da variedade estável de dimensão q com a variedade instável de dimensão p tem dimensão $(p + q - d)$, onde recordamos que a dimensão de \mathbb{M} é d . Além disso, a fronteira da variedade estável é a união de variedades estáveis de dimensão menor. O conjunto de variedades estáveis forma um complexo do qual construímos uma dimensão por vez.

- 0-esqueleto: Adiciona-se todos os mínimos, como variedades estáveis de dimensão 0, para inicializar o processo.

- 1-esqueleto: Adiciona-se todas as variedades estáveis de dimensão 1, cada qual um intervalo aberto que colaremos seus pontos extremos nos dois pontos do 0-esqueleto.
- 2-esqueleto: Adiciona-se todas as variedades estáveis de dimensão 2, cada qual um disco aberto que vamos colar a sua fronteira, o círculo, ao longo do 1-esqueleto.

Observação 3.3. Admitimos a condição degenerada para os d -esqueletos que vamos acrescentar o $(d+1)$ -esqueleto. Por exemplo, permitimos que no 0-esqueleto tenhamos apenas um ponto. Assim, ao colar o 1-esqueleto que consiste em um intervalo aberto, colaremos ambas as extremidades no mesmo ponto, conseguindo assim um círculo.

3.6.5 Desigualdade de Morse

Veremos que a soma alternada dos números de variedades estáveis nos fornece a característica de Euler da variedade. Primeiramente, enunciaremos o teorema sobre a característica de Euler.

Teorema 3.2. Euler-Poincaré *A característica de Euler de um espaço topológico é a soma alternada dos números de Betti*

$$\chi_E = \sum_{p \geq 0} (-1)^p \beta_p.$$

Denotando por c_q o número de pontos críticos de índice q , temos:

Teorema 3.3. Desigualdades de Morse

Seja \mathbb{M} uma variedade de dimensão d e seja $f : \mathbb{M} \rightarrow \mathbb{R}$ uma função de Morse. Então,

1. *Fraca:*

$$c_q \geq \beta_q(\mathbb{M}), \text{ para todo } q.$$

2. *Forte:*

$$\sum_{q=0}^j (-1)^{j-q} c_q \geq \sum_{q=0}^j (-1)^{j-q} \beta_q(\mathbb{M}), \text{ para todo } j.$$

A desigualdade forte de Morse, para $j = d$ é uma igualdade.

Somos capazes de recuperar a desigualdade fraca de Morse por meio da desigualdade forte de Morse.

De fato, pela desigualdade forte,

$$\sum_{q=0}^j (-1)^{j-q} c_q \geq \beta_j(\mathbb{M}) + \sum_{q=0}^{j-1} (-1)^{j-q} \beta_q(\mathbb{M}).$$

Assim,

$$\begin{aligned}
 \sum_{q=0}^j (-1)^{j-q} c_q &= (-1)^j c_0 + (-1)^{j-1} c_1 + \dots + (-1)^1 c_{j-1} + (-1)^0 c_j \\
 &= c_j - \sum_{q=0}^{j-1} (-1)^{j-1-q} c_q \\
 &\geq \beta_j(\mathbb{M}) - \sum_{q=0}^{j-1} (-1)^{j-q-1} \beta_q(\mathbb{M}) \\
 \text{Des. Forte para } j-1 &\geq \beta_j(\mathbb{M}) - \sum_{q=0}^{j-1} (-1)^{j-1-q} c_q.
 \end{aligned}$$

Na última desigualdade acima, utilizamos a desigualdade forte de Morse para $j-1$,

$$\sum_{q=0}^{j-1} (-1)^{j-1-q} c_q \geq \sum_{q=0}^{j-1} (-1)^{j-1-q} \beta_q(\mathbb{M}).$$

Assim, invertemos a desigualdade quando multiplicamos por (-1) .

Portanto, concluímos que

$$c_j - \sum_{q=0}^{j-1} (-1)^{j-1-q} c_q \geq \beta_j(\mathbb{M}) - \sum_{q=0}^{j-1} (-1)^{j-1-q} c_q.$$

Removendo os termos em comum, segue o resultado

$$c_j \geq \beta_j(\mathbb{M}).$$

3.7 Funções Lineares por Partes

3.7.1 Funções Simpliciais

O análogo natural de funções contínuas definidas entre espaços topológicos são as funções simpliciais definidas entre complexos simpliciais.

Seja K complexo simplicial com vértices u_0, u_1, \dots, u_n .

Cada $x \in |K|$ pertence ao interior de exatamente um simplexo de K .

Se σ é o simplexo dado pelo fecho convexo de $\{u_0, \dots, u_k\}$,

$$\sigma = \text{fecho convexo}\{u_0, \dots, u_k\},$$

então $x \in \sigma$ se, e somente se,

$$x = \sum_{i=0}^k \lambda_i u_i, \text{ com } \sum_{i=0}^k \lambda_i = 1 \text{ e } \lambda_i \geq 0, i = 0, \dots, k.$$

Definimos as coordenadas baricêntricas de x por $b_i(x)$, com $b_i(x) = \lambda_i$, para $0 \leq i \leq k$ e $b_i(x) = 0$, quando $k+1 \leq i \leq n$.

Segue que $x = \sum_{i=0}^n b_i(x) u_i$.

3.8 Filtração de Lower Star

Apresentaremos o conceito de Lower Star, que transporta a teoria estabelecida até o presente momento de subníveis e filtrações definidas sobre um parâmetro que percorre valores reais para um conjunto finito de pontos, mantendo a mesma estrutura para a aplicação dos funtores de homologia singular de dimensão p e, por conseguinte, a partir da filtração de Lower Star extraímos as mesmas informações. Este tratamento é imprescindível, uma vez que, computacionalmente, somos aptos apenas a trabalhar com quantidades finitas de dados.

Seja K um complexo simplicial e suponha que cada um de seus vértices tenha um valor real associado. Através de extensão linear sobre os simplexos, obtemos uma função linear por partes (PL),

$$f : |K| \rightarrow \mathbb{R},$$

definida por

$$x \mapsto f(x) = \sum_{i=0}^n b_i(x) f(u_i),$$

onde u_i são os vértices de K e b_i as coordenadas baricêntricas de x .

Além disso, é conveniente assumirmos f genérica, o que significa que $f(u_i) \neq f(u_j)$, sempre que $u_i \neq u_j$, onde u_i e u_j são vértices de K . Dessa forma, poderemos ordenar de maneira crescente o valor assumido pela função em seus vértices

$$f(u_0) < f(u_1) < \dots < f(u_n).$$

Para cada $0 \leq i \leq n$, consideramos o subcomplexo completo K_i definido pelos i primeiros vértices. Em outras palavras, dado $\sigma \in K$ um simplexo, $\sigma \in K_i$ se, e somente se, cada vértice u_j de σ satisfaz $j \leq i$.

Cada face de σ é o fecho convexo de um subconjunto não vazio de u_i 's, onde u_i é vértice de σ , e diremos que é uma face própria de σ se o subconjunto não for o conjunto todo.

Se τ for face de σ , então chamaremos σ de coface de τ . Uma vez que um conjunto de tamanho $(k+1)$ tem 2^{k+1} subconjuntos, incluindo o vazio \emptyset , então σ tem $2^{k+1} - 1$ faces das quais todas, exceto uma, são faces próprias.

Fazemos essas observações para definir Star e Lower Star.

A Star de um vértice u_i é o conjunto das cofaces de u_i em K .

A Lower Star de um vértice u_i é o subconjunto de simplexos, tais que u_i é o vértice com maior valor assumido pela f , isto é,

$$St(u_i) = \{\sigma \in K \mid \text{se } x \in \sigma, \text{ então } f(x) \leq f(u_i)\}.$$

No geral, o Lower Star não é um complexo. Consideramos, portanto o Lower Star Fechado de u_i , que é obtido adicionando as faces ausentes do Lower Star ao conjunto, assim o Lower Star Fechado é um subcomplexo.

Pela hipótese de genericidade, cada simplexo possuirá um único vértice máximo, por conseguinte, pertencerá a uma única Lower Star. Disso, segue que os Lower Star particionam K . Ainda mais, K_i é a união dos i primeiros Lower Stars, o que nos motiva a considerar a sequência encaixante

$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K,$$

chamada a filtração por Lower Stars.

4 Topologia Persistente Computacional

4.1 Homologia Persistente

4.1.1 Persistência

O conceito central deste capítulo é motivado pela necessidade prática de lidar com ruído nos dados. Isso inclui definir, reconhecer e, possivelmente, eliminar os ruídos. Estes são objetivos elevados e o desafio pode ser esmagador. De fato, a diferença entre ruído e característica não está bem definida, mas reside nos olhos do observador. Em qualquer caso particular, o foco está em uma gama de escalas e um desejo de ignorar tudo que é menor ou maior.

Em outras palavras, fazemos a nós mesmos a medida de todas as coisas e, fazendo isso, obtemos uma unidade, um ponto de vista e uma opinião.

Motivado por esse pensamento, tomamos uma abordagem agnóstica e oferecemos um significado a uma escala de medidas, uma ferramenta que pode ser usada para fazer julgamentos baseados em informações quantitativas.

4.1.2 Homologia Persistente

Homologia Persistente pode ser usada para medir a escala ou resolução de uma característica topológica. Existem dois ingredientes, um geométrico, definindo uma função em um espaço topológico, e o outro algébrico, transformando a função em medições.

As medidas fazem sentido somente se a função o faz.

A Regra do Mais Velho

Começamos com um cenário simples onde desenvolveremos nossa intuição. Seja X espaço topológico conexo e $f : X \rightarrow \mathbb{R}$ uma função contínua. Os conjuntos de subníveis de f formam uma família a 1-parâmetro de subespaços encaixados, $X_a \subseteq X_b$ sempre que $a \leq b$, onde $X_a = \{x \in X \mid f(x) \leq a\}$.

É conveniente pensar nessa família como a evolução dos subníveis conforme o parâmetro aumenta. Visualizamos essa evolução desenhando cada componente de X_a como um ponto. O resultado é um grafo de dimensão 1, $G(f)$.

Pensando na função f como a função altura, desenhamos o grafo da base para o topo. Uma vez que as componentes nunca desaparecem, os arcos do grafo podem mergir, mas nunca cindem. No final, para um valor a suficientemente grande, teremos

uma única componente. Segue que $G(f)$ é uma árvore, e nos referimos a ela como a árvore fundida da função.

Decompomos essa árvore em caminhos disjuntos que crescem monotonicamente com f . Para obter os caminhos, desenhamo-nos da base ao topo, simultaneamente, enquanto mantemos seus pontos finais superiores na mesma altura a . Os caminhos se estendem, no entanto, quando eles se fundem, finalizamos aquele que começou por último.

Pensando na diferença entre dois valores da função idade, damos prioridade ao caminho mais velho.

A Regra do Mais Velho: Numa conjuntura, o mais velho dos caminhos que mergiram continua e o mais novo acaba.

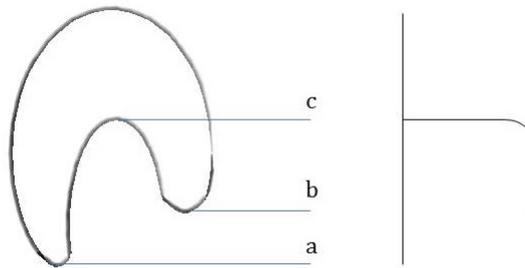


Figura 4.1: A Regra do Mais Velho.

É possível observar pela Figura ?? que obtemos a árvore de fundição ao gravarmos o comportamento dos nascimentos e mortes das classes de homologia segundo a Regra do Mais Velho.

Seja a função altura f definida sobre a variedade e ao aplicar o funtor de homologia singular de dimensão 0 sobre os conjuntos de subníveis, temos:

- $H(X_t) = \emptyset$, para todo $t < a$;
- $H(X_t) = \mathbb{Z}$, para $a \leq t < b$;
- $H(X_t) = \mathbb{Z} \oplus \mathbb{Z}$, para $b \leq t < c$;
- $H(X_t) = \mathbb{Z}$, para todo $t \geq c$.

Conseguimos arquivar na árvore fundida da função o nascimento de uma componente conexa no valor a . Ao passo que o limitante cresce, alcançamos o valor b e temos agora duas componentes conexas, que observamos na ilustração à direita com o surgimento de outro ramo, do qual persistem até o instante em que atingimos o valor c , quando as duas componentes conexas se fundem. E ainda, observamos a Regra do Mais Velho porque a componente mais nova, ou seja, que nasceu por último, se funde com a componente mais velha, isto é, o ramo da árvore fundida que representa a componente conexa nascida em a persiste e a mais nova, que nasceu no instante b , morre.

Seja $a \leq b$ e seja $\beta(a, b)$ o número de componentes em X_b que têm uma interseção não vazia com X_a . Em termos da árvore de fundição, isso é o número de subárvores com pontos superiores no valor b que atingem o nível a ou abaixo.

Cada uma dessas subárvores tem um caminho único, e o mais longo, que gera todo o intervalo entre a e b . Segue-se que $\beta(a, b)$ é também o número de caminhos na decomposição por caminho de $G(f)$ que abrange $[a, b]$. Observamos que qualquer decomposição por caminho que não é gerada usando a Regra do Mais Velho não possui essa propriedade. Em particular, se f é função de Morse, então a Regra do Mais Velho gera uma única decomposição por caminhos, a qual é a única para a qual o número de caminhos abrangendo $[a, b]$ é igual a $\beta(a, b)$, para todos os valores de $a \leq b$.

Filtrações

Obtemos persistência formulando a regra do mais velho para os grupos de homologia de todas as dimensões.

Considere um complexo simplicial K e uma função $f : K \rightarrow \mathbb{R}$. Assumiremos f monótona, isto é, $f(\sigma) \leq f(\tau)$ sempre que σ for uma face de τ . Monotonicidade implica que o conjunto de subnível $K(a) = f^{-1}(-\infty, a]$ é um subcomplexo de K , para todo $a \in \mathbb{R}$.

Seja m o número de simplexes em K . Conseguimos $n + 1 \leq m + 1$ subcomplexos distintos de K , os quais estão arrançados em uma sequência

$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K,$$

do seguinte modo: se $a_1 < a_2 < \dots < a_n$ são os valores da função nos simplexes em K e $a_0 = -\infty$, então $K_i = K(a_i)$, para todo i .

Chamamos essa sequência de complexos de filtração de f e pensamos nela como uma construção onde adicionamos pedaços de simplexes com o tempo. Vimos exemplos anteriormente, a saber, a filtração de Lower Star de uma função linear por partes.

Mais do que na sequência de complexos, estamos interessados na evolução topológica, expressa pela sequência correspondente de grupos de homologia.

Para todo $i \leq j$ temos a aplicação inclusão de K_i em K_j e, portanto, o homomorfismo induzido

$$f_p^{i,j} : H_p(K_i) \rightarrow H_p(K_j),$$

onde H_p é a homologia simplicial de dimensão p com coeficientes em \mathbb{Z}_2 .

A filtração então nos fornece uma sequência de grupos de homologia conectados por homomorfismos

$$0 = H_p(K_0) \rightarrow H_p(K_1) \rightarrow \dots \rightarrow H_p(K_n) = H_p(K),$$

para cada $p \in \mathbb{N}$.

Conforme vamos de K_{i-1} para K_i , podemos ganhar novas classes de homologia, como podemos também perder algumas, que podem se tornar triviais ou se fundir com alguma outra.

Coletamos as classes que nascem em um dado nível ou antes e morrem depois de outro nível em grupos do seguinte modo.

Definição 4.1. Os p -ésimos grupos de homologia persistente são as imagens dos homomorfismos induzidos pela inclusão

$$H_p^{i,j} = \text{Im } f_p^{i,j} \quad \text{para } 0 \leq i \leq j \leq n.$$

Os p -ésimos números de Betti persistentes correspondentes são as dimensões desses espaços vetoriais

$$\beta_p^{i,j} = \dim H_p^{i,j}.$$

Os grupos de homologia persistente consistem das classes de homologia de K_i que permanecem vivas em K_j ou, mais formalmente,

$$H_p^{i,j} \approx \frac{Z_p(K_i)}{B_p(K_j) \cap Z_p(K_i)}.$$

De fato, pelo Teorema do Isomorfismo,

$$H_p^{i,j} \approx \frac{H_p(K_i)}{\ker f_p^{i,j}}.$$

Por outro lado,

$$\ker f_p^{i,j} = \{c + B_p(K_i) \mid c \in B_p(K_j) \cap Z_p(K_i)\}.$$

Portanto,

$$H_p^{i,j} = \frac{\frac{Z_p(K_i)}{B_p(K_i)}}{\frac{B_p(K_j) \cap Z_p(K_i)}{B_p(K_i)}} \approx \frac{Z_p(K_i)}{B_p(K_j) \cap Z_p(K_i)}.$$

Definição 4.2. *Seja γ uma classe de homologia de dimensão p em $H_p(K_i)$.*

Dizemos que γ nasce em K_i se $\gamma \notin H_p^{i-1,i}$.

Além disso, se γ nasce em K_i , diremos que γ morre ao entrar em K_j se γ se fundir com uma classe mais velha, conforme vamos de K_{j-1} para K_j . Assim, $f_p^{i,j-1}(\gamma) \notin H_p^{i-1,j-1}$, mas $f_p^{i,j}(\gamma) \in H_p^{i-1,j}$.

Essa é, novamente, a Regra do Mais Velho.

É relevante e útil notar que a definição acima diz que se γ nasce em K_i , isso significa que, desde que $H_p^{i-1,i} = \text{Im}\{f_p^{i-1,i} : H_p(K_{i-1}) \rightarrow H_p(K_i)\}$, significa que não existe nenhuma classe de homologia de dimensão p , $\alpha \in H_p(K_{i-1})$, tal que $f_p^{i-1,i}(\alpha) = \gamma$.

Se γ morre ao entrar em K_j então significa que não existe nenhuma classe de homologia de dimensão p , $\beta \in H_p(K_{i-1})$, tal que $f_p^{i-1,j-1}(\beta) = f_p^{i,j-1}(\gamma)$ em $H_p(K_{j-1})$, mas existe $\Delta \in H_p(K_{i-1})$ classe de homologia de dimensão p , tal que $f_p^{i-1,j}(\Delta) = f_p^{i,j}(\gamma)$ em $H_p(K_j)$.

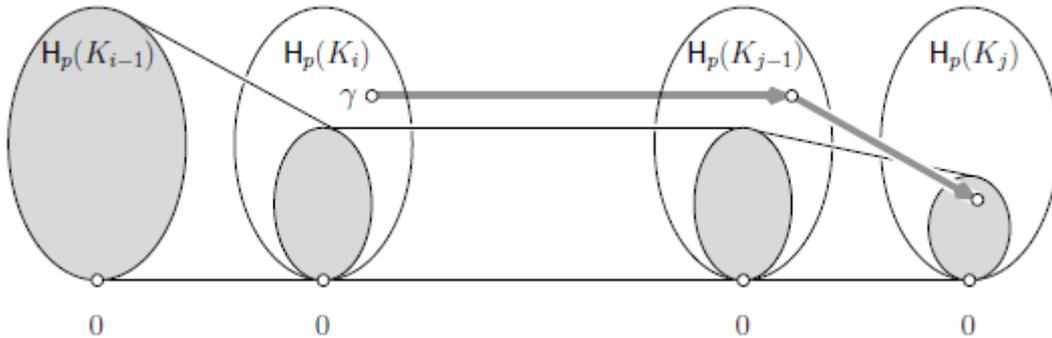


Figura 4.2: A classe γ nasce em K_i , uma vez que não jaz na imagem de $H_p(K_{i-1})$ (sombreada). Além disso, γ morre ao entrar em K_j pois é o primeiro momento em que sua imagem se funde para dentro da imagem de $H_p(K_{i-1})$.

Definição 4.3. Se γ nasce em K_i e morre ao entrar em K_j , chamamos a diferença $a_j - a_i$ de persistência de γ . Denotaremos por $\text{pers}(\gamma) = a_j - a_i$, onde a_l é o valor da função do simplexo em K , $K_l = K(a_l) = f^{-1}(-\infty, a_l]$.

Algumas vezes decidimos ignorar os valores a_i e a_j da função e simplesmente considerar a diferença de índices $j-i$, a qual chamaremos persistência do índice da classe γ .

Caso tenhamos uma classe γ que nasce em K_i , porém nunca morre, então definiremos sua persistência, bem como seu índice de persistência, como sendo infinito.

4.1.3 Diagramas de Persistência

Vizualizamos a coleção dos números de Betti persistentes desenhando pontos no plano. Alguns desses pontos, talvez, tenham coordenadas igual a infinito, e alguns podem coincidir, então falamos realmente de um multiconjunto de pontos no plano estendido

$$\bar{R}^2 = (\bar{\mathbb{R}} \cup \{\infty\})^2.$$

Seja $\mu_p^{i,j}$ o número de classes linearmente independentes em $H_p(K_i)$ que nascem em K_i e morrem quando entram em K_j . Então, temos

$$\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j}),$$

para todo $i < j$ e qualquer p .

O primeiro termo $\beta_p^{i,j-1}$ pode ser interpretado como o número de classes de homologia independentes em $H_p(K_{j-1})$ que nasceram antes de, ou em $H_p(K_i)$. A primeira diferença $(\beta_p^{i,j-1} - \beta_p^{i,j})$ computa as classes em $H_p(K_{j-1})$ que nasceram antes de, ou em $H_p(K_i)$ e morrem antes de $H_p(K_j)$.

Similarmente, a segunda diferença, $(\beta_p^{i-1,j-1} - \beta_p^{i-1,j})$, calcula as classes em $H_p(K_{j-1})$ que nascem antes de, ou em $H_p(K_{i-1})$ e morrem antes de $H_p(K_j)$. Assim, segue que $\mu_p^{i,j}$ conta as classes que nascem entre $H_p(K_{i-1})$ e $H_p(K_i)$ e morrem entre $H_p(K_{j-1})$ e $H_p(K_j)$.

Considerando

$$H_p(K_i) \xrightarrow{f_p^{i,j-1}} H_p(K_{j-1}) \xrightarrow{f_p^{j-1,j}} H_p(K_j),$$

$$\begin{aligned} \beta_p^{i,j} &= \dim \text{Im } f_p^{i,j} = \dim \{ \text{Im}(f_p^{j-1,j} \circ f_p^{i,j-1}) \} = \dim(f_p^{j-1,j} |_{H_p^{i,j-1}}) \\ &= \dim H_p^{i,j-1} - \dim \ker(f_p^{j-1,j} |_{H_p^{i,j-1}}) = \beta_p^{i,j-1} - \dim \ker(f_p^{j-1,j} |_{H_p^{i,j-1}}). \end{aligned}$$

Portanto, obtemos

$$\beta_p^{i,j} = \beta_p^{i,j-1} - \dim \ker(f_p^{j-1,j} |_{H_p^{i,j-1}}),$$

implicando que

$$\dim \ker(f_p^{j-1,j} |_{H_p^{i,j-1}}) = \beta_p^{i,j-1} - \beta_p^{i,j}.$$

Analogamente,

$$\dim \ker(f_p^{j-1,j} |_{H_p^{i-1,j-1}}) = \beta_p^{i-1,j-1} - \beta_p^{i-1,j}.$$

Definição 4.4. Desenhando cada ponto (a_i, a_j) com multiplicidade $\mu_p^{i,j}$, obtemos o p -ésimo diagrama de persistência da filtração, denotado por $Dgm_p(f)$.

Ele representa uma classe por um ponto cuja distância vertical a diagonal é sua persistência.

Uma vez que as multiplicidades são definidas para $i < j$, todos os pontos vivem acima da diagonal.

Por razões técnicas (ficará claro mais tarde quando abordarmos teoria de estabilidade), adicionamos os pontos sobre a diagonal do diagrama, cada um com multiplicidade infinita.

É fácil achar os números de Betti persistentes. Especificamente, $\beta_p^{k,l}$ é o número de pontos no quadrante esquerdo superior com vértice (a_k, a_l) . Uma classe que nasce em K_i e morre ao entrar em K_j é contada se, e somente se, $a_i \leq a_k$ e $a_j > a_l$. O quadrante é, portanto, fechado ao longo de seu lado vertical direito e aberto ao longo de seu lado horizontal inferior.

A teoria de Persistência é edificada, essencialmente, sobre três resultados fundamentais (Teorema 4.1, Teorema 4.3 e o Teorema 5.1), cujos quais podem ser encontrados em [4]. Estes caracterizam a codificação e arquivamento das informações presentes nos diagramas de persistência, bem como o comportamento de diagramas de persistência e sua estabilidade.

Apresentaremos o primeiro deles.

Teorema 4.1. Lema Fundamental da Homologia Persistente

Seja

$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K$$

uma filtração. Para todo par de índices $0 \leq k \leq l \leq n$ e para qualquer dimensão p , o p -ésimo número de Betti persistente é

$$\beta_p^{k,l} = \sum_{i \leq k} \sum_{j > l} \mu_p^{i,j}.$$

Demonstração. A demonstração se dará por indução em l .

Para $l = 0$, então $k = 0$, já que $0 \leq k \leq 0 \leq n$.

Devemos mostrar que

$$\beta_p^{0,0} = \sum_{j \geq 1} \mu_p^{0,j}.$$

Como $K_0 = \emptyset$, temos que $H_p(K_0) = \{0\}$, para todo p .

Assim,

$$\beta_p^{0,0} = \dim H_p^{0,0} = \dim \text{Im}\{f_p^{0,0} : H_p(K_0) \rightarrow H_p(K_0)\} = 0.$$

Além disso, $0 \leq i \leq k = 0$. E, por definição

$$\sum_{j \geq 1} \mu_p^{0,j} = \sum_{j \geq 1} (\beta_p^{0,j-1} - \beta_p^{0,j}) - (\beta_p^{-1,0} - \beta_p^{-1,j}),$$

observamos que cada um dos $\beta_p = 0$.

Concluimos que

$$\beta_p^{0,0} = 0 = \sum_{j \geq 1} \mu_p^{0,j}.$$

Logo para $l = 0$ a fórmula vale.

Supondo válido para l , vamos mostrar para $(l + 1)$.

Note que, utilizando a hipótese de indução, resta mostrar que

$$\beta_p^{k,l} - \beta_p^{k,l+1} = \sum_{i=0}^k \mu_p^{i,l+1}.$$

De fato,

$$\begin{aligned} \sum_{i=0}^k \mu_p^{i,l+1} &= (\beta_p^{i,l} - \beta_p^{i,(l+1)}) - (\beta_p^{(i-1),l} - \beta_p^{(i-1),(l+1)}) \\ &= (\beta_p^{0,l} - \beta_p^{0,(l+1)}) - (\beta_p^{(-1),l} - \beta_p^{(-1),(l+1)}) + \\ &+ (\beta_p^{2,l} - \beta_p^{2,(l+1)}) - (\beta_p^{(1),l} - \beta_p^{(1),(l+1)}) + \dots \\ &+ \dots + (\beta_p^{(k-1),l} - \beta_p^{(k-1),(l+1)}) - (\beta_p^{(k-2),l} - \beta_p^{(k-2),(l+1)}) + \\ &+ (\beta_p^{k,l} - \beta_p^{k,(l+1)}) - (\beta_p^{(k-1),l} - \beta_p^{(k-1),(l+1)}). \end{aligned}$$

Notemos que os termos da parcela anterior se cancelam com os termos da parcela seguinte, conforme i aumenta. Logo, a somatória se resume aos termos $(\beta_p^{k,l} - \beta_p^{k,(l+1)})$. Portanto,

$$\sum_{i=0}^k \mu_p^{i,l+1} = (\beta_p^{k,l} - \beta_p^{k,(l+1)}).$$

□

Essa é uma importante propriedade que diz que o diagrama codifica todas as informações sobre os grupos de homologia persistente.

4.1.4 Redução Matricial

Complexos de Cadeia

Um complexo de cadeia $(C_\bullet, \partial_\bullet)$

$$\dots \xrightarrow{\partial_{i+2}} C_{i+1}(K) \xrightarrow{\partial_{i+1}} C_i(K) \xrightarrow{\partial_i} C_{i-1}(K) \xrightarrow{\partial_{i-1}} C_{i-2}(K) \xrightarrow{\partial_{i-2}} \dots$$

é uma sequência de módulos C_i conectados por homomorfismos bordo ∂_i , tal que a composição entre quaisquer dois homomorfismos bordo consecutivos é a transformação nula, isto é, $\partial_n \partial_{n+1} = 0$, para todo n .

No nosso caso, $\partial_i : C_i(K) \rightarrow C_{i-1}(K)$ é o homomorfismo bordo cuja definição de bordo de um i -simplexo, pertencente ao $C_i(K)$, é a soma de suas faces de dimensão $(i-1)$ (estamos considerando coeficientes em \mathbb{Z}_2).

Uma vez que ∂_i é um homomorfismo, segue que $\partial_i(nc+md) = n\partial_i(c) + m\partial_i(d)$ desde que n, m pertencem ao corpo de coeficientes, \mathbb{Z}_2 , com c e $d \in C_i(K)$. Além disso, já sabemos que toda transformação linear pode ser representada por uma matriz, uma vez fixada a base de domínio e contra-domínio.

Escolhemos o conjunto dos i -simplexos como base de C_i e o conjunto dos $(i-1)$ -simplexos como base de C_{i-1} , e fixamos uma ordenação (ainda que arbitrária) em tais simplexos.

Dessa maneira, ∂_i é representada pela matriz bordo $D_i = \{a_{i,j}\}$ que possui $n_{(i-1)}$ linhas e n_i colunas, da qual cada linha é indexada por $(i-1)$ -simplexos enquanto as colunas são indexadas por i -simplexos.

A j -ésima coluna, que representa um i -simplexo particular, contém 1 para cada linha indexada por uma $(i - 1)$ -face do j -ésimo i -simplexo, ao passo que as demais entradas são nulas.

Definimos o conjunto dos i -ciclos $Z_i(K)$ como:

$$Z_i(K) = \{\alpha \in C_i(K) \mid \partial_i(\alpha) = 0\} = \ker \partial_i$$

e o conjunto dos i -bordos $B_i(K)$ como

$$B_i(K) = \{\gamma \in C_i(K) \mid \exists \sigma \in C_{i+1}(K), \text{ tal que } \partial_{i+1}\sigma = \gamma\} = \text{Im } \partial_{i+1}.$$

O i -ésimo grupo de homologia de K é $H_i(K) = \frac{Z_i(K)}{B_i(K)}$.

Notamos ainda que o i -ésimo número de Betti de K é $\beta_i = \text{rank}(H_i(K)) = \text{rank}\left(\frac{Z_i(K)}{B_i(K)}\right) = \text{rank}(Z_i(K)) - \text{rank}(B_i(K))$. Assim, se definirmos $z_i = \text{rank } Z_i(K)$ e $b_i = \text{rank } B_i(K)$, temos que $\beta_i = z_i - b_i$. Essa observação se tornará muito relevante quando estivermos calculando a homologia persistente por métodos computacionais.

A fim de explicar a teoria apresentada, exibiremos um exemplo que é valioso por trazer elementos fundamentais que estarão presentes na edificação da teoria da construção da matriz bordo e da redução matricial, que é uma de nossas ferramentas computacionais para obtermos diagramas de persistência.

Exemplo 4.1. Considere o complexo simplicial K

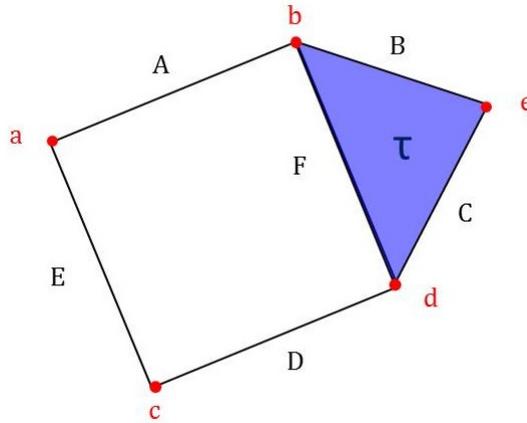


Figura 4.3: Complexo simplicial

A fim de calcular seus complexos de cadeia, observamos que primeiramente necessitamos descrever os grupos de cadeia por meio de seus geradores:

- $C_0(K) = \langle a, b, c, d, e \rangle$,
- $C_1(K) = \langle A, B, C, D, E, F \rangle$,
- $C_2(K) = \langle \tau \rangle$,

considerando

$$C_2(K) \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0.$$

Portanto, para o exemplo acima, podemos escrever ∂_2 e ∂_1 em notação matricial, somente sendo necessário estabelecermos uma base para os espaços.

Logo, o homomorfismo bordo ∂_2 pode ser descrito como

$$\partial_2 = \begin{bmatrix} & \tau \\ A & 0 \\ B & 1 \\ C & 1 \\ D & 0 \\ E & 0 \\ F & 1 \end{bmatrix}.$$

Assim, relembrando o que temos na teoria. A coluna 1 representa o nosso 2-simplexo que denominamos τ e contém 1 para cada linha indexada por 1-simplexos que é uma face de τ . Concluimos que as faces de τ são as arestas B , C e F .

Enquanto, podemos descrever o homomorfismo bordo ∂_1

$$\partial_1 = \begin{bmatrix} & A & B & C & D & E & F \\ a & 1 & 0 & 0 & 0 & 1 & 0 \\ b & 1 & 1 & 0 & 0 & 0 & 1 \\ c & 0 & 0 & 0 & 1 & 1 & 0 \\ d & 0 & 0 & 1 & 1 & 0 & 1 \\ e & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Analogamente, temos que, por exemplo, a coluna 4 representa o nosso 1-simplexo que chamamos de D , cuja as linhas contendo o número 1 são exatamente as linhas que representam os 0-simplexos c e d , suas faces de codimensão 1.

Observação 4.1. Sejam σ_j , σ_k e σ_i os simplexos representados pelos índices j , k e i , respectivamente.

Ao observarmos na matriz bordo ∂ que a coluna j possui 1 nas posições k e i , significa que σ_i e σ_k são faces de codimensão 1 do simplexo σ_j .

Usando a abordagem do homomorfismo bordo, notamos que $\partial(\sigma_j) = \sigma_k + \sigma_i$.

Assim, a matriz bordo criptografa a informação recebida pelo homomorfismo bordo. Ler que determinados índices k e i são face de codimensão 1 de determinada coluna j é o mesmo que afirmar que a coluna j tem determinados índices k e i como seu bordo, pois os índices estão biunivocamente determinados com seus simplexos.

Redução Matricial

Ao contrário do grupo fundamental, existe um algoritmo bem definido para computar grupo de homologia de um complexo simplicial K arbitrário, originalmente feito por Poincaré [6], chamado algoritmo da redução. Além dos próprios grupos, o algoritmo de redução computa bases para cada grupo de homologia: um conjunto de i -ciclos dos quais as classes de homologia geram $H_i(K)$. Jeff Erickson [8] identificou que, na verdade, o algoritmo de redução de Poincaré é equivalente ao algoritmo publicado quatro décadas antes por Smith [7] a fim de calcular certa forma normal de uma matriz inteira. No entanto, Poincaré, aparentemente, estava desavisado do fato. O algoritmo de Smith por sua vez é uma variação do algoritmo padrão de eliminação de Gauss do qual, de fato, foi descoberto pelos matemáticos chineses por volta de 100 DC [8].

Além de ter uma descrição compacta em termos de diagramas, a persistência também pode ser calculada de forma eficiente. O algoritmo particular que usamos é uma versão da redução matricial.

Surpreendentemente, podemos obter todas as informações com uma única redução.

Para descrever isso, usamos uma ordenação compatível dos simplexes, ou seja, uma sequência $\sigma_1, \sigma_2, \dots, \sigma_m$, tal que se $i < j$ ou se tivermos σ_i uma face de σ_j , então $f(\sigma_i) < f(\sigma_j)$. Nesse momento, estamos estabelecendo uma base para nossos grupos de cadeia, o que outrora era irrelevante.

Tal ordenação existe porque f é uma função monotônica. Note que cada subsequência inicial de simplexes formam um subcomplexo de K .

Usaremos essa sequência proveniente da ordenação compatível quando formos definir a matriz bordo $m \times m$, ∂ , que armazena todos os simplexes de todas as dimensões em um único lugar. Em outras palavras, queremos construir uma matriz bordo que contenha todas as matrizes bordos D_i proveniente dos homomorfismos bordo ∂_i .

A matriz bordo será definida da seguinte forma

$$\partial[i, j] = \begin{cases} 1, & \text{se } \sigma_i \text{ for uma face de codimensão 1 de } \sigma_j \\ 0, & \text{caso contrário} \end{cases}.$$

Notamos que o índice k representa o k -ésimo simplexo que segundo a ordenação compatível representa o σ_k .

Tal qual anteriormente, estamos escrevendo o número 1 posicionado na coluna j que representa o σ_j e na linha i que representa o σ_i , sempre que σ_i for uma face de codimensão 1 de σ_j . E colocaremos 0 nessa posição, caso contrário.

Assim, englobamos todos os simplexes em uma única matriz e ainda temos a informação de suas respectivas faces.

A ordenação compatível, nesse caso, serve não somente para organizar os simplexes em ordem crescente de simplexo segundo a função f , isto é, ponto, aresta, face e assim sucessivamente, mas também para armazenar de forma conveniente as faces dos simplexes, pois vamos estudar a matriz em blocos, sendo que cada bloco consiste em simplexes de uma determinada dimensão. Portanto, apesar de termos todos os simplexes juntos, continuaremos mantendo o controle da matriz e de suas informações.

Claramente perderíamos domínio da informação contida em ∂ caso a ordenação fosse arbitrária. Em outras palavras, colunas e linhas são ordenadas da mesma forma que os simplexes na ordenação total e o bordo de determinado simplexo é gravado em suas colunas. O algoritmo utiliza operações de coluna para reduzir a matriz bordo ∂ em outra matriz 0 – 1, que denominaremos R , a matriz reduzida.

Definição 4.5. *Sejam $\{\sigma_i\}_{i=1}^m$ os m simplexes ordenados na ordenação compatível segundo a função f . Considere ∂ , a matriz bordo.*

Definiremos o lowest j , $low(j)$, como o índice da linha que contém o mais baixo 1 da coluna j .

Caso a coluna j seja nula, então o $low(j)$ é indefinido.

O lowest 1, ou também conhecido por mais baixo 1, da coluna j é o último número 1 que aparece na coluna j , no sentido de cima para baixo.

Denominaremos R matriz reduzida se $low(j) \neq low(j_0)$ sempre que $j \neq j_0$ para quaisquer índices de colunas não nulas.

Exemplo 4.2. Nas condições da definição acima, afirmar que na matriz bordo 7×7 , ∂ , temos $3 = low(7)$ significa que na posição $\partial[3, 7] = 1$ e ainda $\partial[i, 7] = 0$ para $4 \leq i \leq 7$. Portanto, $3 = low(7)$ nos fornece que a linha 3 da coluna 7 temos o 1 mais baixo.

Também podemos extrair a informação de que σ_3 é uma das faces de codimensão 1 de σ_7 .

Algoritmo de Redução

O algoritmo que reduz ∂ através de operações coluna da esquerda para a direita é descrito em linguagem de pseudo-código como

```

R = ∂
for j = 1 to m do
  while there exists j0 < j with low(j0) = low(j) do
    add column j0 to column j
  endwhile
endfor

```

Complexidade de Algoritmo: A complexidade de algoritmo analisa e determina a quantidade necessária de tempo para execução do código e também a quantidade de memória necessária, além de eventuais outros recursos. Este estudo é importante para desenvolver códigos eficientes, ou seja, que a medida de complexidade seja pequena.

Ao analisarmos a complexidade do Algoritmo de Redução, conseguimos que sua ordem de complexidade é, no máximo, cúbica em relação ao número de simplexes. Com efeito, temos dois laços, sendo eles um **for** e um **while**, logo já atingimos complexidade de algoritmo quadrática, contudo ainda temos a função **low** que busca e armazena o mais baixo 1.

Assim, para uma fixada coluna j , ocorre a comparação de todos os simplexes à direita dela, logo no pior dos casos temos a comparação da última coluna até a primeira, e ainda, obtemos o pior caso quando na última coluna, tivermos o $\text{low}(j)$ na última linha, assim percorreremos a última coluna inteira. Dessa maneira obtemos ordem cúbica, no máximo.

Em notação matricial, o algoritmo calcula a matriz reduzida R como $R = \partial V$. Desde que cada simplexo é precedido pelos simplexes que são suas faces, por conta da ordenação compatível, segue que ∂ é triangular superior. As colunas j de V codifica as colunas em ∂ que foram adicionadas entre si para obtermos um coluna j em R .

Implica que V guarda o caminho percorrido pelas matrizes através das operações colunas efetuadas para se conseguir a matriz R .

Uma vez que só adicionamos da esquerda para a direita, temos que V é triangular superior. Por conseguinte, R enquanto produto de duas matrizes que são triangulares superior, também o é.

Para extraírmos os ranks dos grupos de homologia de dimensão p de K , nossa tarefa inicial é identificar nas colunas da matriz R os n simplexes de dimensão p e os m simplexes de dimensão $(p + 1)$, feito isso, basta contar o número de colunas nulas de R que correspondem aos p -simplexes e nomearemos de z_p . Agora, contaremos os números de colunas não nulas associadas aos $(p + 1)$ -simplexes e denominaremos b_p .

Por fim, para termos o p -ésimo número de Betti, computamos

$$\beta_p = z_p - b_p.$$

Notemos que esse processo é apenas uma formulação técnica da observação mencionada anteriormente.

O p -ésimo grupo de homologia $H_p(K) = \frac{Z_p(K)}{B_p(K)}$, logo

$$\beta_p = \text{rank}(H_p(K)) = \text{rank}\left(\frac{Z_p(K)}{B_p(K)}\right) = \text{rank}(Z_p(K)) - \text{rank}(B_p(K)),$$

chamando $z_p = \text{rank}(Z_p(K))$ e $b_p = \text{rank}(B_p(K))$ para conseguirmos

$$\beta_p = z_p - b_p.$$

Percebamos que $Z_p = \ker \partial_p$, portanto, são os p -simplexos que avaliados valem 0. O que é equivalente à contar, na matriz reduzida, a quantidade de colunas nulas dos p -simplexos, pois eles gerarão o p -ciclos.

Já $B_p = \text{Im } \partial_{i+1}$ são os $(p+1)$ -simplexos mapeados em p -simplexos não nulos, então resta contar as colunas não nulas, na matriz R , que representam os $(p+1)$ -simplexos, pois esses serão a base do p -bordo.

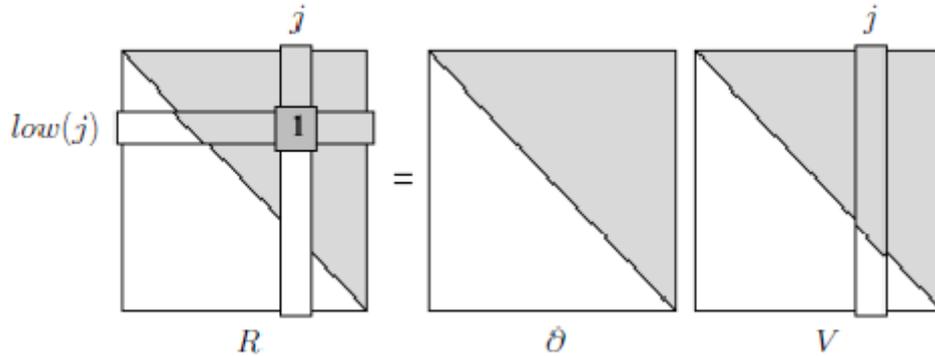


Figura 4.4: A redução da matriz bordo vista como produto de matrizes. O elementos presentes na área em branco são necessariamente 0, enquanto na área sombreada pode ser 0 ou 1.

Pareamento

No entanto, existe significativamente mais informação do que podemos colher. Para vermos isso, precisamos entender como o lowest 1 se relaciona com os grupos de homologia persistente. Iniciaremos mostrando que ele é único apesar de que a matriz reduzida R não é. De fato, R é caracterizado por ser reduzida e obtida por operações colunas da esquerda para a direita. Contudo, podemos ou não continuar as operações uma vez que alcançarmos uma matriz reduzida.

Para ver que lowest 1 são únicos, consideramos a submatriz inferior esquerda R_i^j de R cujo elemento no canto superior direito é $R[i, j]$, em outras palavras, R_i^j é obtida de R removendo as primeiras $(i-1)$ linhas e as colunas da posição $(j+1)$ até m .

Uma vez que as operações da esquerda para a direita preservam o rank de toda submatriz, o rank de R_i^j é o mesmo da submatriz correspondente em ∂ que também é obtida removendo as primeiras $(i-1)$ linhas e as colunas da posição $(j+1)$ até m .

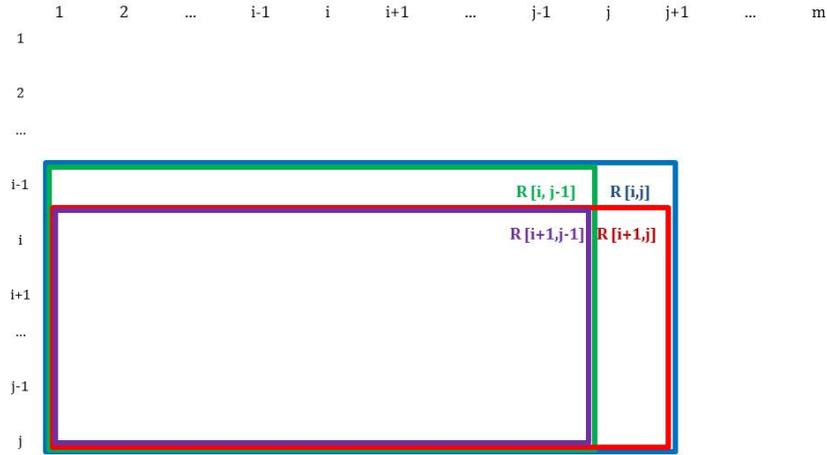


Figura 4.5: Ilustração das submatrizes R_i^j , R_{i+1}^j , R_{i+1}^{j-1} e R_i^{j-1} da matriz R .

Consideremos a expressão

$$r_R(i, j) = \text{rank}R_i^j - \text{rank}R_{i+1}^j + \text{rank}R_{i+1}^{j-1} - \text{rank}R_i^{j-1}.$$

Note que $r_R(i, j) = r_\partial(i, j)$ para todo i e j , onde $r_\partial(i, j)$ tem definição análoga exceto pelo fato que tomamos os ranks das submatrizes de ∂ .

A fim de avaliar essa expressão, observamos que combinação linear de qualquer coleção de colunas não nulas de R_i^j é novamente não-nula. Segue que o rank de R_i^j é igual ao número de colunas não-nulas.

Agora, se $R[i, j]$ for lowest 1, então R_i^j possui uma coluna não-nula a mais do que as três submatrizes.

De fato, notemos que $R[i+1, j] = 0$, pois $R[i, j]$ é lowest 1 e mais que isso $R[k, j] = 0$, para $i+1 \leq k \leq n$, já que se houvesse alguma posição abaixo de $R[i, j]$ com $R[k, j] = 1$, para $i+1 \leq k \leq n$, contrariaria o fato de $R[i, j]$ ser lowest 1.

Além disso, por construção de R temos que se $R[i, j]$ é lowest 1, então $R[i, j-1]$ não é lowest 1, graças ao algoritmo que percorre todas as colunas da esquerda para a direita e sempre que ocorrer dois lowest 1 em colunas distintas, ocorre a soma das colunas para que possamos alcançar a matriz reduzida. De fato, concluímos que $R[i, k]$ não é lowest 1, para $0 \leq k \leq j-1$.

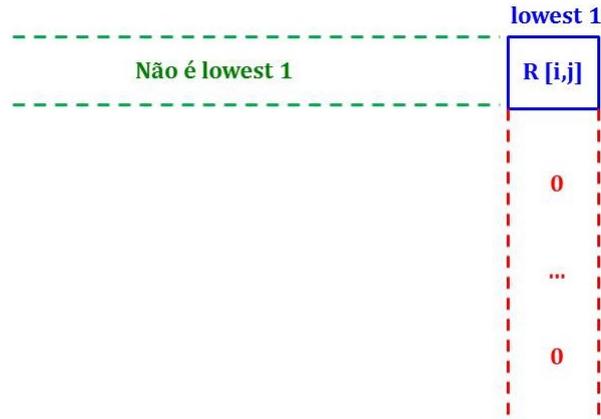


Figura 4.6: se $R[i, j]$ for lowest 1, então R_i^j possui uma coluna não-nula a mais do que as três submatrizes.

Logo R_i^j possui uma coluna não-nula a mais que R_{i+1}^j , R_{i+1}^{j-1} e R_i^{j-1} .

Consequentemente, $r_R(i, j) = 1$.

Caso $R[i, j]$ não é lowest 1, considere os casos abaixo.

Caso 1: $R[i, j]$ não é lowest 1 e nenhuma das colunas de 1 à $j - 1$ possuírem lowest 1 na linha i , então R_i^j e R_{i+1}^j têm o mesmo número de colunas não nulas.

De fato, por hipótese assumimos que $R[i, k]$ não é lowest 1, para $1 \leq k \leq j$, assim para as posições $R[i, k] = 1$ teremos que existirá um outro índice $R[l, k]$, com $i < l \leq n$ tal que $R[l, k] = 1$, ou seja, os lowest 1 estão no interior da matriz R_{i+1}^j .

Logo todas as colunas não-nulas são perceptíveis pelas matrizes R_i^j e R_{i+1}^j , portanto elas têm o mesmo número de colunas não nulas. Analogamente, concluímos que R_i^{j-1} e R_{i+1}^{j-1} têm o mesmo número de colunas não-nulas.

Assim,

$$(\text{rank}R_i^j - \text{rank}R_{i+1}^j) = 0 \text{ e } (\text{rank}R_{i+1}^{j-1} - \text{rank}R_i^{j-1}) = 0,$$

e portanto,

$$r_R(i, j) = 0.$$

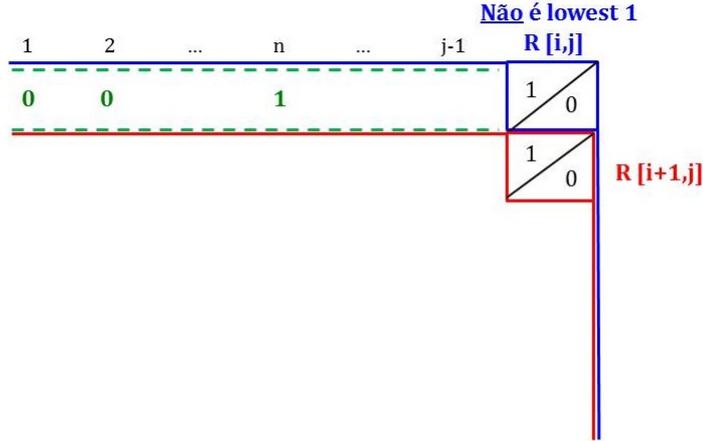


Figura 4.7: $R[i, j]$ não é lowest 1 e nenhuma das colunas de 1 à $j - 1$ possuem lowest 1 na linha i .

Caso 2: $R[i, j]$ não é lowest 1 e uma das colunas de 1 à $j - 1$ possui lowest 1.

Temos que $R[i, n] = 1$ é lowest 1, com $1 \leq n \leq j - 1$. Observamos que R_i^j tem uma coluna não-nula a mais que R_{i+1}^j , pois na posição $R[i, n] = 1$ e está contida na matriz R_i^j , todavia, não está presente na matriz R_{i+1}^j , pois na coluna n , ela se inicia na posição $R[i + 1, n]$ que é necessariamente 0 já que o lowest 1 é em $R[i, n]$.

Implicando em $(rank R_i^j - rank R_{i+1}^j) = 1$.

Analogamente, concluímos que R_i^{j-1} possui uma coluna não nula a mais do que R_{i+1}^{j-1} , assim $(rank R_i^{j-1} - rank R_{i+1}^{j-1}) = -1$.

Então,

$$r_R(i, j) = (rank R_i^j - rank R_{i+1}^j) + (rank R_{i+1}^{j-1} - rank R_i^{j-1}) = +1 - 1 = 0.$$

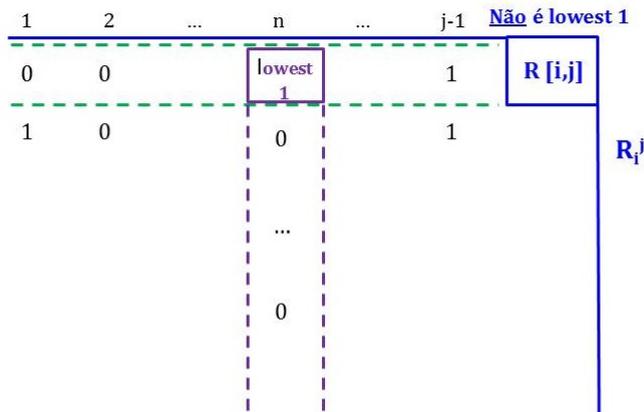


Figura 4.8: $R[i, j]$ não é lowest 1 e uma das colunas de 1 à $j - 1$ possui lowest 1 na linha i .

Desde que os ranks das submatrizes inferiores esquerda de R são os mesmos daqueles de ∂ , nós temos uma caracterização do lowest 1 que não depende do processo de redução.

Teorema 4.2. Lema do Pareamento

Temos que

$$i = \text{low}(j) \text{ se, e somente, se } r_{\partial}(i, j) = 1.$$

Em particular, o pareamento entre linhas e colunas definidas pelo lowest 1 na matriz reduzida não depende de R .

Temos certeza que lowest 1 não é artefato proveniente de uma estratégia particular usada para redução. Perguntamo-nos o que ele realmente significa? Note que a coluna j alcança sua forma final ao término da j -ésima interação.

Nesse momento, temos a matriz reduzida para os complexos consistindo dos j primeiros simplexes na ordenação total. Faremos distinção entre a coluna j que é identicamente nula, da coluna j que possui lowest 1.

Definição 4.6. *Se a coluna j de R for identicamente nula, denominaremos σ_j positivo, porque sua adição cria um novo ciclo e portanto nasce uma classe de homologia.*

Se a coluna j de R for não-nula, logo possui lowest 1, ela armazena o bordo da cadeia acumulada na coluna j da matriz V e é assim um ciclo. Chamamos σ_j negativo, pois sua adição causa a morte de uma classe de homologia.

A classe que morre no caso em que a coluna j possui lowest 1 é representado pela coluna j . Ainda precisamos verificar que a classe nasce no momento em que o simplexo de lowest 1, σ_i , com $i = \text{low}(j)$, é adicionado.

No entanto é claro, porque o ciclo na coluna j de R apenas morreu, bem como todos os demais ciclos que morreram com ele, possuem 1 em linhas acima da linha i , já que caso contrário, se existisse um 1 abaixo da linha i , então esse seria de fato o lowest 1, em outras palavras, poderíamos reduzir a matriz e obter um índice $k = \text{low}(j) > i$, o que contradiz o algoritmo.

Realmente a classe nasce no momento em que adicionamos o simplexo σ_i , com $i = \text{low}(j)$.

Relembramos que $K_i = K(a_i) = \{x \in K \mid f(x) \leq a_i\} = f^{-1}(-\infty, a_i]$ o valor da função avaliado no simplexo.

Segue que o lowest 1 de fato correspondem à pontos no diagrama de persistência. Mais precisamente, (a_i, a_j) é um elemento finito do diagrama de persistência, $Dgm(f)$ se, e somente se, $i = \text{low}(j)$ e, por conseguinte, σ_i é um p -simplexo, acarretando que σ_j é um $(p + 1)$ -simplexo.

Temos que $(a_i, +\infty)$ é um elemento infinito do diagrama de persistência, $Dgm(f)$ se, e somente se, a coluna i for nula e a linha i não tiver lowest 1. Isso significa que σ_i é positivo, pois sua coluna i é nula, contudo não é pareado com um simplexo negativo, porque se fosse ocorreria um lowest 1 na linha i .

Portanto, temos que $(a_i, a_j) \in Dgm(f)$ significa que uma classe de homologia de dimensão p nasce se adicionarmos σ_i e morre quando acrescentamos σ_j .

Exemplo 4.3. Considere o complexo simplicial K , o triângulo com sua face.

Para termos uma filtração, primeiro colocamos todos os vértices, depois as arestas e, finalmente, adicionamos a face do triângulo. Numerando-os de 1 à 7, segundo a ordenação compatível.

Para tornar o exemplo mais interessante acrescentamos um elemento não-nulo de dimensão (-1) do grupo de cadeia reduzida, como um falso simplexo indexado em 0. Assim, calcularemos a homologia reduzida, ao invés da homologia convencional.

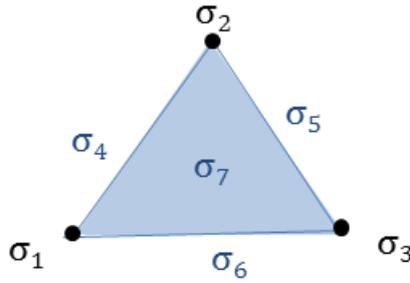


Figura 4.9: O complexo simplicial triângulo.

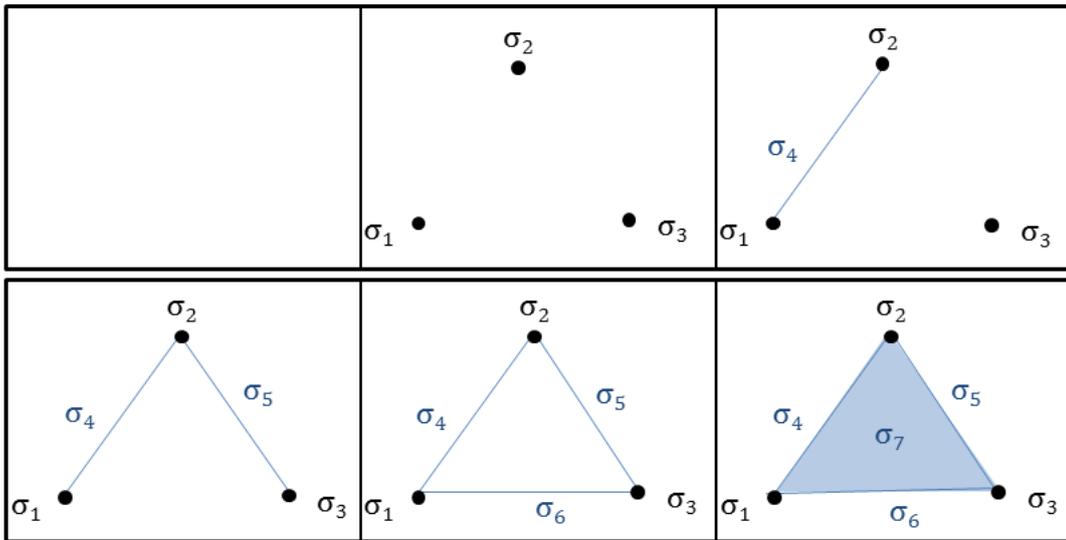


Figura 4.10: A filtração do complexo simplicial triângulo. Inicialmente temos o conjunto vazio, que representa o falso simplexo indexado em 0, após isso adicionamos os vértices, as arestas e finalmente a face, portanto conseguimos o triângulo.

Notemos que o conjunto dos 2-simplexos, ou seja, da face do triângulo K é gerado por σ_7 , $C_2(K) = \langle \sigma_7 \rangle$.

Bem como o conjunto das arestas, 1-simplexo, é gerado por $C_1(K) = \langle \sigma_4, \sigma_5, \sigma_6 \rangle$.

O conjunto dos vértices, 0-simplexos, $C_0(K)$ é gerado por $\langle \sigma_1, \sigma_2, \sigma_3 \rangle$.

E por fim, $\mathbb{Z}_2 = \langle \sigma_0 \rangle$.

Assim,

$$0 \xrightarrow{\partial_3} C_2(K) \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} \mathbb{Z}_2 \xrightarrow{0} 0.$$

Escreveremos a matriz bordo, ∂

$$\partial = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note que podemos tanto analisar a matriz bordo como os simplexes σ_i , faces de codimensão 1 de σ_j quanto como o homomorfismo bordo entre os complexos de cadeia $\partial_i : C_i(K) \rightarrow X_{i-1}(K)$ que associa o σ_j com os simplexes que são seu bordo, σ_i . Ambas as abordagens são equivalentes, visto que fornecem os mesmos resultados, a única diferença é que ao construirmos a matriz bordo ∂ , contornamos aspectos e conceitos mais teóricos da topologia algébrica e a estruturamos de forma mais técnica.

A matriz bordo ∂ nos diz, por exemplo, que o simplexo associado à coluna 4, σ_4 , tem como faces de codimensão 1, os simplexes associados as linhas 1 e 2 que são σ_1 e σ_2 , respectivamente.

Agora, ao aplicarmos o algoritmo de redução matricial, obtemos a matriz reduzida R

$$R = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Exibiremos dois passos do processamentos do algoritmo de redução matricial a fim de ilustrar como se desenvolve a evolução da matriz bordo ∂ para a matriz reduzida R .

Seja o algoritmo de redução matricial:

```

R = ∂
for j = 1 to m do
  while there exists j0 < j with low(j0) = low(j) do
    add column j0 to column j
  endwhile
endfor

```

Para $j = 1$, notamos que temos $j_0 = 0 < j = 1$, sendo $low(0) \neq low(1)$ uma vez que a coluna 0 tem $low(0)$ indefinido por ser identicamente nula.

Já quando $j = 2$, notamos que existe $j_0 = 1 < j = 2$ com $low(1) = 1 = low(2)$, portanto, satisfazemos a condição do laço while e procedemos com a soma da coluna j_0 na coluna j , desde que trabalhamos em \mathbb{Z}_2 , ocorre o anulamento dos termos diferentes de um, e assim atualizamos a coluna 2 com o resultante da soma delas.

Se observarmos quando j alcança o valor 6, percebemos que, fixada a coluna 6, o índice $j_0 = 5$ é tal que ambas colunas possuem mesmo lowest 1, $low(5) = 3 = low(6)$. Logo, adicionamos a coluna 5 na coluna 6, resultando a seguinte configuração

$$\partial = \begin{bmatrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Todavia, agora existe o índice $j_0 = 4$ em que os lowest 1 coincidem, $low(4) = 2 = low(6)$, logo entramos mais uma vez no laço while e somamos a coluna 4 com a coluna 6, tornando-a identicamente nula. Assim, conseguimos a configuração final da matriz ∂ .

Extrairemos as informações arquivadas na matriz reduzida R .

Vamos, primeiramente, calcular os números de Betti do complexo K .

- Armazenamos o (-1) -simplexo, σ_0 , na coluna 0, donde é única e nula.
- Armazenamos os 0-simplexos, os vértices σ_1 , σ_2 e σ_3 nas colunas de 1 à 3, donde temos que a coluna 1 possui lowest 1, enquanto as demais são nulas.
- Armazenamos os 1-simplexos, as arestas σ_4 , σ_5 e σ_6 nas colunas de 4 à 6, donde a única coluna nula é a 6.
- Armazenamos o 2-simplexo, a face σ_7 , na coluna 7, donde temos lowest 1.

Assim calcular o β_0 se resume a computar $z_0 = rank Z_0$ e subtrair de $b_0 = rank B_0$, sendo que essa tarefa se resume à contar a quantidade de linhas nulas de Z_0 e a quantidade de linhas não-nulas de B_0 , assim temos

$$\beta_0 = z_0 - b_0 = 2 - 2 = 0.$$

Analogamente, obtemos:

$$\beta_{-1} = z_{-1} - b_{-1} = 1 - 1 = 0,$$

$$\beta_1 = z_1 - b_1 = 1 - 1 = 0,$$

$$\beta_2 = z_2 - b_2 = 0 - 0 = 0.$$

Além disso, somos aptos à captar ainda mais informações: O nascimento e a morte de classes de homologia.

Os lowest 1 pareados, $i = low(j)$, indicam o nascimento e morte de classes de homologia da seguinte forma: ao adicionarmos o simplexo associado à i , σ_i , anunciamos o nascimento de uma classe de homologia cuja dimensão é a dimensão do simplexo σ_i , enquanto, ao acrescentarmos o simplexo associado à j , σ_j , identificamos a morte da

determinada classe, que nasceu com σ_i . Em outras palavras, σ_i é positivo e σ_j é negativo.

Iniciaremos com a linha, da matriz reduzida R , que representa σ_2 e notamos que $low(4) = 2$ e ainda, ressaltamos que σ_2 é um vértice, portanto, tratamos de informações de classe de homologia de dimensão 0. Assim, temos que uma classe de homologia de dimensão 0 nasce ao adicionarmos σ_2 e morre ao adicionarmos σ_4 . Observe que na Figura 4.10, temos que σ_4 funde σ_1 com σ_2 , logo, uma componente conexa por caminhos morre e elas se tornam uma só.

Já, $6 = low(7)$, ou seja, σ_6 que é uma aresta, dá nascimento à classe de homologia de dimensão 1, um ciclo, que morre quando adicionamos σ_7 que é uma face, assim, preenchendo o buraco.

Procedendo dessa forma, obtemos que

- σ_0 gera o nascimento de uma classe de dimensão (-1) que morre com σ_1 .
- σ_2 gera o nascimento de uma classe de dimensão 0 que morre com σ_4 .
- σ_3 gera o nascimento de uma classe de dimensão 0 que morre com σ_5 .
- σ_6 gera o nascimento de uma classe de dimensão 1 que morre com σ_7 .

Portanto, podemos construir os diagramas de persistência para cada uma dessas dimensões. Além disso, é importante lembrar que, como já mencionamos,

$$(a_i, a_j) \in Dgm(f) \text{ se, e somente se, } i = low(j).$$

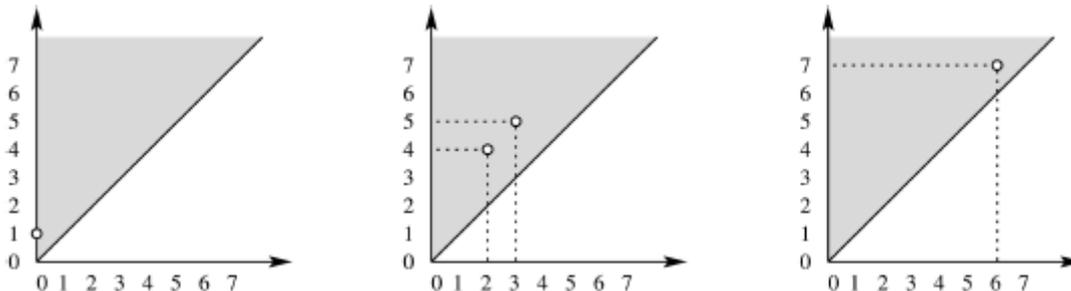


Figura 4.11: Da esquerda para a direita temos ilustrado os diagramas de persistência Dgm_{-1} , Dgm_0 e Dgm_1 .

Finalmente, ressaltamos também que a Regra do Mais Velho é satisfeita.

4.2 Implementação Eficiente

Nas aplicações práticas, o número de simplexes pode ser muito grande, portanto, o armazenamento de uma matriz bordo inteira se torna proibido. Vejamos o exemplo dado no capítulo anterior, com apenas 8 simplexes, obtivemos uma matriz 8×8 , conseqüentemente, foram necessários 64 posições para guardar apenas a matriz bordo.

Como uma alternativa, apresentamos a implementação da matriz esparsa do Algoritmo de Persistência, e entregamos limitantes do tempo de processamento que são melhores do que ordem cúbica, para diversas quantidades de tamanhos de dados de entrada.

4.2.1 Representação da Matriz Esparsa

Bem como no capítulo anterior, assumimos uma função monotônica definida em um complexo simplicial a valores reais, $f : K \rightarrow \mathbb{R}$, juntamente com uma ordenação compatível dos simplexes $\sigma_1, \sigma_2, \dots, \sigma_m$. Armazenamos os dados utilizando um vetor linear $\partial[1..m]$ com uma lista ligada de simplexes por entrada.

A lista em $\partial[j]$ corresponde a j -ésima coluna da matriz bordo, porém guardando apenas os índices dos simplexes que são faces de codimensão 1 com σ_j . Ao final do algoritmo, a lista contida na j -ésima entrada do vetor, corresponde a coluna na matriz reduzida cujo lowest 1 está na j -ésima linha. Se não houver tal coluna, então a lista estará vazia.

Com propósito de enfatizar a mudança, faremos uma transição de nomenclatura do vetor ∂ no início para R ao final do algoritmo.

Todas as listas são classificadas na ordem de índice decrescente, assim o simplexo adicionado mais recentemente estará prontamente disponível no topo.

Nós vemos uma migração geral da lista da direita para a esquerda. A fim de descrever o algoritmo que rege a migração, vamos escrever:

- L a lista ligada à j -ésima entrada do vetor.
- $i = Top(L)$ o índice do simplexo que estiver no topo.
- A i -ésima entrada é dita ocupada se armazenar uma lista não-vazia. Caso contrário, denominaremos como desocupada.

Exemplo 4.4. Com o intuito de ilustrar a diferença de armazenagem da matriz bordo ∂ para o vetor $\partial[0..7]$, vamos reutilizar o último exemplo apresentado na seção anterior.

A informação guardada pelo vetor $\partial[0..7]$ para o simplexo σ_4 seria $\partial[4] = (2, 1)$, onde a ordenação é no sentido decrescente e os índices 2 e 1 significam os simplexes σ_2 e σ_1 . É evidente a economia de espaço, porque deixamos de arquivar os 0's como informação, que sequer útil era. Já a posição $R[6] = \emptyset$, pois ao final do algoritmo a coluna 6 é identicamente nula, assim não salvando e usando espaço algum.

Uma vez esclarecido a forma de guardar as informações do vetor ∂ , apresentamos o algoritmo.

```

R = ∂
for j = 1 to m do
  L = ∂[j].cycle ; R[j].cycle = NULL
  while L ≠ NULL and R[i] with i = Top(L) is occupied do
    L = L + R[i].cycle
  endwhile
  if L ≠ NULL then R[i].cycle = L endif
endfor

```

4.2.2 Análise do Algoritmo

A estrutura principal da Implementação da Matriz Esparsa é que temos dois laços encaixados, o loop interno e o loop externo. O acréscimo de duas listas é outro loop disfarçado, logo a ordem de resposta do algoritmo é, também, no máximo cúbico.

Com a finalidade de melhorar e refinar nossa primeira estimativa, definimos uma colisão como uma tentativa de depositar a lista L que falha pelo fato da entrada estar ocupada.

Para cada colisão, requer-se a fundição de duas listas que leva o tempo proporcional a soma de seus tamanhos.

O loop termina quando L se torna vazia, o que caracteriza σ_j dando o nascimento a uma classe de homologia, ou quando a lista L não-vazia é depositada com sucesso, identificando σ_j como matando, já o simplexo σ_i , onde o depósito ocorre, desencadeando o correspondente nascimento.

Sair do Loop: Se $L = NULL$ significa que $\partial[j].cycle \equiv 0$ pois $L = \partial[j].cycle$ para algum $1 \leq j \leq m$, ou seja, temos uma coluna identicamente nula, assim dizemos que σ_j dá nascimento a uma classe de homologia. Ou ainda, se $L \neq NULL$ é depositada com sucesso, $L = \partial[j].cycle$ e $R[i]$ ocupada, onde $i = Top(L)$, assim $L = L + R[i].cycle$ mas $L = \partial[j].cycle$, segue que $L = \partial[j].cycle + R[i].cycle$. Desde que $L \neq NULL$, daí $R[i].cycle = \partial[j].cycle + R[i].cycle = L$, ou seja, na lista $L = \partial[j].cycle$ temos outra coluna, digamos $R[i]$, ocupada com um top simplexo igual ao de L , o que equivale a dizer, segundo o ponto de vista do algoritmo da matriz bordo ∂ que temos dois lowest 1 na mesma linha, assim ocorre a soma da lista L com $R[i]$ e armazenamos na posição $R[i]$ que agora possui 0 na posição $i = Top(L)$, portanto, diremos que σ_j causa a morte de uma classe de homologia enquanto σ_i gera o nascimento.

Cada lista $R[k].cycle$ contém σ_k como maior top simplexo.

Similarmente, σ_k é o maior top simplexo na lista L se esse colide com a lista em $R[k]$.

Como usamos a aritimética módulo 2, σ_k é excluído, implica que o top simplexo na lista fundida possui índice menor que k .

O loop interno assim procede monotonicamente da direita para a esquerda. Segue que a colisão com um simplexo σ_j acontece apenas nas entradas entre i e j .

Considere agora o loop interno para σ_j . Uma colisão na entrada k pode ocorrer somente se σ_k deu nascimento a uma classe que morre em σ_l , $l < k$, em outras palavras, σ_l vem antes de σ_k .

Analogamente, as colisões durante o loop interno para σ_l corresponde ao par nascimento-morte dentro de $[k, l]$.

Indutivamente, isso implica que listas adicionadas nas colisões contém apenas faces de simplexos com índice $[i, j]$.

Definindo p como a dimensão de σ_j , o número de tais faces é, no máximo, $p + 1$, isso vezes o número de índices no intervalo. O tempo para fundir duas listas é, portanto, no máximo proporcional a esse número. Em suma, o tempo de processamento do loop interno para um p -simplexo, σ_j , é no máximo $(p + 1)(j - i)^2$.

4.2.3 Implementação da Matriz Esparsa

A matriz bordo ∂ é esparsa, por definição. Já a matriz reduzida pode ser densa, ainda que raramente seja, assim usando a estrutura da matriz de dados esparsa podemos obter significantes ganhos de eficiência. Descrevemos a particular implementação.

A estrutura de dados consiste de um vetor linear $\partial[1..m]$ armazenando uma lista associada a cada simplexo. Inicialmente todas as listas são vazias e pelo menos metade delas permanecem dessa forma.

Cada lista guarda um ciclo, mais especificamente, os índices das linhas cuja entrada é não-nula na coluna da matriz de bordo.

Cada lista é classificada com o maior índice da linha, prontamente disponível no topo.

Somar duas listas significa fundi-las e remover os índices em duplicidade. Uma vez que as listas são classificadas, isso leva o tempo linear no comprimento das duas listas.

Armazenar a lista que representa a coluna j em $\partial[i]$ onde fazendo uma analogia com a matriz bordo, $i = low(j)$. Quando o $low(j)$ se modifica, nós movemos a lista, e uma vez que ela apenas pode decrescer, a lista irá movimentar monotonamente para a esquerda.

Seja i o índice de topo na lista L , encontramos dois casos

1. Colisão: $\partial[i]$ é não-vazia e adicionamos $\partial[i]$ a L .
2. Chegada: $\partial[i]$ é vazia e definimos $\partial[i] = L$.

Após adicionar $\partial[i]$ no caso de colisão, o índice de topo da lista L é menor do que antes, assim continuamos a pesquisa. Também é possível que adicionando $\partial[i]$, L se torne vazia, nesse caso, acabamos a pesquisa mas marcamos o simplexo σ_j do qual iniciamos o procedimento como positivo. Se L não se tornar vazia e eventualmente armazenada em algum $\partial[k]$, marcamos σ_j como negativo e pareamos σ_k com σ_j . O número de vezes que a lista se movimenta até alcançar σ_k é, no máximo, $(j - k)$ a persistência do par.

Exemplo 4.5. Vamos descrever o passo-a-passo do procedimento do algoritmo.

Considere o complexo simplicial:

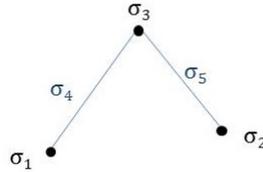


Figura 4.12: Complexo simplicial.

Temos, portanto, que $\partial[j] = NULL$, para $1 \leq j \leq 3$. Já $\partial[4] = (3, 1)$ e $\partial[5] = (3, 2)$.

Para $1 \leq j \leq 3$ temos que $L_j = \partial[j].cycle = NULL$ e definimos $R[j].cycle = NULL$. Como $L_j = NULL$ não acessamos a condição do if também.

Agora, $j = 4$ temos $L_4 = \partial[4].cycle$, onde $3 = i = Top(L_4)$, definimos $R[4].cycle = NULL$ e assim acionamos o loop interno. No entanto, nesse momento, a segunda condição falha, já que $R[3]$ é classificada como desocupada (unoccupied). Logo saltamos do loop interno para a condicional que diz que se $L_4 \neq NULL$ então $R[3].cycle = L_4$. Então temos $R[3] = (3, 1) = L_4$ e $R[k] = \emptyset$, para $k = 1, 2, 4$.

Na última interação, $j = 5$, temos $L_5 = \partial[5].cycle = (3, 2)$ e definimos $R[5].cycle = NULL$. Observemos que $3 = i = Top(L_5)$. Segue que entraremos no loop interno, uma vez que $L_5 = (3, 2) \neq NULL$ e $R[3]$ está ocupada (occupied), por armazenar uma lista não-vazia. Então:

$$L_5 = L_5 + R[3].cycle = \partial[5].cycle + R[3].cycle = (3, 2) + (3, 1) = (0, 2, 1) = (2, 1).$$

Obtemos o seguinte formato para $\partial[5] = (2, 1)$ após o processamento do loop interno. Mais uma vez executaremos o loop interno, pois nos encontramos nas condições exigidas. Agora temos a versão atualizada de $\partial[5] = (2, 1) \neq NULL$, só que nesse instante $i = Top(L_5) = 2$, entretanto, $R[2]$ está desocupada e pulamos do loop interno para a condicional que se $L_5 \neq NULL$, então $R[2].cycle = L_5$, logo $R[2] = (2, 1)$. Obtemos R tal que $R[i] = NULL$, para $i = 1, 4, 5$. Já $R[2] = (2, 1)$ e $R[3] = (3, 1)$. E o algoritmo termina.

Precisamos lê-lo e extrair as informações codificadas.

Sabemos que a lista de $R[j]$ corresponde à matriz reduzida da qual o lowest 1 se encontra na j -ésima linha. Ainda temos que os números armazenados em $R[j].cycle$ significam aonde devemos anexar o número 1 na posição do número guardado.

Primeira observação é que $R[2]$, de fato, arquivou o índice 2 como seu top simplexo, bem como $R[3]$ guardou 3 como seu top simplexo. As demais posições são vazias.

Além disso $R[2].cycle = (2, 1)$ significa que σ_5 tem como vértices σ_2 e σ_1 . Isso ocorre porque $R[2]$ aponta para $\partial[5]$ que representa o bordo de σ_5 . Analogamente, $R[3]$ nos diz que σ_4 , pois $R[3]$ aponta para $\partial[4]$ o bordo de σ_4 , tem como vértices σ_3 e σ_1 .

Escrevendo a matriz de bordo da maneira usual temos:

$$\partial = \begin{bmatrix} & \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 \\ \sigma_1 & 0 & 0 & 0 & 1 & 1 \\ \sigma_2 & 0 & 0 & 0 & 0 & 1 \\ \sigma_3 & 0 & 0 & 0 & 1 & 0 \\ \sigma_4 & 0 & 0 & 0 & 0 & 0 \\ \sigma_5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Notemos ainda que $\partial(\sigma_4) = \sigma_3 + \sigma_1$ e $\partial(\sigma_5) = \sigma_2 + \sigma_1$. Segundo a Regra do Mais Velho, temos que a informação codificada pela matriz é coerente e compatível com o processo estabelecido.

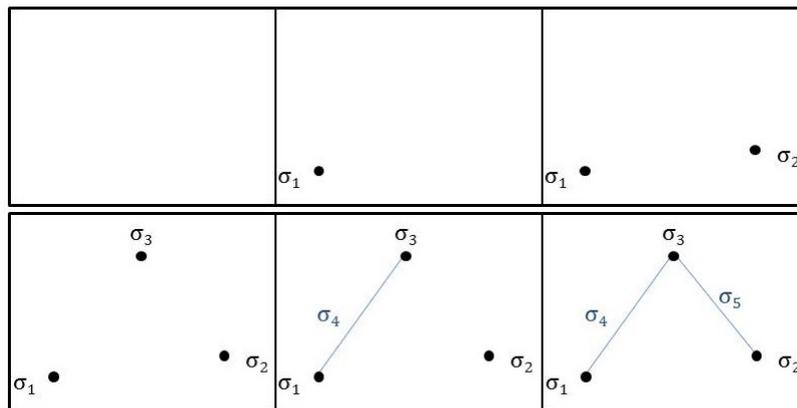


Figura 4.13: As filtrações do complexo simplicial e a Regra do Mais Velho.

4.3 O 0-ésimo Diagrama

A estrutura da lista usada para computar o 0-ésimo diagrama de persistência é mais simples do que para dimensões acima de zero. Esse diagrama depende exclusivamente

dos vértices e arestas de K e de sua sequência de ordenação compatível. Um vértice não possui bordo, logo sempre gera o nascimento de uma componente, aqui, portanto, não há escolhas!

Já uma aresta σ_j possui 2 vértices e seu bordo é $\partial(\sigma_j) = u + w$, onde $u = \sigma_i$ e $w = \sigma_k$ e suponha que u venha primeiro, isto é, $i < k$.

O primeiro passo do algoritmo é então a tentativa de depositar a lista L consistindo de u e w em $R[k]$.

Se $R[k].cycle = L_k$ é vazia, então o depósito é realizado com sucesso, σ_k e σ_j formam um par e o loop interno termina. Caso contrário, L_k é uma lista de dois vértices v e k , do qual $v = \sigma_l$ vem primeiro. Somando as duas listas obtemos $L + L_k$ que consiste de u e v .

Assim, toda lista não-vazia tem comprimento 2, logo cada adição leva um tempo constante. Implicando que o esforço total para dimensão 0 é, no máximo, a soma dos índices, para vértices que geram nascimento e, no máximo, a soma dos índices de persistência para vértices que geram a morte. Em ambos os casos, é limitado por m^2 .

Entretanto, conseguimos fazer ainda melhor.

Considere novamente os dois casos da aresta com bordo $\partial(\sigma_j) = u + w$ que gera o nascimento, se e somente se, u e w pertencerem a mesma componente de K_{j-1} que é o complexo antes de adicionarmos σ_j .

Começando com σ_j , o algoritmo adiciona uma aresta ao caminho crescente em cada colisão, L mantém controle de seu bordo, os dois vértices são os pontos extremos do caminho. Eventualmente, os dois vértices extremos se juntam e L se torna vazia, o caminho se transforma em 1-ciclo.

A aresta σ_j gera morte se, e somente se, u e w pertencem a duas componentes diferentes de K_{j-1} .

Notamos que a lista $R[j].cycle$ se tornará nula e assim temos uma lista vazia, ou seja, dizemos que σ_j dá nascimento a uma classe de homologia.

O loop interno termina quando um dos finais do caminho crescente alcança o vértice mais velho σ_l , de uma componente. Uma vez que o loop trabalha monotonamente da direita para esquerda, isso implica que o vértice mais velho da outra componente é ainda mais velho. Seguindo a Regra do Mais Velho, L se deposita em $R[l]$ e σ_l, σ_j formam um par.

Note que o resultado é previsível. Tudo que precisamos saber é se os vértices u e w pertencem ou não à componentes diferentes de K_{j-1} , e caso façam, qual deles tem o vértice mais velho dessas componentes.

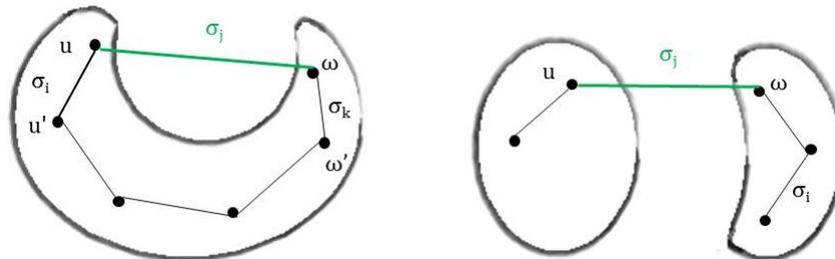


Figura 4.14: Os dois casos estudados para o 0-ésimo diagrama.

4.4 Superfícies

Consideramos um complexo simplicial K que seja uma triangulação para uma variedade de dimensão 2. Esse cenário é de grande importância prática e nos permite uma rápida implementação do Algoritmo de Persistência.

Seja $f : |K| \rightarrow \mathbb{R}$ uma função linear por partes obtida por meio da interpolação dos valores dos vértices.

Possivelmente, existe informação não-trivial no 0-ésimo e Primeiro Diagrama de Persistência de f , mas em nenhum outro diagrama.

Para computar esses dois diagramas rapidamente, devemos responder duas questões:

1. Como podemos transformar uma família de 1 parâmetro de conjuntos de subnível em uma filtração que possamos executar com nosso algoritmo?
2. Como podemos melhorar o tempo de execução mais lento do Primeiro Diagrama de Persistência para aproximadamente o tempo do 0-ésimo Diagrama?

Lidaremos com a primeira questão agora e deixaremos a segunda para mais tarde.

Assumimos, para simplificar, que a função f restrita aos vértices de K seja injetora, isto é, vértices distintos possuem valores avaliados pela f diferentes.

Consideremos a filtração do Lower Star donde conseguimos a sequência encaixante $\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K$, onde K_i é a união dos lower stars dos primeiros i vértices segundo a ordenação compatível de f . Essa também é a filtração que conseguimos através da função monotônica $g : K \rightarrow \mathbb{R}$ definida por $g(\sigma) = \max_{x \in \sigma} f(x)$.

O diagrama de f é definido pelos grupos de homologia dos subníveis de f , $|K_a| = f^{-1}(-\infty, a]$, enquanto os de g são definidos pelos grupos de homologia de g , $K_a = g^{-1}(-\infty, a]$.

Pela definição da filtração do Lower Star, temos $|K_a| \subseteq |K|_a$ e a inclusão é equivalência homotópica.

Segue que os diagramas são isomorfismos

$$\begin{array}{ccc} H_p(K_a) & \longrightarrow & H_p(K_b) \\ \downarrow & & \downarrow \\ H_p(|K|_a) & \longrightarrow & H_p(|K|_b). \end{array}$$

O diagrama comuta, porque todos os quatro homomorfismos são induzidos da inclusão.

Apresentaremos o segundo resultado fundamental da teoria de Persistência que garante que dados dois módulos de persistência \mathbb{U} e \mathbb{V} , tais que exista um isomorfismo $\phi : \mathbb{U} \rightarrow \mathbb{V}$, então os diagramas de persistência são os mesmos.

Teorema 4.3. Teorema de Equivalência da Persistência

Considere duas sequências de espaços vetoriais conectadas por homomorfismo $\phi_i : U_i \rightarrow V_i$

$$\begin{array}{ccccccc} U_0 & \longrightarrow & U_1 & \longrightarrow & \dots & \longrightarrow & U_n \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ V_0 & \longrightarrow & V_1 & \longrightarrow & \dots & \longrightarrow & V_n \end{array}$$

Se para todo i , ϕ_i for isomorfismo e todos os quadrados comutarem, então o diagrama de persistência definido por U_i é o mesmo do definido por V_i .

Demonstração. Supondo que γ é uma classe de homologia que nasce em U_{i+1} , segue por definição que $\gamma \notin U_p^{i,i+1}$, em outras palavras, temos uma caracterização de que não existe $\alpha \in U_i$, tal que $f_p^{i,i+1}(\alpha) = \gamma$ se, e somente se, $f_p^{i,i+1}(\alpha) - \gamma \neq 0$, para todo $\alpha \in U_i$.

Sabemos por hipótese que ϕ_{i+1} é isomorfismo, portanto, consideramos $\phi_{i+1}(\gamma) \in V_{i+1}$. Vamos mostrar que $\phi_{i+1}(\gamma)$ nasce em V_{i+1} .

Assumamos, por absurdo, que não, isto é, que $\phi_{i+1}(\gamma)$ não nasce em V_{i+1} . Assim, existe $\delta \in V_i$, tal que $g_p^{i,i+1}(\delta) = \phi_{i+1}(\gamma)$.

Uma vez que ϕ_i é isomorfismo, podemos considerar $\phi_i^{-1}(\delta) \in U_i$. Logo,

$$f_p^{i,i+1}(\phi_i^{-1}(\delta)) \in U_{i+1},$$

ainda mais

$$\phi_{i+1}(f_p^{i,i+1}(\phi_i^{-1}(\delta))) = g_p^{i,i+1}(\delta),$$

mas,

$$g_p^{i,i+1}(\delta) = \phi_{i+1}(\gamma).$$

Logo,

$$\phi_{i+1}(f_p^{i,i+1}(\phi_i^{-1}(\delta))) = g_p^{i,i+1}(\delta) = \phi_{i+1}(\gamma).$$

Consequentemente, pela injetividade de ϕ_{i+1} , temos

$$f_p^{i,i+1}(\phi_i^{-1}(\delta)) = \gamma.$$

Assim, alcançamos uma contradição. Portanto, ambas as classes nascem ao mesmo tempo.

Agora, vamos mostrar que se a classe γ que nasce em U_{i+1} morre ao entrar em U_{j+1} , então também $\phi_{i+1}(\gamma)$ morre ao entrar em V_{j+1} .

Recordemos que γ morrer ao entrar em U_{j+1} significa que $f_p^{i+1,j}(\gamma) \notin U_p^{i,j}$, mas $f_p^{i+1,j+1}(\gamma) \in U_p^{i,j+1}$. Assim, segue que existe $\Delta \in U_i$, tal que $f_p^{i,j+1}(\Delta) = f_p^{i+1,j+1}(\gamma)$. Para provar que $\phi_{i+1}(\gamma)$ morre ao entrar em V_{j+1} , suponha, por contradição, que não, assumindo que $\phi_{i+1}(\gamma)$ morre ao entrar em V_{k+1} , com $k+1 > j+1$. Assim, $f_p^{i,j+1}(\Delta) = f_p^{i+1,j+1}(\gamma)$, por conseguinte consideremos, $\phi_{j+1}(f_p^{i+1,j+1}(\gamma)) = \phi_{j+1}(f_p^{i,j+1}(\Delta)) \in V_{j+1}$.

Agora, se considerarmos $\gamma \in U_{i+1}$, temos que

$$\phi_{j+1}(f_p^{i+1,j+1}(\gamma)) = g_p^{i+1,j+1}(\phi_{i+1}(\gamma)).$$

Similarmente, para $\Delta \in U_i$, conseguimos

$$\phi_{j+1}(f_p^{i,j+1}(\Delta)) = g_p^{i,j+1}(\phi_i(\Delta)).$$

Estamos supondo que $\phi_{i+1}(\gamma)$ morre ao entrar em V_{k+1} , com $k+1 > j+1$. Assim, $g_p^{i+1,k}(\phi_{i+1}(\gamma)) \notin V_p^{i,k}$. Contudo, se observarmos, temos

$$g_p^{i+1,k}(\phi_{i+1}(\gamma)) = g_p^{j+1,k} g_p^{i+1,j+1}(\phi_{i+1}(\gamma)).$$

Mas,

$$g_p^{i+1,j+1}(\phi_{i+1}(\gamma)) = \phi_{j+1}(f_p^{i+1,j+1}(\gamma)) = \phi_{j+1}(f_p^{i,j+1}(\Delta)) = g_p^{i,j+1}(\phi_i(\Delta)).$$

Logo,

$$g_p^{i+1,k}(\phi_{i+1}(\gamma)) = g_p^{j+1,k}(g_p^{i+1,j+1}(\phi_{i+1}(\gamma))) = g_p^{j+1,k}(g_p^{i,j+1}(\phi_i(\Delta))) = g_p^{i,k}(\phi_i(\Delta)),$$

ou seja, $g_p^{i+1,k}(\phi_{i+1}(\gamma)) \in V_p^{i,k}$, contradição!

Repetindo o processo, mas agora supondo $(k+1) < j$, chegamos no mesmo absurdo.

Então, $\phi_{i+1}(\gamma)$ morre ao mesmo tempo que a classe γ .

Concluimos que, se (a_{i+1}, a_{j+1}) for um elemento do diagrama de persistência definido por U , significa que existe uma classe de homologia γ que nasce em U_{i+1} e morre ao entrar em U_{j+1} , por conseguinte, temos que existe uma classe $\phi_{i+1}(\gamma)$ que nasce em V_{i+1} e morre ao entrar em V_{j+1} , portanto, o elemento (a_{i+1}, a_{j+1}) pertence ao diagrama de persistência definido por V .

Assim, os diagramas de persistência definidos por U e por V são os mesmos. \square

4.5 O Primeiro Diagrama

Ao invés de computar o Primeiro Diagrama de Persistência de f diretamente, vamos construir o 0-ésimo diagrama de persistência de $(-f)$ e derivar o diagrama de f a partir dele.

Iniciamos através da descrição da relação entre $Dgm_1(f)$ e $Dgm_0(-f)$, omitindo as provas, uma vez que as relações são consequências de teoremas mais gerais que serão dados na próxima seção.

O Primeiro Diagrama de Persistência de f consiste da diagonal e uma porção finita de pontos fora da diagonal (a, b) . Construiremos a porção finita através do 0-ésimo diagrama de $(-f)$. Especificamente, o ponto (a, b) representa o nascimento de uma classe de homologia de dimensão 1 em a e que morre em b .

Olhar para $(-f)$ é como tomar o complemento e ir para trás. Portanto, temos o nascimento de uma classe de homologia 0-dimensional em $-b$ e sua morte em $-a$. Segue que:

$$(a, b) \in Dgm_1(f) \iff (-b, -a) \in Dgm_0(-f).$$

Em outras palavras, a porção finita de $Dgm_1(f)$ pode ser obtida pela reflexão dos pontos da porção finita de $Dgm_0(-f)$ através da diagonal menor.

E uma vez que já estudamos como construímos o 0-ésimo diagrama, temos um ganho considerável de tempo para captar informações na porção finita do Primeiro Diagrama.

4.6 Persistência Estendida

4.6.1 Teorema da Dualidade de Lefschetz

Definição 4.7. *O Star de τ consiste de todas as cofaces de τ .*

$$St\tau = \{\sigma \in K \mid \tau \leq \sigma\}.$$

Geralmente, o Star não é fechado segundo suas faces. Podemos torná-lo um complexo adicionando as faces que faltam e obtemos o que resulta no Star Fechado \bar{St} , que é o menor complexo que contém o Star de τ . O link consiste de todos os simplexes no Star Fechado de τ que são disjuntos de τ :

$$Lk\tau = \{v \in \tilde{S}t\tau \mid v \cap \tau = \emptyset\}.$$

Definição 4.8. Uma variedade de dimensão d combinatória é uma variedade de dimensão d munida de uma triangulação, tal que o link de todo i -simplexo triangula a esfera de dimensão $(d - i - 1)$.

Teorema 4.4. Primeiro Teorema da Dualidade de Lefschetz

Seja \mathbb{M} uma variedade de dimensão d , compacta e combinatória, com fronteira $\partial\mathbb{M}$. Então para cada par de dimensões complementares $p+q = d$, existe um isomorfismo

$$H_p(\mathbb{M}, \partial\mathbb{M}) \approx H^q(\mathbb{M}),$$

$$H_p(\mathbb{M}) \approx H^q(\mathbb{M}, \partial\mathbb{M}).$$

Teorema 4.5. Segundo Teorema da Dualidade de Lefschetz

Seja \mathbb{M} uma variedade de dimensão d , compacta e combinatória com fronteira $\partial\mathbb{M}$. Então, a aplicação interseção de classes de homologia

$$\# : H_p(\mathbb{M}) \times H_q(\mathbb{M}, \partial\mathbb{M}) \rightarrow \mathbf{G}$$

é perfeita para todo $p + q = d$.

4.7 Elevação em uma superfície

Daremos um breve esboço da abordagem de “formas encaixantes” e apresentaremos uma descrição mais detalhada nos próximos capítulos.

Seja \mathbb{M} uma variedade suave de dimensão 2, imersa no \mathbb{R}^3 .

Dada uma direção $u \in \mathbb{S}^2$, a função altura nessa direção é a função $f : \mathbb{M} \rightarrow \mathbb{R}$, dada por $f(x) = \langle x, u \rangle$, para todo $x \in \mathbb{M}$.

Usualmente desenhamos u verticalmente apontando para cima e pensando na altura como a distância da base plana horizontal com o sinal.

Dado um limitante $a \in \mathbb{R}$, relembramos que o subnível consiste de todos os pontos com altura a ou menor, $\mathbb{M}_a = f^{-1}(-\infty, a]$, definindo assim a sequência dos grupos de homologia.

Para uma variedade suave genérica, os valores críticos homológicos da função altura são os valores altura avaliados em pontos críticos isolados.

Se, além disso, a direção for genérica, então haverá apenas três tipos diferentes de valores críticos homológicos:

1. Mínimo: no qual se inicia componentes.
2. Sela: que funde componentes ou completa loops.
3. Máximo: que preenche espaços.

Assumindo que os pontos críticos possuam alturas distintas, os elementos no diagrama de persistência de f correspondem a pares de pontos críticos.

Seja (x, y) elemento no diagrama de persistência de f .

Então, a elevação nos pontos x e y é definida por $|f(x) - f(y)|$.

Desde que x é ponto crítico, para duas direções opostas, nós precisamos ter certeza que o pareamento é o mesmo em ambas direções, senão contradiríamos a atribuição de elevação. Também necessitamos que todos os pontos críticos estejam pareados, caso contrário, obteríamos áreas brancas cuja elevação permaneceria indefinida. Graças a isso é que a estendemos a persistência e a primeira é a restrição que precisamos observar nessa extensão.

4.8 Filtração Estendida

Sejam $a_1 < a_2 < \dots < a_n$ valores críticos homológicos da função altura $f : \mathbb{M} \rightarrow \mathbb{R}$, intercalados com os valores

$$b_0 < a_1 < b_1 < a_2 < b_2 < \dots < b_{m-1} < a_m < b_m.$$

Conseguimos os conjuntos de subníveis $\mathbb{M}_i = f^{-1}(-\infty, b_i]$, cada qual uma variedade de dimensão 2 com fronteira.

Simetricamente, definimos o conjunto de supernível

$$\mathbb{M}^i = f^{-1}[b_i, +\infty),$$

variedade de dimensão 2, complementar a \mathbb{M}_i e com a mesma fronteira.

Estendemos a persistência para homologia essencial utilizando a Dualidade de Poincaré que fornece um isomorfismo entre a dimensão r do grupo de homologia com a dimensão $(d-r)$ do grupo de cohomologia, segundo uma variedade compacta de dimensão d ,

$$H_r(\mathbb{M}) \approx H^{d-r}(\mathbb{M}).$$

As inclusões $\mathbb{M}_k \subseteq \mathbb{M}_{k+1}$ induzem o homomorfismo $H^{d-r}(\mathbb{M}_{k+1}) \rightarrow H^{d-r}(\mathbb{M}_k)$ e, assim, obtemos a sequência estendida

$$0 = H_r(\mathbb{M}_0) \rightarrow \dots \rightarrow H_r(\mathbb{M}_m) = H_r(\mathbb{M}) \simeq H^{d-r}(\mathbb{M});$$

$$H_r(\mathbb{M}) \simeq H^{d-r}(\mathbb{M}) = H^{d-r}(\mathbb{M}_m) \rightarrow \dots \rightarrow H^{d-r}(\mathbb{M}_0) = 0.$$

Classes nascem e morrem também durante a nova segunda metade da sequência. Em particular, cada classe que nasce em \mathbb{M}_k e persiste viva toda a primeira metade da sequência, morrerá ao entrar em algum \mathbb{M}_l na segunda metade. Uma vez que vamos do grupo trivial para o grupo trivial, tudo que nasce, morre. Como consequência, todos os nascimentos são pareados com a correspondente morte, como desejávamos.

O único problema com essa definição é que para fazer sentido, requer que entendamos cohomologia persistente. Isso certamente pode ser feito, mas vamos, ao invés, utilizar a Dualidade de Lefschetz para trocar cohomologia por homologia relativa, já que nos é garantido um isomorfismo:

$$H^{d-r}(\mathbb{M}_k) \approx H_r(\mathbb{M}_k, \partial\mathbb{M}_k).$$

Definido $\mathbb{M}^{m-k} = f^{-1}[b_k, +\infty)$ o supernível contendo os últimos $(m - k)$ pontos críticos, usamos a excisão para adquirir um isomorfismo

$$H_r(\mathbb{M}_k, \partial\mathbb{M}_k) \approx H_r(\mathbb{M}, \mathbb{M}^m)$$

e outra versão da sequência acima seria

$$\begin{aligned} 0 &= H_r(\mathbb{M}_0) \rightarrow \cdots \rightarrow H_r(\mathbb{M}_m) = H_r(\mathbb{M}) \\ &\rightarrow H_r(\mathbb{M}) = H_r(\mathbb{M}, \mathbb{M}^0) \rightarrow \cdots \rightarrow H_r(\mathbb{M}, \mathbb{M}^m) = 0, \end{aligned}$$

onde $\mathbb{M}_0 = \emptyset = \mathbb{M}^0$ e $\mathbb{M}_m = \mathbb{M} = \mathbb{M}^m$.

Trabalhando nessa sequência, podemos entender a persistência estendida da seguinte maneira: Suponha que passando de \mathbb{M}_{k-1} para \mathbb{M}_k gere o nascimento de uma classe de homologia de dimensão r essencial. Descendo pelos conjuntos de supernível, olhamos para o primeiro \mathbb{M}^{m-l} , com o maior l possível, que contenha uma classe homóloga àquela classe essencial, em \mathbb{M} . Então essa classe morre ao entrar em \mathbb{M}^{m-l} e pareamos p_k com p_l . Definimos a persistência estendida desse par como $|k - l|$ ou $|f(p_k) - f(p_l)|$, dependendo da aplicação.

Existe um conjunto de regras que descrevem como é regido os nascimentos e mortes na persistência estendida, assim facilitando a retirada de informação sobre as classes.

- **Regra 1.** Uma classe de homologia de dimensão p de \mathbb{M}^b morre ao mesmo tempo que uma classe de homologia relativa de dimensão $(p + 1)$ de $(\mathbb{M}, \mathbb{M}^b)$ morre.

$$\begin{array}{ccccccc} H_p(\mathbb{M}) & \xrightarrow{id} & H_p(\mathbb{M}) & \xrightarrow{id} & \cdots & \xrightarrow{id} & H_p(\mathbb{M}) \\ \xi^{b_n} \downarrow & & \xi^{b_{n-1}} \downarrow & & & & \xi^{b_k} \downarrow \\ H_{p+1}(\mathbb{M}, \mathbb{M}^{b_n}) & \xrightarrow{z^{b_n}} & H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{n-1}}) & \xrightarrow{z^{b_{n-1}}} & \cdots & \longrightarrow & H_{p+1}(\mathbb{M}, \mathbb{M}^{b_k}) \\ \partial \downarrow & & \partial \downarrow & & & & \partial \downarrow \\ H_p(\mathbb{M}^{b_n}) & \xrightarrow{h^{b_n}} & H_p(\mathbb{M}^{b_{n-1}}) & \xrightarrow{h^{b_{n-1}}} & \cdots & \longrightarrow & H_p(\mathbb{M}^{b_k}) \\ (i^{b_n})^* \downarrow & & (i^{b_{n-1}})^* \downarrow & & & & (i^{b_k})^* \downarrow \\ H_p(\mathbb{M}) & \xrightarrow{id} & H_p(\mathbb{M}) & \xrightarrow{id} & \cdots & \longrightarrow & H_p(\mathbb{M}) \end{array}$$

As colunas são sequências exatas longa do par e temos ∂ o operador bordo da sequência exata longa do par. Além disso, $h_k^l : \mathbb{M}^k \rightarrow \mathbb{M}^l$ é induzido pela inclusão e também $z_k^l : (\mathbb{M}, \mathbb{M}^k) \rightarrow (\mathbb{M}, \mathbb{M}^l)$ é induzido pela inclusão.

Dada uma classe $\gamma \in H_{p+1}(\mathbb{M}, \mathbb{M}^{b_i})$, dizemos que ela nasce em $(\mathbb{M}, \mathbb{M}^{b_i})$ se $\gamma \notin (z_{b_{i+1}}^{b_i})^*(H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{i+1}}))$. Dada uma classe γ que nasceu em $(\mathbb{M}, \mathbb{M}^{b_i})$, dizemos que ela morre ao entrar em $(\mathbb{M}, \mathbb{M}^{b_j})$ se $(z_{b_i}^{b_{j-1}})^*(\gamma) \notin (z_{b_{i+1}}^{b_{j-1}})^*(H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{i+1}}))$, mas $(z_{b_i}^{b_j})^*(\gamma) \in (z_{b_{i+1}}^{b_j})^*(H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{i+1}}))$.

Considere $\partial\gamma \in H_p(\mathbb{M}^{b_i})$.

Mostraremos que $\partial\gamma$ morre em \mathbb{M}^{b_j} .

Como γ morre em $(\mathbb{M}, \mathbb{M}^{b_j})$,

$$\exists \alpha \in H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{i+1}}), \quad \text{tal que } (z_{b_i}^{b_j})^*(\gamma) = (z_{b_{i+1}}^{b_j})^*(\alpha).$$

Logo, $(h_{b_i}^{b_j})^*(\partial\gamma) = (h_{b_{i+1}}^{b_j})^*(\partial\alpha)$.

Assim, para garantir que $\partial\gamma$ morreu ao entrar em \mathbb{M}^{b_j} basta garantir que

$$(h_{b_i}^{b_{j+1}})^*(\partial\gamma) \notin (z_{b_{i+1}}^{b_{j+1}})^*(H_p(\mathbb{M}^{b_{i+1}})).$$

Suponha, por contradição, $\beta \in H_p(\mathbb{M}^{b_{i+1}})$, tal que $(h_{b_i}^{b_{j+1}})^*(\partial\gamma) = (h_{b_{i+1}}^{b_{j+1}})^*(\beta)$.

Isto implica que $(i^{b_{j+1}})^*(h_{b_i}^{b_{j+1}})^*(\partial\gamma) = (i^{b_{j+1}})^*(h_{b_{i+1}}^{b_{j+1}})^*(\beta)$.

Mas pela comutatividade, conseguimos $(i^{b_{j+1}})^*\partial(z_{b_i}^{b_{j+1}})^*(\gamma)$ e, ainda, a sequência longa do par é exata, assim $i \circ \partial = 0$.

$(i^{b_{j+1}})^*(h_{b_{i+1}}^{b_{j+1}})^*(\beta) = 0$, assim $(i^{b_{j+1}})^*(\beta) = 0$.

Uma vez que a sequência é exata, $\text{Im } \partial \subseteq \ker(i^{b_{j+1}})^*$, logo $\beta = \partial\beta'$.

Temos

$$\partial(z_{b_{i+1}}^{b_{j+1}})^*(\beta') = \partial(z_{b_i}^{b_{j+1}})^*(\gamma),$$

e portanto,

$$\partial[-(z_{b_{i+1}}^{b_{j+1}})^*(\beta') + (z_{b_i}^{b_{j+1}})^*(\gamma)] = 0.$$

Então,

$$[(z_{b_i}^{b_{j+1}})^*(\gamma) - (z_{b_{i+1}}^{b_{j+1}})^*(\beta')] \in \ker \partial = \text{Im } \xi^{b_{j+1}},$$

daí $(z_{b_i}^{b_{j+1}})^*(\gamma) = (z_{b_{i+1}}^{b_{j+1}})^*(\beta') + \xi^{b_{j+1}}(\Delta)$, para algum $\Delta \in H_{p+1}(\mathbb{M})$.

Mas, $\xi^{b_{j+1}}(\Delta) = (z_{b_{i+1}}^{b_{j+1}})^*(\xi^{b_{i+1}}(\Delta))$.

Portanto, $(z_{b_i}^{b_{j+1}})^*(\gamma) = (z_{b_{i+1}}^{b_{j+1}})^*(\beta' + \xi^{b_{i+1}}(\Delta))$, contradição.

- **Regra 2.** Uma classe de homologia inessencial de dimensão p de \mathbb{M}^b nasce ao mesmo tempo que nasce uma classe de homologia relativa de dimensão $(p+1)$ de $(\mathbb{M}, \mathbb{M}^b)$

Seja $\Gamma \in H_{p+1}(\mathbb{M}, \mathbb{M}^{b_n})$ classe que nasce em $(\mathbb{M}, \mathbb{M}^{b_n})$ de homologia relativa de dimensão $(p+1)$.

Considere $\partial\Gamma \in H_p(\mathbb{M}^{b_n})$ classe de homologia de dimensão p .

Mostraremos que $\partial\Gamma$ nasce em \mathbb{M}^{b_n} . Uma vez que Γ nasce em $(\mathbb{M}, \mathbb{M}^{b_n})$, segue que $\Gamma \notin (z_{b_{n+1}}^{b_n})^*H_{p+1}(\mathbb{M}, \mathbb{M}^{b_n})$.

Suponha que exista $\beta \in H_p(\mathbb{M}^{b_{n+1}})$, tal que $(h_{b_{n+1}}^{b_n})^*(\beta) = \partial\Gamma$, ou seja, $\partial\Gamma$ não nasceria em \mathbb{M}^{b_n} .

Queremos mostrar que $\beta \in \ker i^{b_{n+1}}$ e, por conseguinte, $\beta \in \text{Im } \partial$ uma vez que a sequência longa do par é exata.

O quadrado comuta, assim:

$$(i^{b_n})^*(h_{b_{n+1}}^{b_n})^*(\beta) = (i^{b_n})^*(\beta)$$

contudo, $(h_{b_{n+1}}^{b_n})^*(\beta) = \partial\Gamma$, por hipótese de contradição.

Então:

$$(i^{b_{n+1}})^*(\beta) = (i^{b_n})^*(h_{b_{n+1}}^{b_n})^*(\beta) = (i^{b_n})^*\partial\Gamma = 0,$$

$$(i^{b_{n+1}})^*(\beta) = 0,$$

portanto, $\beta \in \ker(i^{b_{n+1}})^*$, logo $\beta \in \text{Im } \partial$, pois a sequência longa do par é exata.

Temos que existe $\beta' \in H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{n+1}})$ classe de homologia $(p+1)$ -dimensional, tal que $\partial\beta' = \beta$.

Notemos que $\partial(z_{b_{n+1}}^{b_n})^*\beta' = \partial\Gamma \in H_p(\mathbb{M}^{b_n})$, classe de homologia de dimensão p .

Agora,

$$\partial(z_{b_{n+1}}^{b_n})^*\beta' = (h_{b_{n+1}}^{b_n})^*\partial\beta = (h_{b_{n+1}}^{b_n})^*\beta = \partial\Gamma.$$

Portanto,

$$\partial(z_{b_{n+1}}^{b_n})^*\beta' - \partial\Gamma = 0,$$

$$\partial((z_{b_{n+1}}^{b_n})^*\beta' - \Gamma) = 0.$$

Por conseguinte $(z_{b_{n+1}}^{b_n})^*\beta' - \Gamma \in \text{Im}(\xi^{b_n})^*$.

Então, existe $\Delta \in H_p(\mathbb{M})$, tal que $(\xi^{b_n})\Delta = \partial((z_{b_{n+1}}^{b_n})^*\beta' - \Gamma)$.

Portanto, temos $(\xi^{b_n})^*\Delta = (z_{b_{n+1}}^{b_n})^*\beta' - \Gamma$.

Por um lado,

$$(z_{b_{n+1}}^{b_n})^*((\xi^{b_{n+1}})^*\Delta) = (\xi^{b_n})^*\Delta = (z_{b_{n+1}}^{b_n})^*\beta' - \Gamma.$$

Assim,

$$-(z_{b_{n+1}}^{b_n})^*((\xi^{b_{n+1}})^*\Delta) + (z_{b_{n+1}}^{b_n})^*\beta' = \Gamma,$$

o que implica que $\Gamma \in (z_{b_{n+1}}^{b_n})^*(H_{p+1}(\mathbb{M}, \mathbb{M}^{b_{n+1}}))$, absurdo!

4.8.1 Dualidade e Simetria

A simetria surge como consequência da dualidade de Lefschetz entre os grupos de homologia absoluta e relativa de dimensões complementares $H_p(\mathbb{M}_b) \approx H_{d-p}(\mathbb{M}, \mathbb{M}^b)$.

Isso se traduz através do resultado de dualidade para diagramas de persistência que afirmaremos sem prova.

Usaremos o símbolo T para indicar reflexão pela diagonal principal, associando o ponto (a, b) , com $T(a, b) = (b, a)$.

Teorema 4.6. Teorema de Dualidade e Simetria

Seja f função definida em uma variedade de dimensão d sem fronteira. Então, os diagramas de persistência são refletidos uns nos outros conforme:

$$\text{Ord}_p(f) = \text{Rel}_{d-p}^T(f);$$

$$\text{Ext}_p(f) = \text{Ext}_{d-p}^T(f);$$

$$\text{Rel}_p(f) = \text{Ord}_{d-p}^T(f).$$

Equivalentemente, o p -ésimo diagrama de persistência completo é a reflexão do $(d-p)$ -ésimo diagrama de persistência completo

$$\text{Dgm}_p(f) = \text{Dgm}_{d-p}^T(f).$$

Para $2p = d$ os subdiagramas estendidos são reflexões deles próprios e por conseguinte simétricos através da diagonal principal:

$$Ext_p(f) = Ext_{d-p}^T(f) = Ext_p^T(f).$$

Lembrando que a definição de elevação exige que o pareamento dos pontos críticos seja o mesmo para funções altura antipodais. Temos, especificamente, o resultado estrutural, expressado em termos de subdiagramas de persistência. Utilizaremos o símbolo R para representar reflexão na diagonal menor, avaliando (a, b) em $R(a, b) = (-b, -a)$. E o 0 para indicar reflexão central, ou rotação de 180° , tal que $0(a, b) = (-a, -b)$.

Teorema 4.7. *Seja f função definida em uma variedade de dimensão d , sem fronteira e considere $(-f)$ sua negativa.*

Então, os diagramas de persistência das duas funções são reflexões umas das outras:

$$Ord_p(f) = Ord_{d-p-1}^R(-f);$$

$$Ext_p(f) = Ext_{d-p}^0(-f);$$

$$Rel_p(f) = Rel_{d-p+1}^R(-f).$$

5 Estabilidade

Persistência é um conceito teórico de medida, construído e fundamentado sobre as estruturas algébricas.

A propriedade mais importante é a estabilidade sob as perturbações dos dados. Em outras palavras, pequenas mudanças nos dados implicam, no máximo, em pequenas mudanças na medida de persistência, o que é fundamental.

Assim, existem grandes ramificações das aplicações e também pesquisa, incluindo o estudo de famílias a 1-parâmetro e a comparação e classificação de formas.

5.1 Famílias a 1-parâmetro

Estudaremos como deformações contínuas nos dados afetam a medida de persistência.

Focaremos no efeito estrutural e seus cálculos. A consequência dessa análise é a primeira prova de estabilidade.

5.1.1 Homotopia Linear

Sejam $f : K \rightarrow \mathbb{R}$ e $g : K \rightarrow \mathbb{R}$ duas funções monotônicas, definidas no mesmo complexo simplicial à valores reais. Relembramos que isso significa que as funções são não-decrescentes ao longo de cadeias crescentes com respeito a relação de face.

A homotopia linear

$$F : K \times [0, 1] \rightarrow \mathbb{R}$$

definida por:

$$F(\sigma, t) = (1 - t)f(\sigma) + tg(\sigma),$$

é uma deformação entre as funções f e g .

Definimos $f_t(\sigma) = F(\sigma, t)$ e notamos que $f_0 = f$ e $f_1 = g$, como desejávamos.

Além disso, f_t é monôtonica, para cada $t \in [0, 1]$ fixado.

De fato, se σ for face de τ , então $f(\sigma) \leq f(\tau)$ e $g(\sigma) \leq g(\tau)$ pois as funções são monôtonicas, por conseguinte:

$$f_t(\sigma) = (1 - t)f(\sigma) + tg(\sigma) \leq (1 - t)f(\tau) + tg(\tau) = f_t(\tau),$$

para todo $t \in [0, 1]$ arbitrário fixado.

Daí somos capazes de encontrar uma ordenação compatível dos simplexes, isto é, uma ordenação total que estende a ordem parcial estabelecida por f_t e pela relação de face.

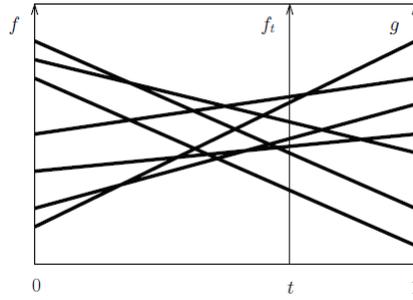


Figura 5.1: Cada linha armazena o valor da função definida nos simplexos enquanto o parâmetro t aumenta. Para qualquer $t \in [0, 1]$ fixado, temos f_t através da interseção com correspondente reta vertical.

No entanto, se de alguma forma já tivermos os diagramas de f , então somos capazes de considerar modificá-los para conseguir os diagramas de f_t . Acontece que essa abordagem é mais eficiente do que recalculer os diagramas providos das duas ordens totais, já que eles não são tão diferentes.

Para descrever exatamente o que isso significa, plotamos os valores da função com o tempo, nos dando uma linha reta para cada simplexo.

É conveniente assumir que f e g são injetoras, o que representa que $f(\sigma_1) = f(\tau_1)$ e $g(\sigma_2) = g(\tau_2)$, então $\sigma_1 = \tau_1$ e $\sigma_2 = \tau_2$. Portanto,

$$f_t(\sigma) = f_t(\tau);$$

$$tg(\sigma) - tf(\sigma) + f(\sigma) = tg(\tau) - tf(\tau) + f(\tau).$$

Rearranjando os termos de maneira mais conveniente e passando os termos para o mesmo lado da igualdade, obtemos

$$t(g(\sigma) - g(\tau)) + (1 - t)(f(\sigma) - f(\tau)) = 0.$$

Assim,

$$t(g(\sigma) - g(\tau)) = -(1 - t)(f(\sigma) - f(\tau)).$$

Uma vez que f e g são injetoras, segue que $(f(\sigma) - f(\tau)) \neq 0$ e $(g(\sigma) - g(\tau)) \neq 0$, portanto somos aptos a considerar

$$\frac{g(\sigma) - g(\tau)}{f(\sigma) - f(\tau)} = 1 - \frac{1}{t}.$$

Para dados dois simplexos distintos σ e τ , como f e g são injetivas, f_t será injetiva, exceto por uma quantidade finita de pontos t , uma vez que a equação acima é uma função de grau 1, e sempre que ocorrer um zero na equação, teremos a não-injetividade de f_t .

A fim de termos uma melhor simplificação da situação, suporemos que mais de um par diferente de linhas não se cruzam ao mesmo tempo.

Equivalentemente, cada f_t tem, no máximo, uma violação de injetividade, ou seja, no máximo dois simplexos com o mesmo valor avaliado pela função.

Conforme varremos da esquerda para a direita, na direção do crescimento de t , passamos pelas violações de injetividade transpondo os dois simplexos na ordenação compatível.

Isso nos motiva a estudar o impacto da transposição na persistência.

5.2 Decomposição Matricial

Relembramos que calcular o diagrama de persistência de $f : K \rightarrow \mathbb{R}$ se faz reduzindo a matriz bordo do qual as colunas e linhas são ordenados segundo a ordenação compatível. Iniciando com $R = \partial$ performamos operações de adição nas colunas da esquerda para a direita até que R seja a matriz reduzida, ou seja, cada coluna não-nula possui um lowest 1 em uma única linha. Assim, o mapeamento das colunas não-nulas para as linhas definidas pelo $low(j)$ é injetor. Cada lowest 1 fornece um par de simplexes, (σ_i, σ_j) , se $i = low(j)$ e um par finito no diagrama de persistência $(f(\sigma_i), f(\sigma_j))$ no diagrama de persistência de f , $Dgm_p(f)$, onde p é a dimensão de σ_i .

Será conveniente assumir bijeção entre lowest 1 e os pontos fora da diagonal nos diagramas de persistência, representando assim que não existem pontos fora da diagonal no infinito, pois se houvesse, existiria um simplexo, digamos σ , que nasce e nunca morre, logo ele não é pareado com nenhum outro simplexo, por conseguinte a matriz reduzida não teria lowest 1 definido, o que implica em não bijetividade.

Não há perda de generalidade, uma vez que sempre podemos adicionar simplexes ao final do processo de filtração, pois dessa maneira eles não alteram a evolução dos subníveis.

A matriz reduzida pode ser escrita como $R = \partial V$, onde V mantém o controle das operações de coluna, onde sua j -ésima coluna armazena a cadeia cujo bordo é armazenado na j -ésima coluna de R .

Desde que apenas utilizamos operações de coluna da esquerda para a direita, V é triangular superior, com $V[i, i] = 1$. Portanto, a matriz V é invertível, já que $det(V) = 1 \neq 0$, consideremos U sua inversa. Mais uma vez, U é matriz triangular superior

$$R = \partial V,$$

multiplicando à direita por U ambos os lados

$$RU = \partial VU = \partial,$$

$\partial = RU$, chamamos de decomposição ru da matriz bordo.

Implicitamente estamos requerindo que U seja triangular superior, invertível e ainda que R seja reduzida. Tais propriedades são satisfeitas pela forma como construímos as matrizes, no entanto, há outras decomposições ru que podem ser obtidas por algoritmos similares.

De fato a decomposição ru de ∂ não é única, mas como observamos anteriormente, o lowest 1 na matriz reduzida o é.

A questão específica que fazemos é: como vamos atualizar a decomposição ru da matriz bordo se vamos transpor dois simplexes em posições contíguas ao longo da ordenação compatível?

5.2.1 Atualização da Decomposição

Suponha ∂ a matriz bordo para a ordenação de simplexes $\sigma_1, \sigma_2, \dots, \sigma_n$. Escreva ∂' a matriz bordo após a transposição de σ_i com σ_{i+1} . Considerando $P = P_i^{i+1}$ correspondente a matriz permutação, temos $\partial' = P\partial P$, onde as únicas diferenças entre P e a matriz identidade ocorrem nas posições $P[i, i] = 0 = P[i + 1, i + 1]$ e nas entradas $P[i, i + 1] = 1 = P[i + 1, i]$.

Multiplicar por P à esquerda é trocar duas linhas.

Multiplicar por P à direita é trocar duas colunas.

É fácil ver que P é sua própria inversa, isto é, $P^2 = PP = Id$.

Assim implica que

$$\delta' = P\delta P = P(RU)P = P(RIdU)P = P(RPPU)P = (PRP)(PUP)$$

no entanto, isso não é necessariamente uma decomposição ru da matriz bordo, a falha pode surgir no caso de $R' = PRP$ não ser reduzida, ou, se $U' = PUP$, não for triangular superior.

Ambas deficiências são corrigidas com um pequeno esforço, expresso em um número constante de operações coluna e linha.

A única maneira da qual R' pode falhar em ser reduzida é quando as linhas i e $(i + 1)$ de R , ambas contiverem lowest 1, $i = low(k)$ e $(i + 1) = low(l)$ e ainda ocorrer 1 na linha i com a coluna l .

Note que $i < k$ ou $i + 1 < l$ pela ordenação compatível, já que simplexes de dimensão menores vem antes e são faces dos de dimensão maiores.

Logo, temos que existem dois casos diferentes: $k < l$ (caso 1) ou $l < k$ (caso 2). Em ambos os casos, adicionamos a coluna da esquerda na coluna da direita, antes de fazermos a transposição. O que conserta a deficiência, trocando R antes mesmo de transpor σ_i com σ_{i+1} .

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1	σ_1							
...								
i	σ_i			1		1		
i+1	σ_{i+1}			0		1		

Figura 5.2: A matriz R exibindo a deficiência ressaltada no caso 1.

Se $k < l$, temos que k está a esquerda de l , procedemos adicionando à coluna k na coluna l .

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1	σ_1							
...								
i	σ_i			1		0		
i+1	σ_{i+1}			0		1		

Figura 5.3: Após somar as colunas indicadas, corrigimos a matriz R para o caso 1.

Como vamos multiplicar P à esquerda de R temos que a posição das linhas se alteram, conseguindo R' reduzida.

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1		σ_1						
...								
i		σ_{i+1}		0		1		
i+1		σ_i		1		0		

Figura 5.4: Ao final do processo matriz R' é reduzida, para o caso 1.

Agora,

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1		σ_1						
...								
i		σ_i		1		1		
i+1		σ_{i+1}		1		1		
...				0		0		
m		σ_m		0		0		

Figura 5.5: A matriz R exibindo a deficiência ressaltada para o caso 2.

Já no caso $l < k$ vamos adicionar a coluna l na coluna k e conseguimos a seguinte configuração:

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1		σ_1						
...								
i		σ_i		1		0		
i+1		σ_{i+1}		1		1		
...				0		0		
m		σ_m		0		0		

Figura 5.6: Após somar as colunas indicadas, corrigimos a matriz R para o caso 2

Nessa situação perdemos o requisito da matriz reduzida, porém abrimos mão dessa exigência, pois com a aplicação da matriz P à esquerda conseguiremos R' reduzida.

	Coluna	1	...	k	...	l	...	m
Linha		σ_1		σ_k		σ_l		σ_m
1		σ_1						
...								
i		σ_i		1		1		
i+1		σ_{i+1}		1		0		
...				0		0		
m		σ_m		0		0		

Figura 5.7: Apesar de R não ser reduzida, R' é, por conta da aplicação da matriz P .

A única maneira que U falha na condição de ser triangular superior é se existir i , tal que $U[i, i + 1] = 1$. Corrigimos esse problema adicionando a linha $(i + 1)$ à linha i em U e adicionando a coluna i à coluna $(i + 1)$ em R .

Como a aplicação da matriz P , o elemento $U[i, i + 1]$ transcreverá o seguinte caminho: Quando aplicarmos UP teremos $\tilde{U}[i, i]$ armazenando o antigo $U[i, i + 1] = 1$, segue que ao avaliarmos $P(UP)$ o elemento $\tilde{\tilde{U}}[i + 1, i]$ contém $U[i, i + 1] = 1$, assim, teremos o número 1 abaixo da diagonal principal, logo, a matriz não é triangular superior.

Por questões técnicas, adicionaremos a coluna i à coluna $(i + 1)$ em R .

Considerando $S = S_i^{i+1}$ a matriz da qual sua única diferença para a matriz identidade é que $S[i, i + 1] = 1$, por conseguinte podemos considerar $SU = RS$.

Desde que $P^2 = Id = S^2$ não altera o produto matricial pois

$$\delta' = (PRP)(PUP) = (PRSP)(PSUP) = PRUP,$$

como anteriormente.

Com essa modificação $PSUP$ é triangular superior, contudo $PRSP$ pode falhar novamente.

Caso a coluna i seja nula ou $low(i) < low(i + 1)$ então a multiplicação à direita por S preservará o lowest 1 e RS será reduzida. Nesse caso, teremos uma decomposição ru após a transposição.

Porém, temos mais dois casos a se considerar: se caso a coluna $(i + 1)$ for nula enquanto a coluna i não for (caso 3), ou, se caso $low(i) > low(j)$ (caso 4). Então, precisaremos tornar lowest 1 único de novo. Para fazê-lo, adicionamos a coluna $(i + 1)$ à coluna i e depois da transposição resultando uma operação de adição de colunas da esquerda para a direita.

De fato, observando os casos, temos:

Se a coluna i tiver lowest 1 na posição l , enquanto a coluna $(i + 1)$ for nula:

	Coluna	1	2	...	i	$i+1$...	m
Linha		σ_1			σ_i	σ_{i+1}		σ_m
1	σ_1					0		
2	σ_2					0		
...						...		
l	σ_l				1	0		
$l+1$	σ_{l+1}					0		
...						...		
m						0		

Figura 5.8: Deficiência ressaltada na matriz R no caso 3.

A configuração de R passará ser a seguinte após multiplicação à direita de S :

	Coluna	1	2	...	i	$i+1$...	m
Linha		σ_1			σ_i	σ_{i+1}		σ_m
1	σ_1					0		
2	σ_2					0		
...						...		
l	σ_l				1	1		
$l+1$	σ_{l+1}				0	0		
...						...		
m						0		

Figura 5.9: Deficiência na matriz R após multiplicação por S , caso 3.

Não tendo mais lowest 1 e ainda $P(RS)P$.

R	Coluna	1	2	...	i	i+1	...	m
Linha		σ_1			σ_i	σ_{i+1}		σ_m
1	σ_1					0		
2	σ_2					0		
...						...		
i	σ_i				0	?		
i+1	σ_{i+1}				0	?		
...								
l	σ_l				1	1		

Figura 5.10: Deficiência se propaga e não obtemos *PRSP* reduzida.

Obtendo *PRSP* não reduzida.

Agora, se $low(i) = l$ e $low(i + 1) = k$, com $k < l$, então:

R	Coluna	1	...	i	i+1	...	m
Linha		σ_1		σ_i	σ_{i+1}		σ_m
1	σ_1						
...							
k	σ_k				1		
...					0		
l	σ_l			1	0		
...				0	0		
m	σ_m			0	0		

Figura 5.11: A deficiência da matriz *R* no caso 4.

Aplicando *RS* temos:

RS	Coluna	1	...	i	i+1	...	m
Linha		σ_1		σ_i	σ_{i+1}		σ_m
1	σ_1						
...							
k	σ_k				1		
...					0		
l	σ_l			1	1		
...				0	0		
m	σ_m			0	0		

Figura 5.12: A deficiência da matriz *R* após aplicar a matriz *S*, no caso 4.

E por fim *PRSP*:

PRSP	Coluna	1	...	i	i+1	...	m
Linha		σ_1		σ_i	σ_{i+1}		σ_m
1	σ_1						
...							
k	σ_k			1			
...							
l	σ_l			1	1		
...				0	0		
m	σ_m			0	0		

Figura 5.13: A matriz *PRSP* não é reduzida para o caso 4.

Assim, para ambos os casos, adicionando a coluna $(i + 1)$ na coluna i , e após o resultado das transposições reparamos as deficiências e conseguimos uma decomposição ru de ∂' .

5.2.2 Como o pareamento se altera

Precisamos entender quais transposições tem efeito no pareamento. As que assim o fazem, chamamos de trocas.

Como um princípio básico, a maior parte das transposições não são trocas. Por exemplo, se σ_i e σ_{i+1} possuem dimensões diferentes, então sua transposição não requer nenhuma alteração, além da obrigatória troca de colunas e linhas. Isso porque, com dimensões diferentes, suas respectivas faces de codimensão 1 terão também dimensões diferentes entre si, logo os lowest 1 não conflitam, como vimos no estudo de casos anteriores. Ainda que possuam a mesma dimensão, porém, se σ_i for positivo e σ_{i+1} for negativo, implica que a transposição não será uma troca, isso porque a linha i não terá nenhum lowest 1, assim $R' = PRP$ será reduzida e não exigirá nenhum esforço adicional. Em outras palavras, a decomposição ru é mantida sem qualquer operação de reparação que altere o pareamento.

Após tais observações e estudos de casos, chegamos à seguinte conclusão:

Lema 5.1. Lema da Transposição

Sejam ∂ e ∂' matrizes de bordo de uma ordenação compatível de duas funções monotônicas em um complexo simplicial que difere apenas pela transposição de dois simplices contíguos σ_i e σ_{i+1} .

Então, o pareamento definido pela decomposição ru $\partial = RU$ e $\partial' = R'U'$ difere apenas se $\dim\sigma_i = \dim\sigma_{i+1}$ e caso difiram, será somente pela troca de lugares de σ_i e σ_{i+1} .

5.3 Teorema de Estabilidade

5.3.1 Distância Bottleneck

O diagrama de persistência é um multiconjunto de pontos no plano estendido, $\bar{\mathbb{R}}^2$. Segundo as suposições estabelecidas sobre a função f nesse trabalho, segue que os diagramas consistem em uma porção finita de pontos acima da diagonal principal e a esse multiconjunto adicionamos os infinitos pontos da diagonal, cada um com multiplicidade infinita. Esses pontos extras não são essenciais para o diagrama, contudo sua presença simplificará definições e resultados que estão por vir.

Consideremos dois diagramas de persistência, X e Y . Para definir a distância entre eles, vamos considerar bijeções $\eta : X \rightarrow Y$ e gravamos o supremo das distâncias entre os correspondentes pontos de cada.

Medindo a distância entre os pontos $x = (x_1, x_2)$ e $y = (y_1, y_2)$ com a norma L_∞ :

$$\|x - y\|_\infty = \max_{i=1,2} \{|x_i - y_i|\}$$

e tomando o ínfimo sobre todas as bijeções, obtemos a distância Bottleneck entre os diagramas:

$$W_\infty(X, Y) = \inf_{\eta: X \rightarrow Y} \sup_{x \in X} \|x - \eta(x)\|_\infty.$$

Podemos desenhar quadrados centrados nos pontos de X , de lado cuja medida é duas vezes a distância Bottleneck, assim cada quadrado conterá o correspondente ponto de Y .

Observação 5.1. O fato de adicionarmos infinitos pontos na diagonal é o que nos permite considerar bijeções quaisquer sobre os conjuntos X e Y , uma vez que ambos tem infinitos pontos.

A distância Bottleneck é, de fato, uma distância.

Sejam X, Y dois diagramas de persistência. Assim, $W_\infty(X, Y) = 0$ se, e somente se, $X = Y$.

Se $W_\infty(X, Y) = 0$, significa que $\inf_{\eta: X \rightarrow Y} \sup_{x \in X} \|x - \eta(x)\|_\infty = 0$, considerando a bijeção $x = (x_1, x_2) \mapsto \eta(x) = (\eta x_1, \eta x_2)$, temos:

$$\sup_{x \in X} \{\max_{i=1,2} \{|x_i - \eta x_i|\}\} = 0$$

e ainda $|x - \eta x| \geq 0$, para cada $x \in X$.

Uma vez que tomamos o máximo das distâncias, temos

$$|x_1 - \eta x_1| = 0 = |x_2 - \eta x_2|,$$

portanto, $x_i = \eta x_i$, para cada $i = 1, 2$ e todo $x \in X$. Assim, como tomamos o máximo das diferenças e temos o resultado 0, segue que para todos os $x \in X$, $|x - \eta(x)| = 0$, portanto, $x = \eta x$, e desde que tomamos o ínfimo das bijeções, segue que $\eta = Id$, função identidade. Logo, $Y = \eta(X) = Id(X) = X$.

Reciprocamente, se $X = Y$, ao tomarmos o ínfimo das bijeções $\eta : X \rightarrow X$, teremos que $\|x - Id(x)\| = 0$ e o resultado segue.

É claro que $W_\infty(X, Y) = W_\infty(Y, X)$ desde que trabalhamos com bijeções $\eta : X \rightarrow Y$, implica que existe $\eta^{-1} : Y \rightarrow X$, assim $x = \eta^{-1}(y)$, para algum $y \in Y$, daí $\|x - \eta(x)\|_\infty = \|\eta^{-1}(y) - \eta\eta^{-1}(y)\|_\infty = \|\eta^{-1}(y) - y\|_\infty = \|y - \eta^{-1}(y)\|_\infty$, pois L_∞ é norma. Obtemos o desejado.

Sejam $\alpha : X \rightarrow Y$ e $\beta : Y \rightarrow Z$ bijeções. Tome $\eta = \beta\alpha : X \rightarrow Z$.

Assim,

$$\begin{aligned} \|x - \beta(\alpha(x))\|_\infty &= \|x - \alpha(x) + \alpha(x) - \beta(\alpha(x))\|_\infty \\ &\leq \|x - \alpha(x)\|_\infty + \|\alpha(x) - \beta(\alpha(x))\|_\infty \\ &\leq \sup_{x \in X} \|x - \alpha(x)\|_\infty + \sup_{y \in Y} \|y - \beta(y)\|_\infty, \text{ para todo } x \in X. \end{aligned}$$

Portanto,

$$\sup_{x \in X} \|x - \beta(\alpha(x))\|_\infty \leq \sup_{x \in X} \|x - \alpha(x)\|_\infty + \sup_{y \in Y} \|y - \beta(y)\|_\infty.$$

E temos, pela definição de distância Bottleneck,

$$W_\infty(X, Z) \leq \sup_{x \in X} \|x - \beta(\alpha(x))\|_\infty \leq \sup_{x \in X} \|x - \alpha(x)\|_\infty + \sup_{y \in Y} \|y - \beta(y)\|_\infty,$$

para qualquer $\alpha : X \rightarrow Y$ e para qualquer $\beta : Y \rightarrow Z$, bijeções.

Em particular,

$$W_\infty(X, Z) - \sup_{x \in X} \|x - \alpha(x)\|_\infty \leq \sup_{y \in Y} \|y - \beta(y)\|_\infty \text{ para todo } \beta : Y \rightarrow Z.$$

Tomando o ínfimo das bijeções $\beta : Y \rightarrow Z$ em ambos os lados conseguimos

$$W_\infty(X, Z) - \sup_{x \in X} \|x - \alpha(x)\|_\infty \leq W_\infty(Y, Z), \text{ para todo } \alpha : X \rightarrow Y.$$

Finalmente consideramos

$$W_\infty(X, Z) - W_\infty(Y, Z) \leq \sup_{x \in X} \|x - \alpha(x)\|_\infty, \text{ para todo } \alpha : X \rightarrow Y.$$

Mais uma vez tomamos o ínfimo das bijeções, mas agora de $\alpha : X \rightarrow Y$,

$$W_\infty(X, Z) - W_\infty(Y, Z) \leq W_\infty(X, Y).$$

Então,

$$W_\infty(X, Z) \leq W_\infty(X, Y) + W_\infty(Y, Z).$$

Vemos que W_∞ satisfaz todos os axiomas de métrica, portanto, assim merece ser chamado de distância.

5.4 Estabilidade Bottleneck

Sejam $f : K \rightarrow \mathbb{R}$ e $g : K \rightarrow \mathbb{R}$ duas funções monotônicas definidas em um mesmo complexo simplicial K a valores em \mathbb{R} . Considere a homotopia linear entre as funções f e g , $f_t = (1 - t)f + tg$.

Para cada $t \in [0, 1]$ e cada dimensão p , dispomos de uma função monotônica f_t e um diagrama de persistência.

Fixada uma dimensão p , a família de diagramas persistentes é um multiconjunto em $\mathbb{R}^2 \times [0, 1]$.

Desenhando t ao longo do terceiro eixo coordenado, ganhamos uma visualização tridimensional de como a homologia persistente evolui de $f_0 = f$ até $f_1 = g$.

A fim de descrever esse processo, assumimos K com homologia reduzida trivial.

Acrescentando a terceira coordenada, cada um dos elementos fora da diagonal de $X_t = Dgm_p(f_t)$ é da forma $x(t) = (f_t(\sigma), f_t(\tau), t)$, onde $\sigma, \tau \in K$ simplexos, representando o fato que enquanto construímos K através das filtrações regidas pela ordenação estabelecida por f_t , temos que σ gera o nascimento de uma classe de homologia de dimensão p , já τ causa a morte da mesma.

Há apenas finitos valores dos quais o pareamento dos simplexos muda, denotaremos como $0 = t_0 < t_1 < \dots < t_n < t_{n+1} = 1$. Além disso, em cada intervalo (t_i, t_{i+1}) o pareamento é constante e cada par σ, τ gera um segmento de reta de pontos $x(t)$ conectando os pontos no plano $t = t_i$ e $t = t_{i+1}$.

Se o ponto final for ponto fora da diagonal em t_{i+1} , então existirá outro segmento de reta, único, que começa nesse ponto e ligará o plano $t = t_{i+1}$, com $t = t_{i+2}$.

Esse segmento de reta pode corresponder ao mesmo par de simplexos e assim continuar na mesma linha reta, ou pode corresponder a um par diferente, criado durante uma troca e a linha reta, a partir desse ponto, fará um desvio em relação à reta anterior no ponto em que compartilham.

Também é possível que o ponto final caia na diagonal no nível t_i , nesse caso, não haverá continuação do segmento de reta. Em suma, os segmentos de reta formam caminhos poligonais que monotonamente crescem em t .

Cada caminho se inicia em um ponto fora da diagonal principal, ou se inicia em um ponto da diagonal e termina em um ponto final fora da diagonal.

Definição 5.1. *Denominaremos cada caminho poligonal que conecta os diagramas de persistência de vinha e um multiconjunto de vinhas, vinhedo.*

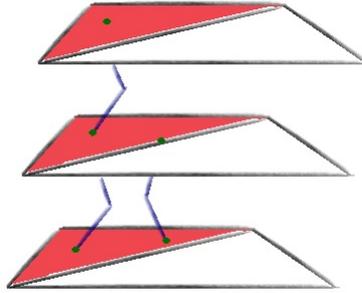


Figura 5.14: A família de diagramas de persistência à 1-parâmetro da homotopia linear entre $f_0 = f$ e $f_1 = g$. Alguns pontos geram vinhas que percorrem todas as camadas, enquanto outras morrem.

O fato que pontos na família de diagramas de persistência formam vinhas é importante. É uma maneira de dizer que o diagrama de persistência é estável.

Finalmente alcançamos o terceiro resultado fundamental da teoria de Persistência, Teorema de Estabilidade para Filtrações, que exhibe condições necessárias para que dois diagramas de persistência sejam tão próximos quanto queiramos:

Teorema 5.1. Teorema de Estabilidade para Filtrações

Sejam $f : K \rightarrow \mathbb{R}$ e $g : K \rightarrow \mathbb{R}$ duas funções monotônicas definidas em K , um complexo simplicial, à valores reais.

Para cada dimensão p , a distância Bottleneck entre os diagramas $X = Dgm_p(f)$ e $Y = Dgm_p(g)$ é limitada superiormente pela distância entre as funções f e g , segundo a norma L_∞ :

$$W_\infty(X, Y) \leq \|f - g\|_\infty.$$

Demonstração. Tome p dimensão arbitrária fixada.

Seja $X_t = Dgm_p(f_t)$, para algum $t \in [t_0, t_{n+1}]$.

Se $x \in X_t$, então $x(t) = (f_t(\sigma), f_t(\tau))$ onde σ e τ são simplexes positivo e negativo, respectivamente.

Para quantificar essa noção, observamos que

$$x(t) = (f_t(\sigma), f_t(\tau), t) = ((1-t)f(\sigma) + tg(\sigma), (1-t)f(\tau) + tg(\tau), t),$$

$$x(t) = (1-t)(f(\sigma), f(\tau), 0) + t(g(\sigma), g(\tau), 1).$$

Diferenciando em relação à t , conseguimos

$$\frac{dx}{dt} = (g(\sigma) - f(\sigma), g(\tau) - f(\tau), 1).$$

Agora, projetando os pontos finais do segmento de volta no plano estendido $\bar{\mathbb{R}}^2$, obtemos os pontos

$$\pi \frac{dx}{dt} = (g(\sigma) - f(\sigma), g(\tau) - f(\tau)).$$

Considere $Z = Dgm_p(f_{t+h})$ onde, para $\epsilon > 0$ suficientemente pequeno escolhido, tal que $h \in (0, \epsilon)$ com $t_h \in (t_i, t_{i+1})$.

Seja $z \in Z$, notemos que

$$\begin{aligned} z(t+h) &= (f_{t+h}(\sigma), f_{t+h}(\tau), t+h) \\ &= ((1-(t+h))f(\sigma) + (t+h)g(\sigma), (1-(t+h))f(\tau) + (t+h)g(\tau), t+h). \end{aligned}$$

Doravante, assuma que aplicamos a projeção e trabalhamos, novamente, no plano estendido.

Construamos a bijeção

$$\eta : X_t \rightarrow Z$$

dada por

$$\eta(x) = (f_t(\sigma) - hf(\sigma) + hg(\sigma), f_t(\tau) - hf(\tau) + hg(\tau)),$$

onde $x = (f_t(\sigma), f_t(\tau))$.

Veja que η é claramente uma bijeção, já que consiste em adicionar os termos $(hg - hf)$ em cada uma das entradas.

E ainda mais, $z = x(t+h) = \eta(x)$, assim,

$$W_\infty(X_t, Z) \leq \sup_{x \in X} \|x - \eta(x)\|_\infty,$$

pois o ínfimo das bijeções é necessariamente menor ou igual do que uma bijeção tomada de forma particular.

Além disso,

$$\|x(t+h) - x(t)\|_\infty \leq \|h\| \sup_{0 \leq t_0 \leq 1} \left\| \frac{dx}{dt}(t_0) \right\|_\infty,$$

segue pela Desigualdade do Valor Médio.

Contudo, $|h| < |t_i - t_{i+1}|$ desde que foi escolhido dessa forma e como já tínhamos calculado $\frac{dx}{dt}$ independente do parâmetro, assim

$$\|x(t+h) - x(t)\|_\infty \leq |t_i - t_{i+1}| \sup_{0 \leq t_0 \leq 1} \|(g(\sigma) - f(\sigma), g(\tau) - f(\tau))\|.$$

Portanto,

$$\|x(t+h) - x(t)\|_\infty \leq |t_i - t_{i+1}| \|f - g\|_\infty.$$

Finalmente,

$$W_\infty(X_t, Z) \leq \sup_{x \in X} \|x - \eta(x)\|_\infty \leq |t_i - t_{i+1}| \|f - g\|_\infty.$$

Desde que W_∞ é uma distância, aplicaremos o processo para cada $X_t = Dgm_p(f_t)$ e $X_{t+1} = Dgm_p(f_{t+1})$, com $t \in [0, 1]$ e conseguiremos:

$$\begin{aligned} W_\infty(X, Y) &\leq W_\infty(X_0, X_1) + W_\infty(X_1, X_2) + \dots + W_\infty(X_n, X_{n+1}) \\ &\leq |t_1 - t_0| \|f - g\|_\infty + |t_2 - t_1| \|f - g\|_\infty + \dots + |t_{n+1} - t_n| \|f - g\|_\infty \\ &= |1 - 0| \|f - g\|_\infty. \end{aligned}$$

Concluimos que:

$$W_\infty(X, Y) \leq \|f - g\|_\infty.$$

□

5.4.1 Funções Tame

Para aplicar o Teorema de Estabilidade, é conveniente permitir que classes mais gerais de funções sejam válidas. Existem fatos que garantem que qualquer função contínua definida em um espaço topológico triangularizável pode ser aproximada por uma função linear por partes, e ainda, para toda função linear por partes existe uma função monotônica que gera o mesmo diagrama de persistência. Assim sendo, ainda que exijamos funções monotônicas, não restringimos demasiado a classe de funções válidas nesse teorema.

Agora prosseguiremos com a recordação de algumas definições e após isso, enunciaremos um teorema mais geral, contudo, sem demonstração.

Definição 5.2. Chamamos $a \in \mathbb{R}$ valor crítico homológico se existir um inteiro k , tal que para todo ϵ suficientemente pequeno, o homomorfismo

$$H_k(f^{-1}(-\infty, a - \epsilon]) \rightarrow H_k(f^{-1}(-\infty, a + \epsilon])$$

induzido pela inclusão não seja isomorfismo.

Definição 5.3. A função $f : X \rightarrow \mathbb{R}$ definida em um espaço topológico triangularizável a valores reais é dita tame se possuir apenas um número finito de valores críticos homológicos e os grupos de homologia $H_p(X_a)$, onde $X_a = f^{-1}(-\infty, a]$ conjunto de subnível, tiverem dimensão finita para todo $p \in \mathbb{Z}$ e todo $a \in \mathbb{R}$.

Em particular, funções de Morse definidas em variedades compactas são tame, bem como funções lineares por partes definidas em complexo simplicial finito.

Lema 5.2. Se algum intervalo $[x, y] \subset \mathbb{R}$ não contiver valor crítico homológico de f , então

$$H_k(f^{-1}(-\infty, x]) \rightarrow H_k(f^{-1}(-\infty, y])$$

é isomorfismo para todo $k \in \mathbb{Z}$.

Teorema 5.2. Teorema de Estabilidade para Funções Tame

Seja X um espaço topológico triangularizável e considere $f : X \rightarrow \mathbb{R}$ e $g : X \rightarrow \mathbb{R}$ duas funções tame.

Para cada dimensão p , a distância Bottleneck entre $X = Dgm_p(f)$ e $Y = Dgm_p(g)$ é limitada superiormente pela distância entre as funções f e g segundo a norma L_∞ :

$$W_\infty(X, Y) \leq \|f - g\|_\infty.$$

6 Aplicações

A primeira aplicação das ferramentas matemáticas e computacionais introduzidas nos capítulos anteriores é para análise de dados. Uma atividade que atinge cada disciplina em ciência e engenharia. Os dados podem incluir as leituras de um variedade de sensores, pixels de uma imagem ou até mesmo podem ser o acúmulo de informações extraídas de observações. Invariavelmente, há ruído nos dados, que podem ser sistemáticos ou aleatórios. O ruído pode, também, refletir propriedades genuínas do fenômeno medido, mas em uma escala que está fora da janela de interesse, ou seja, características que sejam muito grandes ou muito pequenas. A abordagem tradicional para tratar o ruído é “suavizar” ou “regularizar” os dados, o que invariavelmente significa que nós os mudamos. Assim, alteramos, por consequência, a informação obtida deles. Contrastando com essa filosofia, as abordagens via homologia persistente propõem medir o ruído sem alterar os dados. É o que as ferramentas estudadas nesse trabalho oferecem.

Apresentaremos, no que segue, dois exemplos de aplicação.

Começamos com dados biológicos.

6.1 Medidas para Dados de Expressão Genética

Nossa primeira aplicação trata de funções definidas em variedade de dimensão 1 à valores reais. A mais simples classe de objetos do qual homologia persistente pode trazer informações significantes. Tais funções surgem da modelagem do desenvolvimento de somitos em vertebrados.

6.1.1 Pré-requisitos

Vertebrados são caracterizados pela espinha dorsal, consistindo em uma sequência de vértebras que providenciam segmentação periódica ao longo do eixo. Ratos são um exemplo, donde sua espinha dorsal possui em torno de sessenta e cinco vértebras. Já as cobras tem um número muito maior, podem ter algumas poucas centenas de vértebras.

Essa estrutura surge no desenvolvimento do embrião, quando os precursores vertebrais, somitos, são formados ritmicamente a partir da mesoderme pressomática. Esse processo é associado com o oscilador molecular que entrega a expressão do gene com um período correspondente ao da formação do somito. Referimo-nos a esse oscilador por relógio de segmentação. O desejo de compreender totalmente esse relógio é a motivação para o trabalho dessa seção.

6.1.2 Funções Lipschitzianas no Círculo

Modelamos os resultados da série temporal como um conjunto de funções definidas no círculo à valores reais, $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ uma para cada gene.

O círculo representa o período de $2l$ e a função localiza abundância de produto genético por período. A mudança requer energia, isto é, a produção e degradação de RNA. Utilizamos disso para justificar nossa hipótese de que f é Lipschitz, isto é

$$|f(s) - f(t)| \leq Lip(f) \|s - t\|,$$

onde $Lip(f)$ é a constante de Lipschitz da função f .

Para funções diferenciáveis, equivale a restringir a derivada entre $\pm Lip(f)$.

Definimos a variação total como a integral da norma da derivada, assim conseguimos:

$$Var(f) = \int_0^{2\pi} |f'(s)| \leq 2\pi Lip(f).$$

Com efeito,

$$|f'(s)| = \left| \lim_{t \rightarrow s} \frac{|f(t) - f(s)|}{\|t - s\|} \right| = \lim_{t \rightarrow s} \frac{|f(t) - f(s)|}{\|t - s\|} \leq \lim_{t \rightarrow s} Lip(f) = Lip(f).$$

Logo,

$$\int_0^{2\pi} |f'(s)| ds \leq \int_0^{2\pi} Lip(f) ds = 2\pi Lip(f).$$

Tal desigualdade será relevante em breve, quando estudarmos estabilidade em diferentes maneiras de medir funções no círculo.

Para preparar esse estudo, consideramos o diagrama de persistência de f que expressa o histórico de nascimentos e mortes na sequência de conjuntos de subnível.

Assumindo que f é função de Morse. Temos o nascimento de uma componente em todo ponto mínimo, e a morte de uma componente em cada máximo, exceto no máximo global que dá o nascimento de uma classe de dimensão 1 que nunca morrerá. Analogamente, a classe de dimensão 0 que nasce e nunca morre, ocorre no mínimo global. As demais classes nascem e morrem em valores finitos.

Escrevemos c_0 como o número de mínimos da f e c_1 como o número de máximos de f . Claramente $c_0 = c_1$.

Desde que $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ é Lipschitz, segue que f é função contínua, assim sendo, f admite valor máximo global e mínimo global, já que é definida no compacto \mathbb{S}^1 . Intuitivamente, se f parte do mínimo global, a partir daí deve começar a crescer, portanto, alcançará um máximo local e, após isso, começará a decrescer novamente. Como já atingiu o mínimo global e ainda não passou pelo máximo global, necessariamente a função voltará a crescer, passando por um mínimo local. Se a função encontrar outro ponto crítico não-degenerado, precisa ser um máximo. Dessa forma notamos que a quantidade de mínimos e máximos é igual.

O diagrama de persistência de f que contém informações relevantes é o 0-ésimo diagrama já que nossa variedade tem dimensão 1. Assim, $Dgm_0(f)$ contém $c_0 - 1 = n = c_1 - 1$ pontos na porção finita, pois retiramos o mínimo global e máximo global dessa contagem pois ambos tem persistência infinita como já mencionamos.

Além disso, cada ponto finito corresponde ao mínimo pareado com um máximo. Escrevendo $x = (b_i, d_i)$, onde b_i é o valor do ponto mínimo avaliado pela f e d_i é o valor do ponto máximo avaliado pela f .

Sejam b_0 e b_{n+1} os valores do ponto mínimo global e máximo global, avaliados pela f , na devida ordem. Logo, (b_0, ∞) é ponto no infinito de $Dgm_0(f)$ e (b_{n+1}, ∞) é ponto no infinito de $Dgm_1(f)$.

Escrevemos

$$\Phi(f) = \sum_{i=1}^n (d_i - b_i)$$

a persistência total de f .

E afirmamos que

$$\Phi(f) = \frac{1}{2}Var(f) - (b_{n+1} - b_0).$$

6.1.3 Simplificações

Antes de introduzirmos a medida de quão próximo a função $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ está de ser periódica, primeiramente precisamos entender como podemos simplificar.

Chamamos uma outra função contínua $f_\varepsilon : \mathbb{S}^1 \rightarrow \mathbb{R}$ de uma ε -simplificação de f se:

1. $|f(s) - f_\varepsilon(s)| \leq \varepsilon$ para todo $s \in \mathbb{S}^1$.
2. Um ponto fora da diagonal pertence à $Dgm_p(f_\varepsilon)$ se, e somente se, pertencer ao $Dgm_p(f)$ e a distância da diagonal exceder ε .

A segunda condição nos diz que o diagrama de persistência das duas funções são os mesmos exceto que o diagrama de f_ε não tem pontos com persistência ε ou menor. Em outras palavras, a ε -simplificação só é sensível a persistências maiores que ε , descartando o que se considera de informação desprezível por ter pouca duração.

Se f possuir apenas um máximo e um mínimo, esses serão os globais e a f é sua própria ε -simplificação, já que o processo consiste em retirar pontos de persistência ε ou menor da porção finita do diagrama, contudo máximo e mínimo globais são pontos da diagonal, de persistência infinita.

Suponha que f tenha pelo menos dois mínimos e dois máximos. Sejam u e v simples que formam o par com menor persistência e considere o ponto $x = (f(u), f(v)) \in Dgm_0(f)$, com $pers_f(x) < \varepsilon$. Uma vez que esses pontos estão pareados, significa que u é mínimo e consequentemente, v máximo.

Como $f(v)$ é máximo local, segue que $f(v) \geq f(x)$, para todo x em uma vizinhança de v , logo os valores $f(x)$ crescem, por ambos os lados, até atingir $f(v)$, sendo ele o maior valor em uma determinada vizinhança. Em $f(u)$, mínimo local, o valor ao se afastar dele, cresce até atingir $f(v)$. Considere um ponto t tal que $f(t) = f(v)$. Podemos alterar f definindo $f(s) = f(v)$, para todo s contido no arco que vai de t até v e que contenha u . Aos demais pontos preservamos seus valores originais.

Assim conseguimos

$$f_\varepsilon : \mathbb{S}^1 \rightarrow \mathbb{R}$$

dada por $f_\varepsilon(s) = f(v)$ se s pertence ao arco que liga t à v , contendo u e $f_\varepsilon(s) = f(s)$, caso contrário.

Notemos que f_ε é contínua pois o ponto t , avaliado tanto pela f quanto por f_ε , é associado à $f(v)$, bem como ocorre no ponto v . Todos os pontos pertencentes ao arco de t até v passando por u , segundo f_ε , valem constante igual à $f(v)$, ou seja, o ponto,

antes mínimo u , salta seu valor para o mesmo valor que o ponto máximo, e assim perante f_ε ambos deixam de ser mínimo e máximo local, e conseqüentemente, esse par deixa de existir. Ou seja, no diagrama de f_ε não temos esse ponto com persistência menor que ε .

6.1.4 Medidas

A função seno que mapeia pontos de \mathbb{S}^1 em sua segunda coordenada no \mathbb{R}^2 é o protótipo de função periódica. Possui apenas um máximo e um mínimo, variando suavemente entre estes valores. Permitindo casos com padrões mais gerais de crescimento e decrescimento, retemos a propriedade de ter apenas dois pontos críticos como característica ideal de uma função periódica.

A fim de quantificar mais geralmente a periodicidade, atribuímos o número zero às funções com $c_0 + c_1 = 2$, onde mais uma vez c_0 representa a quantidade de mínimos e c_1 de máximos, e atribuímos um número positivo para todas as outras funções.

Restringiremos nossa discussão à funções lineares por partes. Especificamente definimos grau de periodicidade:

$$\mu_0(f) = \frac{1}{2}(c_0 + c_1) - 1.$$

Assim, se f for uma função periódica no sentido tradicional, como a função seno por exemplo, então $\mu_0(f) = 0$, contudo se a função g possuir dois pontos máximos e dois mínimos, então $\mu_0(g) = 1$, logo tem um par de periodicidade extra do que o padrão convencional.

Para cada $q \in \mathbb{Z}_+$ definimos o grau q de periodicidade:

$$\mu_q(f) = \int_{\varepsilon \geq 0} \mu_{q-1}(f_\varepsilon) d\varepsilon.$$

Integrando a medida de grau $(q-1)$ sobre a ε -simplificação de f . Note que $\mu_1(f)$ é proporcional ao número médio de pontos críticos da ε -simplificação:

$$\mu_1(f) = \int_{\varepsilon \geq 0} \mu_0(f_\varepsilon) d\varepsilon.$$

Mas,

$$\mu_0(f_\varepsilon) = \frac{1}{2}(c_0 + c_1) - 1,$$

assim

$$\int_{\varepsilon \geq 0} \left(\frac{1}{2}(c_0 + c_1) - 1\right) d\varepsilon = \frac{\varepsilon}{2}(c_0 + c_1) - \varepsilon.$$

Portanto,

$$\mu_1(f) = \frac{\varepsilon}{2}[(c_0 + c_1) - 2].$$

Definimos, também, o grau q de total persistência de f :

$$\Phi^q(f) = \sum_{i=1}^n (d_i - b_i)^q.$$

Para mostrar que $\mu_q(f)$ é bem definido, demonstraremos que a medida não depende de qual ε -simplificação nós usamos. Faremos através de provar que $\mu_q(f)$ é igual ao grau q de total persistência de f , $\Phi^q(f)$.

Lema 6.1. Lema da Medida de Periodicidade

Seja $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ função linear por partes de Morse. Então, $\mu_q(f) = \Phi^q(f)$, para todo $q \in \mathbb{Z}_{\geq 0}$.

Demonstração. Usaremos indução sobre q para mostrar que a contribuição do ponto $x_i = (b_i, d_i) \in Dgm_0(f)$, no grau q de periodicidade é $(d_i - b_i)^q$.

Para $q = 0$, temos $\mu_0(f) = \frac{1}{2}(c_0 + c_1) - 1 = \frac{1}{2}(2 + 2) - 1 = 2 - 1 = 1$, pois c_0 conta o número de mínimos e temos um global mais um local que é a contribuição de x_i . Analogamente, $c_1 = 2$.

Já o grau 0 de persistência total de f ,

$$\Phi^0(f) = (d_i - b_i)^0 = 1.$$

Note que Φ conta apenas os pontos da porção finita, por isso computamos apenas a contribuição de x_i da qual $(d_i - b_i) > 0$, pois é a persistência do ponto.

Para $q > 1$, a contribuição de x_i ao grau $(q - 1)$ de medida de periodicidade é, por hipótese de indução,

$$\mu_{q-1}(f) = \int_{\varepsilon \geq 0} \mu_{q-2}(f_\varepsilon) d\varepsilon = \Phi^{q-1}(f) = (d_i - b_i)^{q-1}.$$

Resta mostrar que vale para q , isto é,

$$\mu_q(f) = \int_{\varepsilon \geq 0} \mu_{q-1}(f_\varepsilon) d\varepsilon = \int_{\varepsilon \geq 0} (d_i - b_i)^{q-1} d\varepsilon.$$

Agora, desde que f_ε é uma ε -simplificação de f e estamos assumindo que existe apenas um ponto, nomeado x_i , que contribui para a porção finita no diagrama de f , a única forma de simplificação que temos é se retirarmos o único ponto que contribui na porção finita, senão f_ε não seria uma simplificação, e sim a própria f . Logo, $\varepsilon = pers_f(x_i) = (d_i - b_i)$.

Assim, voltando aos cálculos temos

$$\begin{aligned} \mu_q(f) &= \int_{\varepsilon \geq 0} \mu_{q-1}(f_\varepsilon) d\varepsilon = \int_{\varepsilon \geq 0} (d_i - b_i)^{q-1} d\varepsilon \\ &= (d_i - b_i)^{q-1} \int_{b_i}^{d_i} d\varepsilon = (d_i - b_i)^{q-1} (d_i - b_i) = (d_i - b_i)^q \\ &= \Phi^q(f). \end{aligned}$$

Concluimos que

$$\mu_q(f) = \Phi^q(f).$$

Agora, se vale para cada x_i , somando para todos os pontos da porção finita, o resultado segue. □

O Lema de Medida de Periodicidade implica em um algoritmo para computar $\mu_q(f)$, ou seja, construindo o 0-ésimo diagrama de persistência e somando as q -ésimas potências das persistências de seus pontos finitos. Além disso, o Lema também fornece uma definição de μ_q para valores reais de q .

6.1.5 Instabilidade para pequeno grau

Se a medida de periodicidade é estável ou não, isso depende da escolha do índice q . Claramente μ_0 não é estável, pois perturbações arbitrariamente pequenas podem mudar a medida por uma quantidade arbitrária. Menos obviamente μ_1 também não é estável.

Para ver isso, vamos construir uma série de funções Lipschitz $g_k : \mathbb{S}^1 \rightarrow \mathbb{R}$ que se aproxima da função zero, enquanto sua persistência total se aproxima de π .

Substituindo cada ponto de \mathbb{S}^1 por seu ângulo $\varphi \in [0, 2\pi)$, definimos

$$g_0(\varphi) = \min\{\varphi, 2\pi - \varphi\} \quad \text{e} \quad g_k(\varphi) = \frac{1}{2}g_{k-1}(2\varphi) = \frac{1}{2^k}g_0(2^k\varphi).$$

Assim,

$$\|g_k\|_\infty = \sup_{0 \leq \varphi < 2\pi} \left\{ \frac{1}{2^k}g_0(2^k\varphi) \right\} = \frac{1}{2^k} \sup_{0 \leq \varphi < 2\pi} \{ \min\{2^k\varphi, 2\pi - 2^k\varphi\} \},$$

mas, observamos que a função atinge seu máximo em $\varphi = \frac{2\pi}{2^{k+1}}$, portanto

$$\|g_k\|_\infty = \sup \frac{1}{2^k}g_0(2^k\varphi) = \frac{1}{2^k}\pi \rightarrow 0, \quad \text{quando } k \rightarrow \infty.$$

Além disso

$$\frac{dg_k}{d\varphi} = \begin{cases} (\varphi)' = 1, & \text{se } 0 < \varphi < \pi \\ (2\pi - \varphi)' = -1, & \text{se } \pi < \varphi < 2\pi \end{cases}.$$

Logo,

$$Var(g_k) = \int_0^{2\pi} \left| \frac{dg_k}{d\varphi}(\varphi) \right| d\varphi = \int_0^{2\pi} d\varphi = 2\pi, \quad \text{pois } \frac{dg_k}{d\varphi} = \pm 1,$$

enquanto a persistência total

$$\Phi(f) = \sum_{i=1}^n (d_i - b_i)$$

e temos o resultado que garante $\Phi(f) = \frac{1}{2}Var(f) - (b_{n+1} - b_0)$.

Já temos $Var(g_k)$, resta calcular a diferença entre o máximo e mínimo global.

Segue que $b_0 = 0$ e $b_{n+1} = \frac{\pi}{2^k}$.

Portanto,

$$\Phi(g_k) = \pi - \frac{\pi}{2^k} \rightarrow \pi \quad \text{quando } k \rightarrow \infty.$$

6.1.6 Estabilidade para grau pelo menos 2

Existe uma diferença qualitativa para a medida de periodicidade quando q passa de 1 para 2. Em particular, μ_q é estável para toda constante $q \geq 2$.

Teorema 6.1. *Sejam f e g funções lipschitzianas, com constante de Lipschitz igual a um, definidas de \mathbb{S}^1 à valores na reta \mathbb{R} . Então,*

$$|\Phi^q(f) - \Phi^q(g)| \leq 4q\pi^{q-1}\|f - g\|_\infty.$$

Demonstração. Notamos primeiramente que

$$\int_x^y qt^{q-1} dt = q \int_x^y t^{q-1} = q \frac{t^q}{q} \Big|_x^y = y^q - x^q$$

e ainda, $y^q - x^q \leq q|y - x| \max\{x, y\}^{q-1}$ para quaisquer $x, y \geq 0$ e todo $q \geq 1$.

De fato,

$$x^q - y^q = (x - y)(x^{q-1} + x^{q-2}y + \dots + xy^{q-2} + y^q),$$

e seja $\alpha = \max\{x, y\}$.

Assim

$$\begin{aligned} (x - y)(x^{q-1} + x^{q-2}y + \dots + xy^{q-2} + y^q) &\leq (x - y)(\alpha^{q-1} + \alpha^{q-1} + \dots + \alpha^{q-1}) \\ &= (x - y)q\alpha^{q-1} \\ &\leq q|x - y| \max\{x, y\}^{q-1}. \end{aligned}$$

Logo,

$$y^q - x^q \leq q|y - x| \max\{x, y\}^{q-1}.$$

Utilizaremos o Teorema de Estabilidade para Funções Tame, para indexar a persistência dos pontos finitos no 0-ésimo diagrama de persistência de f e g tal que

$$\Phi^1(f) = \sum_{i=1}^n (d_i^f - b_i^f)^1 \text{ e cada } (d_i^f - b_i^f) \text{ chamaremos de } \phi_i,$$

$$\Phi^1(g) = \sum_{i=1}^n (d_i^g - b_i^g)^1 \text{ e cada } (d_i^g - b_i^g) \text{ chamaremos de } \gamma_i$$

e $|\phi_i - \gamma_i| \leq 2\epsilon$ para todo i , onde $\epsilon = \|f - g\|_\infty$, após possivelmente adicionarmos zeros para que ambos os índices sejam iguais.

Se $x_f \in Dgm_p(f)$, segue que $x_f = (b_i^f, d_i^f)$, assim $\|x_f\| = |d_i^f - b_i^f| = \phi_i$.

Se $y_g \in Dgm_p(g)$, segue que $y_g = (b_i^g, d_i^g)$, assim $\|y_g\| = |d_i^g - b_i^g| = \gamma_i$.

Agora notamos que

$$\Phi^1(f) = \sum_{i=1}^n (d_i^f - b_i^f) = \Phi(f) = \frac{1}{2}Var(f) - (d_{n+1}^f - b_0^f) \leq \frac{1}{2}Var(f)$$

e ainda,

$$Var(f) = \int_0^{2\pi} |f'(s)| ds \leq 2\pi Lip(f),$$

desde que $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ é função lipschitziana com constante 1.

Logo,

$$\Phi(f) = \sum_{i=1}^n (d_i^f - b_i^f) \leq \frac{1}{2}Var(f) \leq \frac{1}{2}2\pi Lip(f) = \pi Lip(f) = \pi.$$

Portanto, $\Phi(f) \leq \pi$. Igualmente concluímos que $\Phi(g) \leq \pi$, por conseguinte, $\Phi(f) + \Phi(g) \leq 2\pi$.

Notamos ainda que

$$\phi_1 + \phi_2 + \dots + \phi_n = \sum_{i=1}^n \phi_i = \Phi(f) \leq \pi,$$

logo $\phi_i \leq \pi$, respectivamente $\gamma_i \leq \pi$, para todo $1 \leq i \leq n$.

Escrevendo $\Delta = \Phi^q(f) - \Phi^q(g)$ obtemos

$$\begin{aligned} |\Delta| &= |\Phi^q(f) - \Phi^q(g)| = \left| \sum_{i=1}^n (d_i^f - b_i^f)^q - \sum_{i=1}^n (d_i^g - b_i^g)^q \right| \\ &= \left| \sum_{i=1}^n \phi_i^q - \sum_{i=1}^n \gamma_i^q \right| = \left| \sum_{i=1}^n \phi_i^q - \gamma_i^q \right| \leq \sum_{i=1}^n |\phi_i^q - \gamma_i^q| \\ &\leq \sum_{i=1}^n q |\phi_i - \gamma_i| \max\{\phi_i, \gamma_i\}^{q-1}. \end{aligned}$$

Então,

$$|\Delta| \leq \sum_{i=1}^n q |\phi_i - \gamma_i| \max\{\phi_i, \gamma_i\}^{q-1}.$$

Contudo, $|\phi_i - \gamma_i| \leq 2\epsilon$, com $\epsilon = \|f - g\|_\infty$, então

$$\begin{aligned} \sum_{i=1}^n q 2\epsilon \max\{\phi_i, \gamma_i\}^{q-1} &= 2\epsilon q \sum_{i=1}^n \max\{\phi_i, \gamma_i\}^{q-1} \\ &= 2\epsilon q \left(\sum_{i=1}^n \max\{\phi_i, \gamma_i\}^{q-2} \max\{\phi_i, \gamma_i\} \right). \end{aligned}$$

Independente de quem for o $\max\{\phi_i, \gamma_i\}$ temos que $\phi_i \leq \pi$ e $\gamma_i \leq \pi$, logo $\max\{\phi_i, \gamma_i\} \leq \pi$, por conseguinte $\max\{\phi_i, \gamma_i\}^{q-2} \leq \pi^{q-2}$.

Daí,

$$2\epsilon q \sum_{i=1}^n \pi^{q-2} \max\{\phi_i, \gamma_i\} = 2q\epsilon\pi^{q-2} \sum_{i=1}^n \max\{\phi_i, \gamma_i\}.$$

Mas, $\phi_1 + \dots + \phi_n = \Phi(f) \leq \pi$ e $\gamma_1 + \dots + \gamma_n = \Phi(g) \leq \pi$.

Para cada i fixado, $\max\{\phi_i, \gamma_i\} \leq \phi_i + \gamma_i$, portanto

$$\begin{aligned} \sum_{i=1}^n \max\{\phi_i, \gamma_i\} &\leq \sum_{i=1}^n \phi_i + \gamma_i = \sum_{i=1}^n \phi_i + \sum_{i=1}^n \gamma_i \\ &= \Phi(f) + \Phi(g) \leq \pi + \pi = 2\pi. \end{aligned}$$

Implicando que

$$2q\epsilon\pi^{q-2} \sum_{i=1}^n \max\{\phi_i, \gamma_i\} \leq 2q\epsilon\pi^{q-2} 2\pi = 4q\pi^{q-1}\epsilon,$$

mas $\epsilon = \|f - g\|_\infty$.

Assim,

$$|\Delta| = |\Phi^q(f) - \Phi^q(g)| \leq 4q\pi^{q-1} \|f - g\|_\infty.$$

□

Para constante $q \geq 2$ o lado direito da desigualdade do teorema nos fornece que no máximo temos uma constante vezes a diferença entre as duas funções segundo a norma L_∞ .

Isso significa que a diferença entre as persistências totais de grau q vai para zero quando a diferença, via norma L_∞ , das funções vai para zero.

O Teorema 6.1 é, portanto, uma afirmação sobre estabilidade para persistência total e, conseqüentemente, para medida de periodicidade:

$$|\mu_1(f) - \mu_1(g)| = |\Phi(f) - \Phi(g)| \leq |\Phi^q(f) - \Phi^q(g)| \leq 4q\pi^{q-1} \|f - g\|_\infty.$$

6.1.7 Notas

As informações para o material nesta seção foram fornecidos pelo trabalho sobre o desenvolvimento biológico de somitos no grupo de Pourquié.

Os dados das séries temporais de microarray do genoma do rato formam a motivação para esse trabalho.

Inicialmente, ele havia sido analisado usando uma variante da análise de Fourier. Por causa de limitações nos padrões discernidos, os mesmos dados foram posteriormente reanalisados utilizando quatro métodos matemáticos feitos especificamente para essa aplicação. Todos os métodos concebidos para reconhecer o gene rítmico de expressão exibida por um pequeno número (menos de 30) de genes verificados que participam do desenvolvimento somite.

Um desses métodos foi a medida de periodicidade descrita nesta seção.

Avaliando todos os sete mil e quinhentos perfis de expressão, cada método gerou uma lista de compatibilidade dos genes (ordenados do mais compatível para o menos compatível) com o ritmo de desenvolvimento do somito.

Essas listas foram então comparadas com base no seu ranking dos genes verificados.

A medida de periodicidade específica utilizada na reanálise foi a μ_2 , preferida sobre as demais escolhas, porque $q = 2$ é a menor potência para a qual sabemos que a medida é estável.

Existe, de fato, prova de que a estabilidade é uma propriedade importante para a tarefa em questão. Isso está ilustrado na seguinte comparação das medidas para $q = 0, 1, 2, 3, 4$.

Para cada q , geramos uma lista ordenada de genes, como antes, e nós construímos uma função escada, $f_q : [0, 1] \rightarrow [0, 1]$, que conta os genes verificados em cada segmento inicial da lista. Em outras palavras, $f_q(x)$ é igual à porcentagem do número total de genes verificados que se enquadram nos primeiros x por cento da lista.

Se os genes verificados forem distribuídos uniformemente entre os outros, então temos uma função escada cujo gráfico está próximo da diagonal.

Por outro lado, se os genes verificados estiverem todos listados primeiro, a função dispara até o valor 1 e permanece assim até o fim.

Em geral, uma medida tem performace melhor do que outra se a primeira função majorar os resultados da segunda.

De fato, evidenciamos uma grande diferença entre medidas instáveis, como μ_0 e μ_1 , contrapondo medidas estáveis, como μ_2 , μ_3 e μ_4 . Em suma, o gráfico de f_0 fica um pouco acima da diagonal indicada por μ_0 , indicando que tal medida fornece uma ordenação melhor somente do que se fosse produzido por um processo aleatório.

Conforme nos aproximamos da função escada ideal, melhorias se tornam mais difíceis de serem feitas, portanto mudar da medida μ_1 para μ_2 é realmente muito signifi-

cante. Além disso, o gráfico para $q = 2, 3$ e 4 é quase indistinguível, um reflexo claro da estabilidade.

A medida de periodicidade é definida em termos de ε -simplificação que expressa perfis. O conceito de ε -simplificação foi introduzido para construir resultados em variedades de dimensão 2, mas a situação ainda é pouco compreendida para variedades de dimensão 3. Nosso trabalho ilustrou para variedades de dimensão 1.

6.2 Elevação para Ancoragem de Proteína

Descreveremos protruções e concavidades de uma superfície utilizando uma função real que foi desenvolvida por um problema de encaixe entre variedades de dimensão 3.

De acordo com o dogma central da biologia, os fios de DNA são transcritos em pedaços de RNA, que são então traduzidos para proteínas.

Transcrição funciona por complementaridade, enquanto a tradução está mais envolvida no processo de transferir uma informação codificada em um alfabeto de quatro nucleótidos para um alfabeto de vinte aminoácidos.

As proteínas são feitas de cordas de aminoácidos, que são altamente variáveis segundo seu comprimento. A princípio, isso sugere um número astronômico de diferentes proteínas possíveis, contudo a natureza aparentemente usa apenas uma pequena fração de talvez, alguns milhares de proteínas.

Uma vez formada uma proteína, ela se dobra em uma forma característica. Esta forma determina sua função, ou seja, como a proteína interagirá e se conectará com outras proteínas.

A interação entre proteínas é um dos processos mais fundamentais em Biologia e mantém a chave para o funcionamento dos sistemas biológicos. As células enviam sinais umas para as outras e criam máquinas que executem as muitas tarefas que tornam a vida possível.

Para entender esses processos, seria maravilhoso se pudéssemos prever quais proteínas interagem com as outras proteínas, simplesmente conhecendo suas forma e as forças exercidas pelos seus átomos.

Este é o problema da ancoragem de proteínas, a predição computacional da interação protéica.

No entanto, esta previsão provou-se notoriamente difícil. Há um debate significativo na comunidade da bioquímica sobre a importância relativa da geometria (forma) e da física (forças), mas é evidente que ambas estão envolvidas.

Pode-se argumentar que a importância relativa da geometria aumenta com o tamanho das moléculas envolvidas. Mas, as proteínas são flexíveis, assim sendo a geometria sozinha não pode prever a correspondência de proteínas na ancoragem. Mesmo assim, a análise geométrica é o primeiro passo.

6.2.1 Curvas no Plano

Suponha $M \subset \mathbb{R}^2$ a imagem de uma imersão suave do círculo.

Defina $F : M \times S^1 \rightarrow \mathbb{R}$ dada por $F(x, u) = \langle x, u \rangle$ a função altura de x na direção u .

Fixando uma direção u , obtemos $f_u : M \rightarrow \mathbb{R}$, $x \mapsto f_u(x) = F(x, u) = \langle x, u \rangle$ função altura na direção u .

Estamos interessados em condições sobre as quais a função altura é Morse. Relembramos que f_u pode falhar em ser Morse por duas razões:

1. Contenha ponto crítico degenerado.
2. Possua dois valores críticos, provenientes de pontos críticos diferentes, iguais.

Um ponto crítico degenerado é modelado por uma família de funções, a qual exemplificaremos por $g_s(t) = t^3 + st$, onde $g'_s(t) = 3t^2 + s$.

A fim de encontrar pontos críticos, variaremos s .

Para $s < 0$, obtemos $-\sqrt{\frac{-s}{3}}$ e $\sqrt{\frac{-s}{3}}$, pontos críticos não degenerados, pois $g''_s(t) = 6t \neq 0$.

Agora, caso $s = 0$, temos um ponto crítico degenerado, pois $g''_s(s) = 0$.

Já para $s > 0$ não existem pontos críticos.

Notamos que conforme s varia de valores negativos à positivos, o par de valores críticos se cancelam. Chamamos esse evento de cancelamento. Ainda podemos considerar o anti-cancelamento se o evento ocorrer na direção contrária, ou seja, se partirmos de s positivo para valores negativos, obteremos o surgimento dos valores críticos.

Nesse caso, varrendo o parâmetro s , corresponde à mover a direção de modo que o ponto crítico deslize na curva.

Um cancelamento ocorre quando dois pontos críticos colidem, o que acontece em um ponto de inflexão. O que nos motiva a assumir que a curva tenha apenas uma quantidade finita de pontos de inflexão e um número finito de linhas que são tangentes a dois ou mais pontos. Tal modelagem é necessária para termos apenas uma quantidade finita de direções quais f_u não é Morse.

Equivalentemente, a família à 1-parâmetro de funções altura passam por uma quantidade finita de cancelamentos e anti-cancelamentos.

6.2.2 Função Altura e Função Elevação

Desejamos que f_u seja função de Morse, pois assim podemos usar persistência estendida para parear os pontos críticos. Esses são os pontos dos quais u é normal à curva.

De fato, $f_u(x) = \langle x, u \rangle$, portanto $f'_u(x) = \langle x, u \rangle$ e ao supormos que o ponto é crítico temos $\langle x, u \rangle = 0$, para todo $x \in \mathbb{M}$, logo u é normal à curva.

Associando a cada x a persistência do par do qual pertence, chamamos esse número de elevação do ponto.

Temos que $f_{-u}(x) = \langle x, -u \rangle = -\langle x, u \rangle = -f_u(x)$. Portanto, pelo teorema de Simetria de Persistência, o pareamento é o mesmo se substituirmos $-u$ por u . Assim, f_u e f_{-u} definem os mesmos valores de elevação para os mesmos pontos.

Em outras palavras, a elevação depende apenas da linha normal à curva dada. Desde que cada ponto de \mathbb{M} tem uma única linha normal, isso define a função elevação, exceto nos pontos em que a função altura não é Morse.

2. A função elevação possui uma quantidade finita de máximos locais.

A segunda condição proíbe curvas de largura constante, tal qual o círculo, por exemplo, que para todo aberto que tome nele, sempre terá um máximo local diferente do aberto anterior, assim possuindo infinitos máximos locais.

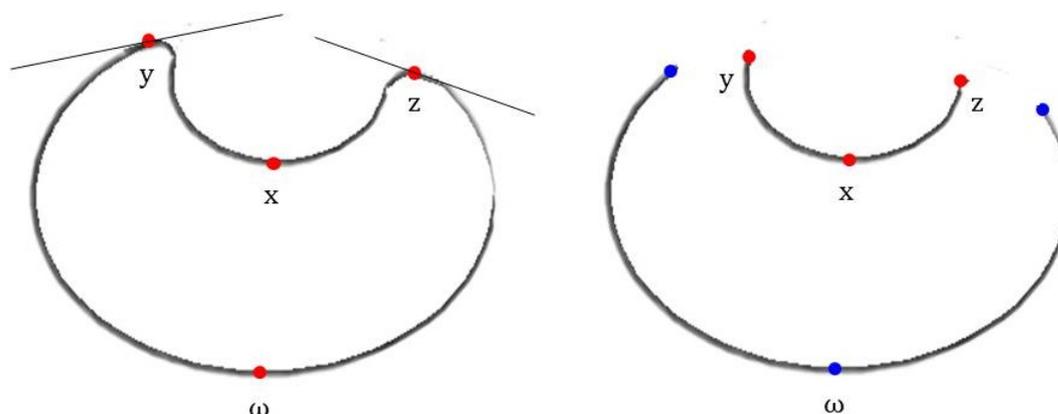


Figura 6.2: À esquerda, as indicações em quais pontos da curva ocorrerá a cirurgia. À direita, cirurgia realizada e os pontos extremos de mesma cor serão identificados, gerando a nova curva N .

Considere a curva acima.

Quando pareamos dois pontos, sendo que nenhum deles é ponto de corte, então chamamos tal par de elevação de uma perna.

Quando um par é formado com um ponto intacto (que não foi cortado), juntamente com outro que foi obtido na identificação e colagem de outros n pontos, então chamaremos o par de n pernas.

Por exemplo, no nosso caso, um par é formado do ponto x juntamente com o ponto que é produto da identificação e colagem do extremo direito de y com o extremo esquerdo de z . Assim, identificamos e colamos dois pontos, y e z e esse é pareado com x , portanto temos um par de duas pernas.

No caso de variedades de dimensão 2, o par de uma perna representa um protrusão, enquanto o par de duas pernas indica concavidade. Essas qualidades atribuídas à variedade, chamamos de características da curva.

6.2.3 Representação da Proteína

A partir desse momento, utilizaremos como apoio o capítulo "Coarse and Reliable Geometric Alignment for Protein Docking", do livro "Pacific Symposium on Biocomputing" [5], que descreve o algoritmo para ancoragem de proteínas, baseando-se nas técnicas vistas previamente. Em suma, o algoritmo de alinhamento grosseiro procura, a partir de dadas as características da proteína, isto é, após a cirurgia para reparar a continuidade da função elevação, por bons encaixes entre duas proteínas.

Representamos cada proteína como um conjunto de quantidade finita de bolas no espaço Euclidiano de dimensão 3, \mathbb{R}^3 .

Especificamente, dadas duas proteínas $A = \{A_1, \dots, A_n\}$ e $B = \{B_1, \dots, B_m\}$, onde $A_i = B(a_i, r_i)$ a bola de centro em $a_i \in \mathbb{R}^3$ e raio $r_i > 0$, bem como $B_j = B(b_j, s_j)$ a bola de centro em $b_j \in \mathbb{R}^3$ e raio $s_j > 0$.

A fim de modelar a ancoragem, fixamos $A \in \mathbb{R}^3$ e descrevemos um algoritmo que encontra um conjunto pequeno de candidatos de transformações para B .

Cada transformação é um movimento rígido μ que produz um candidato à configuração $(A, \mu B)$.

Começamos descrevendo a função pontuação que avalia o ajuste entre duas proteínas.

6.2.4 Função Pontuação

Uma boa função pontuação favorece configurações quase nativas sobre configurações que estão longe de serem nativas.

Definimos

$$d_{i,j} = \|a_i - b_j\| - r_i - s_j,$$

a distância entre as bolas A_i e B_j .

Estabelecemos *cont* e *col* como os quantificadores de contato e colisão, respectivamente,

$$\text{cont}(i, j); \text{col}(i, j) = \begin{cases} -3; 1, & \text{se } d_{i,j} < 0 \\ 1; 0, & \text{se } 0 \leq d_{i,j} < \lambda \\ 0; 0, & \text{se } \lambda < d_{i,j} \end{cases}$$

onde λ é a constante que nos referimos como limitante de contato.

A pontuação de (A, B) é baseada no número total de contatos $\#\text{cont}(A, B) = \sum_{i,j} \text{cont}(i, j)$ e o número total de colisões $\#\text{col}(A, B) = \sum_{i,j} \text{col}(i, j)$.

Frisamos que a configuração (A, B) é válida se $\#\text{col}(A, B) \leq \chi$ onde χ é a constante de limite de colisões, que define o máximo de colisões toleradas para tentarmos acoplar as duas proteínas.

6.2.5 Características

O algoritmo gera movimentos rígidos das características dos conjuntos Φ_A e Φ_B obtidos através da análise da forma de duas proteínas.

Computamos essas características a partir de superfícies aproximadamente suaves que representam essas duas formas.

Seja \mathbb{M} a superfície que representa A .

A função elevação $E : \mathbb{M} \rightarrow \mathbb{R}$ é premissa de nossa definição de característica, como mencionamos anteriormente.

A função elevação associa cada $x \in \mathbb{M}$ com um parceiro canônico, $y \in \mathbb{M}$ de mesma direção normal, $\eta_x = \eta_y$, e definimos $E(x) = |f_u(x) - f_u(y)|$.

Para fins de Ancoragem de Proteína, nos interessamos nas elevações máximas locais pois localmente são mais significantes.

Como já mencionamos, a maioria dos pontos $x \in \mathbb{M}$ tem exatamente um ponto $y \in \mathbb{M}$ que é pareado. Entretanto, a maioria dos máximos surgem em posições onde seu parceiro é ambíguo. Mais especificamente, tipos diferentes de máximo descrevem diferentes tipos de características. Assim, necessitamos das cirurgias para extrair e quantificar essas informações.

A quantidade de pontos pareados após a cirurgia e as colagens apropriadas, geram a característica, Φ_A .

Seja (u, v) um par. O comprimento entre u e v é definido por $\|u - v\|$.

O comprimento e a elevação de uma característica são usados para estimar sua importância, e ambos juntamente com a direção normal são usados para parear as características dos conjuntos Φ_A e Φ_B .

6.2.6 Alinhamento Grosseiro

Dadas duas proteínas A e B , unido de suas respectivas características Φ_A e Φ_B , o algoritmo denominado Alinhamento Grosseiro computa o conjunto de potencial alinhamento grosseiro Γ da seguinte forma:

```

For every  $\alpha \in \Phi_A$  and for every  $\beta \in \Phi_B$  do
  if  $\alpha, \beta$  form a plausible alignment then
     $\mu = \text{Align}(\alpha, \beta)$ 
    compute the contact and collision numbers of  $(A, \mu B)$ ;
    if  $(A, \mu B)$  is valid then add  $\mu$  to  $\Gamma$  endif
  endif
endfor
Sort  $\Gamma$  by contact number

```

A razão por trás do algoritmo é que bons encaixes entre os dados de entrada das proteínas tem características alinhadas, como por exemplo, a protrusão de A se encaixa na concavidade de B , ou vice-versa.

Se parearmos todas as características de A com todas as características de B , nós certamente cobriremos todos bons encaixes.

Por outro lado, a informação que vem com cada característica pode ser usada para discriminar entre os pares e ganhar eficiência ao filtrar os alinhamentos que consideramos importantes ou plausíveis.

Especificamente, introduzimos o filtro de importância que elimina características de Φ_A e de Φ_B cujos comprimentos ou elevações são menores que o limitante.

As características restantes, que sobreviveram ao filtro de importância, formam pares (α, β) que são submetidos ao filtro de plausibilidade que compara α e β com respeito a não serem muito diferentes em comprimento e ainda, se seus tipos são complementares, isto é, se um deles é uma protrusão, então necessariamente, o outro precisa ser uma concavidade.

Para os pares (α, β) que passaram no filtro de importância e no filtro de plausibilidade, computamos um alinhamento de movimento rígido μ como seguinte:

Escrevemos os pontos com uma terceira coordenada, para aprimorar a qualidade da informação retida e extraída.

Se $\alpha \in \Phi_A$ uma característica da proteína A , então $\alpha = (u, v)$ onde u e v são os pontos pareados segundo a função elevação (após o eventual corte e colagem de A).

Passamos a escrever $\alpha = (u, v, \eta_\alpha)$, onde η_α vetor normal à superfície \mathbb{M} . Analogamente, $\beta = (p, q, \eta_\beta)$.

A representação do ponto e de seu vetor normal.

Definimos os vetores binormais $b_\alpha = \eta_\alpha \times \frac{v-u}{\|v-u\|}$ e $b_\beta = \eta_\beta \times \frac{p-q}{\|p-q\|}$ de α e β respectivamente.

Obtemos um movimento rígido μ em três passos.

1. Translade β para que os dois pontos médios coincidam: $\frac{u+v}{2} = \frac{p+q}{2}$.
2. Rotacione β sobre o ponto médio comum para que u, v, p e q sejam colineares.
3. Rotacione β sobre a linha comum para que $b_\alpha = b_\beta$.

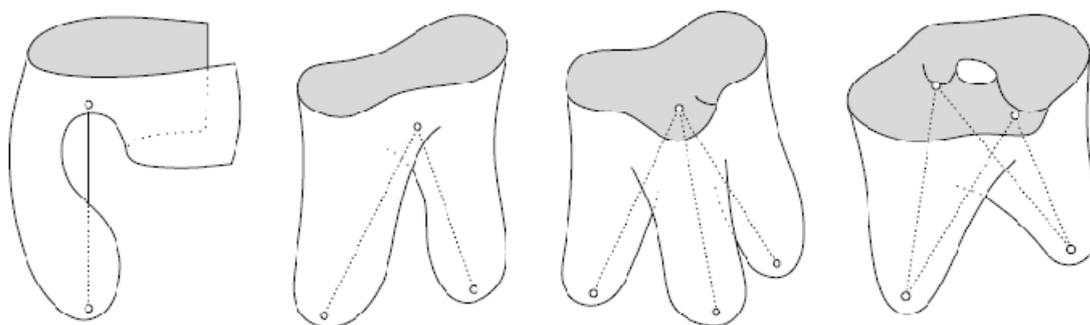


Figura 6.3: Exemplo de representação de proteínas e suas diferentes protrusões e concavidades.

Dado o movimento rígido μ calculamos a função pontuação e o número de colisões utilizando uma estrutura hierárquica de dados que armazena A e B . Sendo n e m o número de bolas nos conjuntos A e B , segue que a complexidade do algoritmo é da ordem de $(n + m)\log(n + m)$.

O algoritmo de Alinhamento Grosseiro foi testado em setenta e quatro problemas de ancoragem de proteínas e em todos os casos exibiu um candidato à configuração nativa ou bem próxima. Além de ser razoavelmente rápido e com possibilidade de melhoras.

Referências

- [1] Azumaya, G., *Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem*. Nagoya Math. J. **1**, (1950). 117–124.
- [2] Chazal, F., de Silva, V., Glisse, M., Oudot, S., *The structure and stability of persistence modules*. arXiv:1207.3674, 2013.
- [3] Chazal, F., de Silva, V., Glisse, M., Oudot, S., *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, 2016.
- [4] Edelsbrunner, H., Harer, J. L., *Computational Topology. An Introduction*. American Mathematical Society, Providence, RI, 2010.
- [5] Wang, Y., Agarwal, P. K., Brown, P., Edelsbrunner, H., Rudolph, J., *Coarse and Reliable Geometric Alignment for Protein Docking*. in: Pacific Symposium on Biocomputing 2005, World Scientific, 2005, pp. 64–75.
- [6] Poincaré, H., *Second complément à l'analysis*. Proc. London Math. Soc, 1900.
- [7] Smith, H. J. S. *On systems of linear indeterminate equations and congruences*. in: Royal Soc. London, 1861.
- [8] Erickson, J. *Computational Topology - Homology* in: <<http://jeffe.cs.illinois.edu/teaching/comptop/2009/notes/homology.pdf>>



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

TERMO DE REPRODUÇÃO XEROGRÁFICA

Autorizo a reprodução xerográfica do presente Trabalho de Conclusão, na íntegra ou em partes, para fins de pesquisa.

São José do Rio Preto, 13 / 03 / 17

Fernanda Japaneiro

Assinatura do autor