

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Distância, na matemática e no cotidiano

Davi Carlos Uehara Approbato

Dissertação de Mestrado do Programa de Mestrado Profissional em
Matemática em Rede Nacional (PROFMAT)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Davi Carlos Uehara Approbato

Distância, na matemática e no cotidiano

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Mestrado Profissional em Matemática em Rede Nacional. *VERSÃO REVISADA*

Área de Concentração: Mestrado Profissional em Matemática em Rede Nacional

Orientador: Prof. Dr. Américo López Gálvez

USP – São Carlos
Agosto de 2019

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

A646d Approbato, Davi Carlos Uehara
 Distância, na matemática e no cotidiano / Davi
 Carlos Uehara Approbato; orientador Américo López
 Gálvez. -- São Carlos, 2019.
 107 p.

 Dissertação (Mestrado - Programa de Pós-Graduação
 em Mestrado Profissional em Matemática em Rede
 Nacional) -- Instituto de Ciências Matemáticas e de
 Computação, Universidade de São Paulo, 2019.

 1. Distância. 2. Espaços Métricos. 3. Distância de
 Hamming. 4. Códigos. 5. Distância genética. I.
 Gálvez, Américo López, orient. II. Título.

Davi Carlos Uehara Approbato

Distance, in math and everyday life

Dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP – in accordance with the requirements of the Professional Master's Program in Mathematics in National Network, for the degree of Master in Science. *FINAL VERSION*

Concentration Area: Professional Master Degree Program in Mathematics in National Network

Advisor: Prof. Dr. Américo López Gálvez

USP – São Carlos
August 2019

Aos meus pais, Carlos e Emiko, e minhas irmãs, Aline e Andressa, que dignamente me apresentaram à importância da família e ao caminho da honestidade e persistência. Aos meus avós, que cuidaram com amor e dedicação e que hoje lá de cima me iluminam e me guiam para o meu sucesso.

AGRADECIMENTOS

Ao Prof. Américo, o meu reconhecimento pela oportunidade de realizar este trabalho ao lado de alguém que transpira sabedoria, paciência e dedicação: meu respeito e admiração.

Meus colegas de profissão, que me ajudaram na construção e orientação. Amigos e familiares, pela força e incentivos e às instituições em que trabalho, pela paciência e compreensão.

“Eduquem as crianças, para que não seja necessário punir os adultos.”
(Pitágoras)

RESUMO

APPROBATO, D. C. U. **Distância, na matemática e no cotidiano**. 2019. 107 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Este trabalho tem como objetivo discutir o conceito formal de distância em matemática, visando depois apresentar exemplos do conceito de distância em situações do dia a dia. Em geral com esse trabalho pretendemos que o leitor menos familiarizado entenda a importância do conceito matemático de distância. Distância é muito mais que o comprimento do segmento entre dois pontos e isso será apresentado em cada capítulo. O assunto foi inspirado pelo livro *Encyclopedia of Distances* Deza Michel Marie (2009), no qual são apresentados, espaços métricos, métricas em várias áreas e aplicações.

No segundo capítulo, será apresentado a definição de espaços métricos. No terceiro capítulo serão apresentados alguns exemplos de métricas. As três primeiras métricas, as mais comuns: métricas euclidiana e máxima em \mathbb{R} e \mathbb{R}^2 . Também serão apresentadas as generalizações de cada uma delas em \mathbb{R}^n . O próximo capítulo, o quarto, é destinado a apresentar o estudo sobre espaços normados, pois por meio desses conceitos pode-se analisar as distâncias entre vetores e matrizes. Veremos que a relevância dessas distâncias auxilia, por exemplo, na compreensão de aproximações de soluções de sistemas.

No capítulo de distância de funções será apresentado um breve comentário sobre a série de Fourier, com relação ao método da aproximação através da decomposição de funções periódicas. Para analisar o quanto as funções trigonométricas estão se aproximando, usa-se o conceito de distância entre funções, as medições são feitas de acordo com as aproximações vão aumentando, essa distância "erro" entre elas tende a zero.

Na teoria dos códigos, é preciso introduzir o conceito de distância entre "palavras", isso permite verificar se o código enviado teve alguma alteração, provocada por uma interferência ou ruídos durante a trajetória. Em algumas situações, o código consegue corrigir e compreender a palavra enviada mesmo tendo sofrido alterações no percurso. Nestes casos, há o estudo da métrica de Hamming.

Já pela métrica de Hausdorff, proposta pelo matemático de mesmo nome, é possível calcular com maior precisão a distância entre conjuntos fechados e limitados. Esta métrica pode ser utilizada em estudos de reconhecimento facial, por exemplo, pois as imagens das faces são transformadas em nuvens de pontos.

Além disso, através do algoritmo de Dijkstra será apresentado a distância entre os vértices de um grafo convexo. Existem várias aplicações de distância entre grafos e uma delas é a questão de

minimizar o custo decorrente do deslocamento entre uma transportadora e o local de entrega por exemplo.

Para finalizar a discussão da importância do consenso de distância, será apresentada uma distância entre genes. Dentro deste tema, o principal cientista foi Thomas Morgan, que por meio de seus estudos conseguiu criar o primeiro mapeamento genético. Com isto, pode relacionar o conceito de distância entre genes à taxa de recombinação gênica.

Finalmente, foi elaborada uma atividade com alunos do ensino médio com o objetivo de analisar os conhecimentos que os estudantes têm sobre distância. Esta atividade também foi importante para que os alunos pudessem compreender a necessidade de formalizar matematicamente este conceito e, principalmente, motivá-los por meio da apresentação de aplicações sobre distância, em diferentes âmbitos.

Palavras-chave: distância, espaços métricos, distância de Hamming, códigos, distância genética.

ABSTRACT

APPROBATO, D. C. U. **Distance, in math and everyday life**. 2019. 107 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Abstract Title: Distance, in mathematics and everyday life

This work has as objective to discuss the formal concept of distance in mathematics, aiming to present examples of the distance concept in everyday situations. In general with this work we want the less familiar reader to understand the importance of the mathematical concept of distance. Distance is much more than the length of the segment between two points and this will be presented in each chapter. The subject was inspired by the book *Encyclopedia of Distances* [Deza Michel Marie \(2009\)](#), in which are presented, metric spaces, metrics in different areas and applications.

In the second chapter, the definition of metric spaces will be presented. In the third chapter some examples of metrics will be presented. The first three metrics, the most common: usual, Euclidean, and maximum metrics in \mathbb{R} and \mathbb{R}^2 . Also the generalizations of each of them were presented in \mathbb{R}^n . The next chapter, the fourth, is intended to show the study on normed spaces, because through these concepts we can analyze the distances between vectors and matrices. We will see that the relevance of these distances helps in the understanding of systems solutions approximation.

In the chapter on distance of functions, a brief comment about Fourier series was presented, regarding the method of approximation through the decomposition of periodic functions. In order to analyze how the trigonometric functions are approaching, the concept of distance between functions is used, the measurements are made as the approximations increase, this distance "error" between them tends to zero.

In code's theory, it is necessary to introduce the concept of distance between "words", this allows to verify if the code had some alteration, caused by an interference or noises during the trajectory. In some situations, the code can correct and understand the sent word even though it has undergone changes in the route. In these cases, there is Hamming's metrics study.

By the Hausdoorf metric, proposed by the mathematician of the same name, it is possible to calculate with more precision the distance between closed and limited sets. This metric can be used in face recognition studies, for example, because face images are transformed into clouds of dots.

Then, through the Dijkstra's algorithm will be presented the distance between the vertices of a convex graphic. There are several applications of distance between graphics and one of them is

the issue of minimizing the cost of moving between a local carrier company and the place of delivery, for example.

To finish the discussion about the importance of distance consensus, the distance between genes will be presented. Within this theme, the main scientist was Thomas Morgan, who through his studies managed to create the first genetic mapping. With this, he was able to relate the concept of distance between genes to the rate of gene recombination.

Finally, an activity was elaborated with high school students with the objective of analyzing students' knowledge about distance. This activity was also important so that the students could understand about a necessity to formalize this concept mathematically and, mainly, to motivate them through the presentation of applications on distance, in different scopes.

Keywords: distance, metric space, Hamming distance, code, genetic distance.

LISTA DE ILUSTRAÇÕES

Figura 1 – Desigualdade triangular	24
Figura 2 – "Esfera na métrica zero - um"	25
Figura 3 – Métrica usual em \mathbb{R}	26
Figura 4 – Métrica Euclidiana em \mathbb{R}^2	27
Figura 5 – Mapa ilustrativo da ilha de Manhattan em Nova Iorque	27
Figura 6 – Exemplo da métrica da soma	28
Figura 7 – Métrica da máxima é igual a 4	29
Figura 8 – Bolas unitárias $B(0, 1)$ com as normas soma, máxima e euclidiana	38
Figura 9 – Gráfico da análise das soluções encontradas com as retas r_1 e r_2	43
Figura 10 – Pontos plotados da tabela	49
Figura 11 – Pontos plotados com característica de uma exponencial	50
Figura 12 – A menor distância entre v e $proj_W(v) : v - proj_W(v)$	51
Figura 13 – Tabela de parametrização dos anos e expectativa de vida	53
Figura 14 – Reta que minimiza os erros	54
Figura 15 – O menor erro quadrático é a reta r_3	57
Figura 16 – A projeção ortogonal de b no espaço coluna de A	59
Figura 17 – Expectativa de vida	60
Figura 18 – Pontos plotados em relação aos dados da população brasileira de 1940 até 2010, considerando uma parametrização nos anos.	61
Figura 19 – População brasileira de 1940 até 2010	63
Figura 20 – Eletrocardiograma	63
Figura 21 – Primeira aproximação $P_1(x)$	64
Figura 22 – Quarta aproximação $P_4(x)$	64
Figura 23 – Décima aproximação $P_{10}(x)$	64
Figura 24 – Vigésima aproximação $P_{20}(x)$	64
Figura 25 – Distância entre as funções f e g : $d(f, g) = 2$	66
Figura 26 – "Distância" das palavras com relação as palavras gato e cachorro	68
Figura 27 – Representação da distância mínima e corretores de erros através do cubo unitário	72
Figura 28 – A distância entre o conjunto A e B é distinto de B e A : $H(A, B) \neq H(B, A)$	75
Figura 29 – Distância usual vale 7 e distância de Hausdorff vale 13	76
Figura 30 – Distância de Hausdorff é de 5.66	76
Figura 31 – As Setes pontes de Königsberg	79

Figura 32 – Vista da atual Kaliningrado	79
Figura 33 – A, B, C e D representam as intersecções entre as pontes e os P_1, P_2, \dots, P_7 as pontes.	80
Figura 34 – Mapa das pontes de Recife	80
Figura 35 – Grafo representando os pontos de entrega da transportadora e os caminhos com os "custo"(pesos) de cada viagem	84
Figura 36 – Iniciando o algoritmo de Dijkstra	84
Figura 37 – Analisando o caminho até o mercado e fechando o vértice do mercado que precede da transportadora	85
Figura 38 – Analisando o caminho até o mercado e fechando o vértice do mercado que precede da transportadora	85
Figura 39 – Analisando os caminhos até a loja	86
Figura 40 – Analisando o caminho até hospital	87
Figura 41 – Analisando o caminho até o clube	87
Figura 42 – Analisando o caminho até o shopping	88
Figura 43 – Localização do DNA e dos genes no cromossomo	90
Figura 44 – Representação de genes em linkage	91
Figura 45 – Meiose sem crossing-over (1) ou com crossing-over(2)	91
Figura 46 – Cruzamentos	92
Figura 47 – Posicionamento dos genes y, v e m	94
Figura 48 – Distância numa esfera e num objeto plano	98
Figura 49 – Distância entre pontos em objetos com curvatura negativa	98
Figura 50 – Medição com barbante e fita métrica	99
Figura 51 – Frases digitadas pelos alunos e corrigidas com a sugestão do corretor ortográfico	99
Figura 52 – Questão 3: parte A	100
Figura 53 – Questão 4: parte A	100
Figura 54 – Questão 4: parte B	101
Figura 55 – Questão 4: parte B	101
Figura 56 – Questionário parte A	104
Figura 57 – Questionário parte B	105

LISTA DE TABELAS

Tabela 1 – Expectativa de vida dos brasileiros de 1940 até 2016	49
Tabela 2 – População brasileira em milhões de pessoas	61
Tabela 3 – Códigos Fonte	68
Tabela 4 – Códigos de Canal	68
Tabela 5 – Códigos do robô	71
Tabela 6 – Início do algoritmo	84
Tabela 7 – Fechando o caminho até o mercado	85
Tabela 8 – Fechando o caminho até o restaurante	86
Tabela 9 – Fechando o caminho até a loja	86
Tabela 10 – Fechando o caminho até hospital	87
Tabela 11 – Fechando o caminho até o clube	88
Tabela 12 – Fechando o caminho até o shopping	88

SUMÁRIO

1	INTRODUÇÃO	21
2	DEFINIÇÃO DE UM ESPAÇO MÉTRICO	23
2.1	Espaços Métricos	23
2.2	Métrica induzida	32
3	ESPAÇOS E DISTÂNCIAS	35
3.1	Espaços vetoriais	35
3.2	Bolas e esferas	38
3.3	Distância entre elementos de um espaço vetorial normado	39
3.4	Distância entre matrizes	40
3.4.1	<i>Normas nos espaço das matrizes</i>	40
3.5	Condicionamento de sistemas lineares	42
3.5.1	<i>Erro absoluto e erro relativo</i>	42
3.5.2	<i>Condicionamento de sistemas lineares</i>	45
3.5.3	<i>Condicionamento numérico de uma Matriz</i>	46
3.6	Distância e o método dos mínimos quadrados	48
3.6.1	<i>Teorema da melhor aproximação</i>	50
3.6.2	<i>Método de aproximação por mínimos quadrados</i>	52
3.7	Distância entre funções	63
3.8	Distância nos códigos	66
3.8.1	<i>Métrica de Hamming</i>	69
3.8.2	<i>Representação da distância mínima e geometricamente</i>	71
3.9	Espaço de Hausdorff	73
3.9.1	<i>Distância de Hausdorff</i>	74
4	GRAFOS	79
4.1	Grafos	79
4.1.1	<i>Distância entre vértices</i>	82
4.1.2	<i>Algoritmos para achar o caminho mais curto</i>	83
5	DISTÂNCIA ENTRE GENES	89
5.1	Distância entre genes	89
5.2	Definições	89

5.3	Ligação gênica ou linkage	90
5.4	Determinação de taxa de recombinação	92
5.4.1	<i>Construção dos mapas gênicos</i>	94
5.4.2	<i>Sob o conceito de espaço métrico e o mapeamento genético</i>	95
6	ATIVIDADE E CONCLUSÃO	97
6.1	Atividade com alunos do ensino médio	97
6.2	Considerações finais	101
7	ANEXOS	103
	REFERÊNCIAS	107

INTRODUÇÃO

As pessoas, na infância, já demonstram uma intuição sobre o conceito de distância, no caso das brincadeiras, por exemplo, como corrida entre amigos, em que o vencedor é o mais rápido e o que consegue percorrer o menor caminho. Outra noção de distância é entre cidades, pois trata-se de medidas positivas.

Talvez seja difícil formalizar ou definir matematicamente o conceito distância com apenas essas intuições, como em relação ao menor caminho ou compreender que é um valor positivo.

É possível observar dois fatos curiosos do cotidiano das pessoas nos quais aparece o conceito de distância, mas que nem sempre é perceptível. Imagine-se uma pessoa que deseja realizar uma pesquisa pela internet e utiliza um site de busca. No próprio site existe um campo para digitar a palavra ou frase sobre o assunto de seu interesse, com a informação do conteúdo desejado, o buscador faz uma varredura pelos sites pelo mundo da internet. Logo, como deve ser realizada a associação entre a palavra digitada e os sites relacionados ao assunto? Existe por trás um conceito de distância entre as palavras? No caso de existir uma distância entre palavras, será que pode-se definir um conceito de distância entre elas ou forma de se medir? Observe outro exemplo em que aparece e não é perceptível por um leigo em matemática. São os reconhecimentos faciais, muito utilizados em smartphones para desbloqueio. Porém, como utilizam a distância para o reconhecimento facial? Se existir um conceito de distância no reconhecimento, como deve ser feita essa medida? Esses questionamentos serão debatidos no decorrer dos capítulos.

Portanto, no desenvolvimento da dissertação, será apresentada a formalização do conceito de distância, definições de espaços métricos e métricas, além de apresentar vários outros tipos de métricas, como a distância entre palavras, pontos sobre o rosto, funções, matrizes, conjuntos, imagem, grafos e genes. Toda teoria de espaços métricos será a base para as definições de métricas, por isso deve-se compreender o conceito matemático inicialmente.

Com base nesses questionamentos em relação as distâncias entre palavras e faces, foi elaborada uma atividade com alunos do ensino médio. O objetivo foi fazer um breve comentário sobre a formalização matemática em relação a distância e estimulá-los através de apresentações de exemplos e medições entre dois pontos em objetos com superfícies irregulares e de formatos distintos. Com isso, é possível criar uma noção de que é preciso ter um instrumento adequado para efetuar uma medição e que a métrica deve ser a mais adequada com o contexto para que possa obter um resultado mais preciso.

Em resumo, o objetivo da dissertação é discutir e mostrar, por meio de exemplos concretos, que existe uma definição matemática e que existem diversas maneiras de se trabalhar com distâncias entre elementos em vários contextos. Além disso, estimular os alunos a pesquisar sobre o assunto e salientar que a matemática não é apenas uma ciência que deve ser mostrada com baseamento teóricos e, sim, prático.

DEFINIÇÃO DE UM ESPAÇO MÉTRICO

Neste capítulo será apresentado a definição de espaço métrico, daremos alguns exemplos e veremos a relevância deste conceito em diversas situações.

2.1 Espaços Métricos

Definição 1. Seja M , um conjunto qualquer não vazio. Uma métrica em M é uma função $d : M \times M \rightarrow \mathbb{R}$, que associa a cada par ordenado $(x, y) \in M \times M$ um número real $d(x, y)$, de forma que satisfazem as seguintes condições:

$$M_1: d(x, x) = 0, \text{ para todo } x \in M.$$

$$M_2: d(x, y) > 0, \text{ para todo } x, y \in M, \text{ então } x \neq y.$$

$$M_3: d(x, y) = d(y, x), \text{ para todo } x, y \in M.$$

$$M_4: d(x, y) \leq d(x, z) + d(z, y), \text{ para todo } x, y, z \in M.$$

Em vista das propriedades apresentadas, se d é uma métrica em M , então o número $d(x, y)$ é chamado de distância entre os elementos x e y em relação à métrica d . Na propriedade (M_1) temos a ideia intuitiva da distância entre o mesmo ponto será nula. Na propriedade (M_2) formaliza que uma distância de um ponto a outro qualquer terá uma distância positiva. A propriedade M_3 , nos diz que de x até y é igual à distância de y até x . A propriedade (M_4) é um pouco menos intuitiva, nela é exigido que a distância entre dois pontos não pode ser maior que a soma das distâncias considerando um terceiro ponto. Na matemática a propriedades (M_1) é conhecida como propriedade não nula, (M_2) como positiva, (M_3) simétrica e (M_4) como desigualdade triangular.

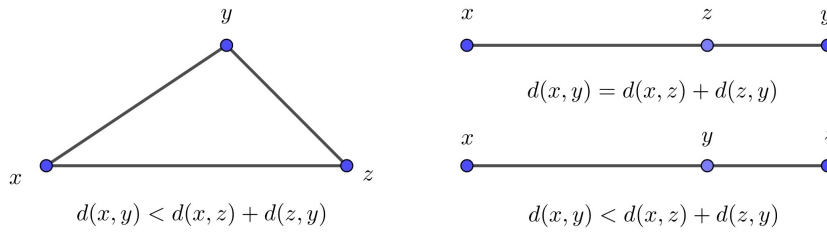


Figura 1 – Desigualdade triangular

Se M e d são como na descrição dada acima, então o par (M, d) é chamado de *espaço métrico* e para calcular a distância entre $x, y \in \mathbb{R}$, basta obter $d(x, y)$.

Com os conceitos definidos de espaço métrico e métrica, apresentaremos a seguir alguns exemplos de métricas. Note-se que os elementos do conjunto M podem ser bastante variados, por exemplo: números, pontos, vetores, matrizes, funções, conjuntos, etc. e em cada um deles existem várias métricas importantes a serem estudadas. Neste trabalho serão apresentadas algumas delas e será mostrado importantes aplicações em situações, bastante relevantes. Devemos observar que para demonstrar que uma dada função é uma métrica precisaremos verificar as condições $(M_1), (M_2), (M_3)$ e (M_4) apresentadas acima.

Exemplo 1: Métrica zero - um.

A métrica *zero - um* permite afirmar que qualquer conjunto não vazio possui uma métrica. A métrica, é conhecida como "*zero - um*", associa a quaisquer dois elementos do conjunto a distância "um", se os elementos forem distintos e "zero" se forem iguais. Dado um conjunto qualquer M não vazio, considere a função $d : M \times M \rightarrow \mathbb{R}$ definida por:

$$d(x, y) = \begin{cases} 1, & \text{se } x \neq y \\ 0, & \text{se } x = y \end{cases} \quad (2.1)$$

A seguir será demonstrada que a função d é uma métrica em M . De fato, serão verificadas as condições da [Definição 1](#) sejam $x, y, z \in \mathbb{R}$.

Se $x = y$, temos por definição da função d que $d(x, x) = 0$. Verificando a condição M_1 . Pela definição da função d , se $x \neq y$, então $d(x, y) > 0$, e em qualquer caso, tem que $d(x, y) = d(y, x)$.

A demonstração da desigualdade triangular será dividida em dois casos:

1º Caso: É imediato se $x = y$, pois $d(x, y) = 0 \leq d(x, z) + d(z, y)$ para qualquer $z \in M$.

2º Caso: Se $x \neq y$, por definição da função d dada acima temos que $d(x, y) = 1 \leq d(x, z) + d(z, y)$ para qualquer $z \in M$, pois se $z = x$, então $d(z, y) = 1$, se $z \neq x$ então $d(x, z) = 1$.

Temos verificado as condições necessárias para que d seja uma métrica.

Exemplo 2: Dado o conjunto $M = \{a, b, c, d\}$ a métrica "zero - um" para calcular a distância entre os elementos de M .

Pela definição dada, temos a distância entre os elementos do conjunto M .

Distância entre os elementos de M serão $d(a, d) = 1$, $d(b, c) = 1$, $d(a, c) = 1$ e $d(a, a) = d(b, b) = d(c, c) = d(d, d) = 0$.

Observa-se que essa métrica é pouco usual embora é uma maneira de se medir distância em qualquer conjunto.

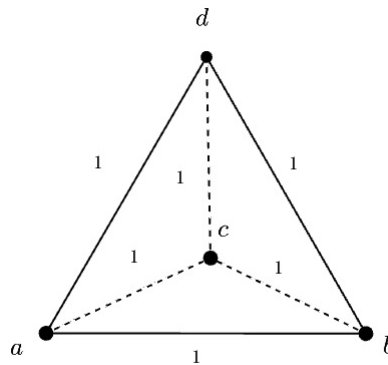


Figura 2 – "Esfera na métrica zero - um"

Exemplo 3: Métrica usual em \mathbb{R} .

No ensino fundamental e médio os alunos aprendem a calcular a distância entre dois pontos em \mathbb{R} e em \mathbb{R}^2 , porém não é estudado estudam a definição da métrica, apenas aprendem a calcular a distância. Por isso, veremos que a distância vista pelos alunos na sala de aula é definida por uma métrica.

Defina a função $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ por:

$$d(x, y) = |x - y|. \quad (2.2)$$

Deve-se verificar que a função d satisfaz as condições da [Definição 1](#) para garantir que d seja uma métrica no sentido definido acima. As condições M_1, M_2 e M_3 são imediatas pois $d(x, x) = |x - x| = 0$, $d(x, y) = |x - y| > 0$, se $x \neq y$, e $d(x, y) = |x - y| = |y - x| = d(y, x)$, pela propriedade do valor absoluto. Para verificar M_4 adicionamos e subtraímos z na definição, mas precisamente

$$|x - y| = |(x - z) + (z - y)|.$$

Pela desigualdade triangular, $|x + y| \leq |x| + |y|$, tem-se:

$$|x - y| \leq |x - z| + |z - y|$$

isto é

$$d(x, y) \leq d(x, z) + d(z, y).$$

Portanto $d(x, y) = |x - y|$ define uma métrica em \mathbb{R} .

Como caso particular vemos que a distância entre os pontos $x = 6$ e $y = 2$ e $x = -2$ e $y = -4$ é calculada com a métrica usual dada por

$$d(6, 2) = |6 - 2| = 4.$$

$$d(-2, -4) = |-2 - (-4)| = |-2 + 4| = 2.$$

Geometricamente, tem-se

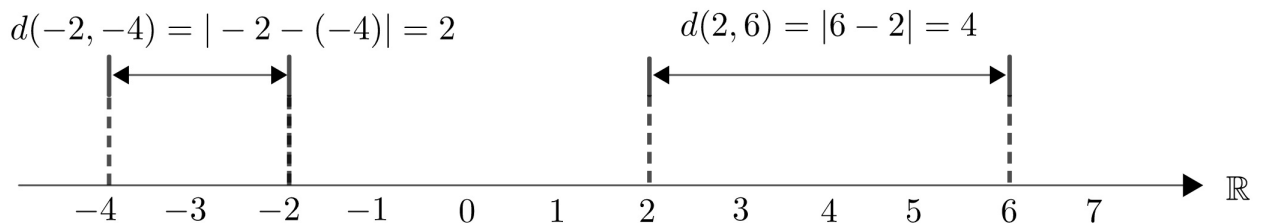


Figura 3 – Métrica usual em \mathbb{R}

Exemplo 4: Métrica usual em \mathbb{R}^2

O conceito de distância entre pontos no plano cartesiano é ensinada aos alunos no ensino médio, porém, os estudantes não relacionam com o conceito de métrica. Veremos isto agora.

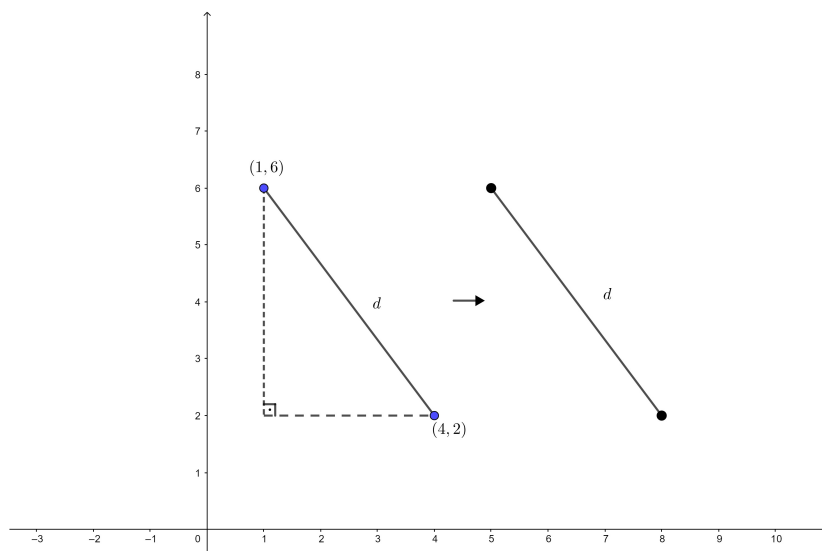
Dados dois pontos $x = (x_1, x_2)$ e $y = (y_1, y_2) \in \mathbb{R}^2$, considere a função $d : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

Prova-se de modo similar ao exemplo anterior que d é uma métrica.

Como caso particular, a distância Euclidiana entre os pontos $x = (1, 6)$ e $y = (4, 2)$ de \mathbb{R}^2 será

$$d((1, 6), (4, 2)) = \sqrt{(1 - 4)^2 + (6 - 2)^2} = 5.$$

Figura 4 – Métrica Euclidiana em \mathbb{R}^2

Note-se que a distância Euclidiana é o comprimento de um segmento de reta de extremidade x e y .

Exemplo 5: Métrica da soma em \mathbb{R}^2 .

Com a métrica da soma podemos calcular a distância entre dois pontos no plano cartesiano, em que soma se as distâncias horizontais e verticais, ela também é conhecida como métrica taxi ou métrica de Manhattan, pois faz alusão ao formato quadriculado da maior parte das ruas na ilha de *Manhattan* e em que se calcula distância percorrida de um taxi.



Figura 5 – Mapa ilustrativo da ilha de Manhattan em Nova Iorque

Considere a função $d : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ e definida por:

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2| \quad (2.3)$$

para cada $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$.

Veamos um exemplo particular. A distância entre os pontos $x = (1, 6)$ e $y = (4, 2)$ calculada com a métrica da soma será

$$d((1, 6), (4, 2)) = |1 - 4| + |6 - 2| = 7.$$

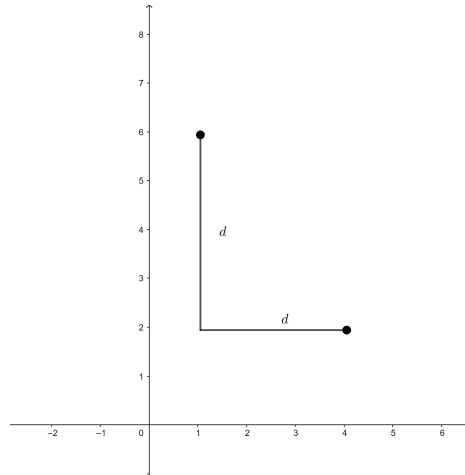


Figura 6 – Exemplo da métrica da soma

Exemplo 6: Métrica máxima em \mathbb{R}^2

Defina a função $d : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ por:

$$d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}. \quad (2.4)$$

para cada $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$.

Vejamos que a função d acima definida satisfaz às condições exigidas de métrica. Sejam $x = (x_1, x_2), y = (y_1, y_2)$ e $z = (z_1, z_2)$ elementos arbitrários em \mathbb{R}^2 .

Como $d(x, x) = \max\{|x_1 - x_1|, |x_2 - x_2|\}$, então tem-se $d(x, x) = 0$, logo M_1 é satisfeito. Se $x \neq y$, então segue da definição do valor absoluto $d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\} > 0$, verificando assim a condição M_2 . A condição M_3 segue imediatamente da definição da função d e propriedade do valor absoluto já que $d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$ e $d(y, x) = \max\{|y_1 - x_1|, |y_2 - x_2|\}$, logo $d(x, y) = d(y, x)$.

Para provar a condição M_4 deve adicionar e subtrair o elemento z .

$$\begin{aligned} d(x, y) &= \max\{|x_1 - z_1 + z_1 - y_1|, |x_2 - z_2 + z_2 - y_2|\} \\ &\leq \max\{|x_1 - z_1| + |z_1 - y_1|, |x_2 - z_2| + |z_2 - y_2|\} \\ &\leq \max\{|x_1 - z_1|, |x_2 - z_2|\} + \max\{|z_1 - y_1|, |z_2 - y_2|\} \end{aligned}$$

Portanto, $d(x, y) \leq d(x, z) + d(z, y)$.

Como exemplo particular podemos ver que a distância entre os pontos $x = (1, 6)$ e $y = (4, 2)$ calculada com a métrica da máxima será

$$d((1, 6), (4, 2)) = \max\{|1 - 4|, |6 - 2|\} = 4.$$

As métricas mostradas anteriormente estão definidas em \mathbb{R} e \mathbb{R}^2 , não é difícil imaginar que elas podem ser também definidas no espaço \mathbb{R}^n .

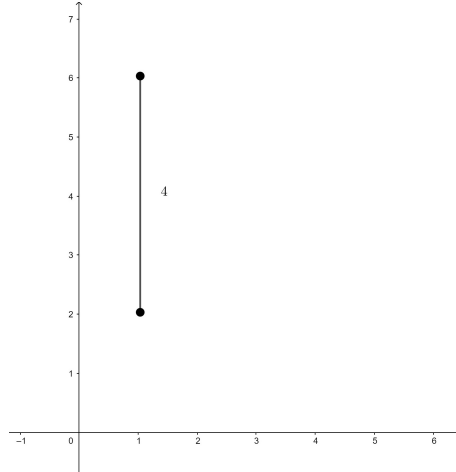


Figura 7 – Métrica da máxima é igual a 4

Lembremos que $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n); x_i \in \mathbb{R} \text{ para } i = 1, 2, \dots, n\}$. A métrica Euclidiana é definida pela distância usual entre dois pontos. Já as métricas da soma e métrica do máximo parecem não ser tão naturais numa breve observação, porém, no ponto de vista prático, em muitos casos visivelmente mais vantajosas.

Exemplo 7: Métrica euclidiana em \mathbb{R}^n

Seja $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ a função definida por:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.5)$$

para cada $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$. A função d acima definida é uma métrica. Sejam $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ pontos em \mathbb{R}^n . Se $x = y$, então $d(x, x) = \sqrt{(x_1 - x_1)^2 + \dots + (x_n - x_n)^2} = 0$ provando assim a condição M_1 . Por outro lado, se $x \neq y$, isto é, se existe $1 \leq i \leq n$ tal que $x_i \neq y_i$, então pela propriedade do valor absoluto, temos que a condição M_2 é imediato pois $d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \geq \sqrt{(x_i - y_i)^2} = |x_i - y_i| > 0$.

Como $(x_i - y_i)^2 = (y_i - x_i)^2$ para $1 \leq i \leq n$, então

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \sqrt{(y_1 - x_1)^2 + \dots + (y_n - x_n)^2} = d(y, x).$$

logo a condição M_3 também é válida.

Para verificar a desigualdade triangular, lembremos a desigualdade de Cauchy-Schwarz para cada (x_1, \dots, x_n) e (y_1, \dots, y_n) em \mathbb{R}^n é verdade que

$$\sum_{i=1}^n |x_i \cdot y_i| \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}.$$

De fato, como

$$p^2 - 2pq + q^2 = (p - q)^2 \geq 0 \text{ então, } 2pq \leq p^2 + q^2$$

é verdadeira para todo p e $q \in \mathbb{R}$.

Agora, considerando $\frac{|x_i|}{a}$ e $\frac{|y_i|}{b}$ na desigualdade anterior no lugar de p e q respectivamente, onde

$$a = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \text{ e } b = \sqrt{y_1^2 + y_2^2 + \dots + y_n^2}$$

obtemos a relação de desigualdade

$$2 \frac{|x_i|}{a} \cdot \frac{|y_i|}{b} \leq \frac{x_i^2}{a^2} + \frac{y_i^2}{b^2}$$

para todo $i = 1, \dots, n$.

Somando em relação ao índice i obtemos

$$\frac{2}{ab} \sum_{i=1}^n |x_i y_i| \leq 1 + 1 = 2$$

escrito de outro modo

$$\sum_{i=1}^n |x_i y_i| \leq a \cdot b$$

de onde ao substituir os valores de a e b concluímos que

$$\sum_{i=1}^n |x_i y_i| \leq \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \cdot \sqrt{y_1^2 + y_2^2 + \dots + y_n^2}$$

mas precisamente

$$\sum_{i=1}^n |x_i \cdot y_i| \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}$$

Provando, assim, a desigualdade de Cauchy-Schwarz. Agora estamos em condições de verificar a condição M_4 para a função d . Como

$$\begin{aligned} [d(x, y)]^2 &= \sum_{i=1}^n (x_i - y_i)^2 = \sum_{i=1}^n (x_i - z_i + z_i - y_i)^2 \\ &= \sum_{i=1}^n (x_i - z_i)^2 + 2 \cdot \sum_{i=1}^n (x_i - z_i) \cdot (z_i - y_i) + \sum_{i=1}^n (z_i - y_i)^2. \end{aligned}$$

Aplicando a desigualdade de Cauchy-Schwarz, podemos concluir que

$$\begin{aligned}
[d(x,y)]^2 &\leq \sum_{i=1}^n (x_i - z_i)^2 + 2 \cdot \sqrt{\sum_{i=1}^n (x_i - z_i)^2} \cdot \sqrt{\sum_{i=1}^n (z_i - y_i)^2} + \sum_{i=1}^n (z_i - y_i)^2 \\
&= \left[\sqrt{\sum_{i=1}^n (x_i - z_i)^2} + \sqrt{\sum_{i=1}^n (z_i - y_i)^2} \right]^2 \\
&= [d(x,z) + d(z,y)]^2.
\end{aligned}$$

de que, $d(x,y) \leq d(x,z) + d(z,y)$, logo M_4 é válida.

Note-se que esta métrica é a métrica usual em \mathbb{R}^n tendo em vista que $d(x,y)$ é o comprimento do segmento de reta de extremidade x e y .

Exemplo 8: Métrica da soma em \mathbb{R}^n .

Considere a função $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, definida por

$$d(x,y) = \sum_{i=1}^n |x_i - y_i| \quad (2.6)$$

para cada $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ em \mathbb{R}^n .

Vejam os que a função d satisfaz as condições para garantir que é uma métrica. Tem-se $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ e $z = (z_1, z_2, \dots, z_n)$ elementos em \mathbb{R}^n . Claramente as condições M_1, M_2 e M_3 são imediatas tendo em vista que $d(x,x) = \sum_{i=1}^n |x_i - x_i| = 0$, $d(x,y) = \sum_{i=1}^n |x_i - y_i| \geq |x_{i_0} - y_{i_0}| > 0$, se $x \neq y$ com $x_{i_0} - y_{i_0} \neq 0$, e $d(x,y) = \sum_{i=1}^n |x_i - y_i| = \sum_{i=1}^n |y_i - x_i| = d(y,x)$.

Vejam agora que a condição M_4 também é válida. Como $x_1 - y_1 = x_1 - z_1 + z_1 - y_1$, assim até n -ésima coordenada, $x_n - y_n = x_n - z_n + z_n - y_n$, aplicando o valor absoluto e a desigualdade triangular, tem-se o seguinte

$$\begin{aligned}
|x_1 - y_1| &\leq |x_1 - z_1| + |z_1 - y_1| \\
|x_2 - y_2| &\leq |x_2 - z_2| + |z_2 - y_2| \\
&\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\
|x_n - y_n| &\leq |x_n - z_n| + |z_n - y_n|
\end{aligned}$$

Agora, somando as n desigualdades, obtém que a condição M_4 , $d(x,y) \leq d(x,z) + d(z,y)$ é válida.

Observa-se que esta métrica é a generalização da métrica soma apresentado em \mathbb{R}^2 .

Exemplo 9: Métrica máxima em \mathbb{R}^n .

Considere a função $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, definida por

$$d(x, y) = \max_{(1 \leq i \leq n)} \{ |x_i - y_i| \}$$

para cada $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ em \mathbb{R}^n

Como nos exemplos anteriores, vejamos que esta função satisfaz as condições M_1, M_2, M_3 e M_4 . Exigidas na [Definição 1](#).

Como $d(x, x) = \max_{(1 \leq i \leq n)} \{ |x_i - x_i| \} = 0$, então M_1 é válida.

As condições M_2 e M_3 também são imediatas pois $d(x, y) = \max_{(1 \leq i \leq n)} \{ |x_i - y_i| \} > 0$, para todo $x \neq y$ e $d(x, y) = \max_{(1 \leq i \leq n)} \{ |x_i - y_i| \} = \max_{(1 \leq i \leq n)} \{ |y_i - x_i| \} = d(y, x)$.

Para provar a condição M_4 , sejam $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ e $z = (z_1, z_2, \dots, z_n)$ pontos em \mathbb{R}^n .

Pela definição de d , temos

$$\begin{aligned} \max_i \{ |x_i - y_i| \} &= |x_{i_0} - y_{i_0}|, \text{ para algum } 1 \leq i_0 \leq n; \\ \max_i \{ |x_i - z_i| \} &= |x_{j_0} - z_{j_0}|, \text{ para algum } 1 \leq j_0 \leq n; \\ &e \\ \max_i \{ |z_i - y_i| \} &= |z_{k_0} - y_{k_0}|, \text{ para algum } 1 \leq k_0 \leq n. \end{aligned}$$

Segue da desigualdade triangular que $|x_{i_0} - y_{i_0}| \leq |x_{i_0} - z_{i_0}| + |z_{i_0} - y_{i_0}|$

Como

$$|z_{i_0} - y_{i_0}| \leq |z_{k_0} - y_{k_0}|$$

e

$$|x_{i_0} - z_{i_0}| \leq |x_{j_0} - y_{j_0}|$$

por serem os máximos, obtém-se

$$|x_{i_0} - y_{i_0}| \leq |x_{i_0} - z_{i_0}| + |z_{i_0} - y_{i_0}| \leq |x_{j_0} - y_{j_0}| + |z_{k_0} - y_{k_0}|$$

Isto é, $d(x, y) \leq d(x, z) + d(z, y)$.

2.2 Métrica induzida

Na métrica induzida, veremos como induzir uma métrica dada num subconjunto. Seja (M, d) um espaço métrico, em que M é um conjunto qualquer não vazio e S um subconjunto de M . Note-se que podemos considerar em S , de uma maneira natural, a métrica d . Para isso basta considerar a restrição a métrica d ao conjunto $S \times S$, isto é, usar entre os elementos de S a mesma

"distância" que os elementos de M possuem. Quando isso ocorre, dizemos que a métrica em S é a *métrica induzida* de M .

Essa ideia nos permite ter inúmeras variedades de exemplos de espaços métricos, considerando os diversos subconjuntos de um espaço métrico dado. Veja a seguir um exemplo de *métrica induzida*.

Vejamos um exemplo de métrica induzida.

Considere o espaço métrico (\mathbb{R}^2, d) , onde d é a métrica usual e considere \mathbb{R} como subconjunto de \mathbb{R}^2 , ou seja, identificamos \mathbb{R} como o subconjunto, $\mathbb{R}^2 = \{(x, 0); x \in \mathbb{R}\}$. Para cada $x = (x, 0)$ e $y = (y, 0)$. Segundo a identificação acima, temos que

$$d(x, y) = \sqrt{(x-y)^2 + (0-0)^2} = \sqrt{(x-y)^2} = |x-y|$$

isto é, a métrica usual de \mathbb{R} é a métrica induzida pela métrica usual de \mathbb{R}^2 .

Exemplo 10: Métrica p-ádico.

Fixe um número primo p . É possível provar que cada número $x \in \mathbb{Q}$, $x \neq 0$, pode ser escrito do modo único, como

$$x = p^r \cdot \frac{a}{b}$$

$r, a, b \in \mathbb{Z}$ e $b \neq 0$, com a e b inteiros não múltiplos de p .

Vejamos alguns exemplos.

Considerando $p = 3$, temos $x = 24$, se escreve como $3^1 \cdot \frac{8}{1}$. Já $x = 10$ se escreve da seguinte forma, $3^0 \cdot 10$ e $x = \frac{12}{5}$ como $3 \cdot \frac{4}{5}$.

Os fatos descritos acima permitem, definir a função $v: \mathbb{Q} - \{0\} \rightarrow \mathbb{Z}$ por $v(x) = r$, para cada $x = p^r \cdot \frac{a}{b}$.

Considerando estas notações podemos definir a função

$$d_p: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{R}$$

por

$$d_p(x, y) = p^{-v(x-y)} \text{ se } x \neq y$$

e por

$$d_p(x, y) = 0 \text{ se } x = y.$$

É possível provar que a função d_p assim definida é uma métrica em \mathbb{Q} .

Vejamos um exemplo

Considere $p = 3$ $x = \frac{9}{4}$ e $y = \frac{3}{2}$ então

$$d_3 \left(\frac{9}{4}, \frac{3}{2} \right) = 3^{-1}$$

$$\text{pois } \frac{9}{4} - \frac{3}{2} = \frac{3}{4} = 3^1 \cdot \frac{1}{4}$$

Exemplo 11: Métrica p -ádica em \mathbb{N}

Fixado o primo p , defina $d_p : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ como sendo uma função que associa a cada par $(m, n) \in \mathbb{N} \times \mathbb{N}$ o número $p^{-\alpha}$, onde α é o maior inteiro tal que p^α divide $(m - n)$, se $m \neq n$ e igual a zero se $m = n$.

É possível provar que esta aplicação é uma métrica em \mathbb{N} , mas ainda, que é a métrica p -ádica em \mathbb{Q} induzida em \mathbb{N}

ESPAÇOS E DISTÂNCIAS

Foi visto no capítulo anterior várias métricas definidas no espaço \mathbb{R}^n . Se considerarmos os espaços das matrizes, polinômios, funções e outros, é possível definir uma métrica de tal modo que permita medir a distância entre elementos desses espaços? E qual a importância de obter a distância entre esses elementos?

Vários projetos de engenharias são solucionados através de problemas que envolvem uma grande quantidade de cálculos e isto naturalmente envolve sistemas lineares com uma gama enorme de sistemas de equações e matrizes, então não é difícil imaginar que sejam utilizados softwares matemáticos para obter soluções (numéricas) e que alguns valores acabam sendo arredondados, logo as soluções numéricas diferem da exata, portanto é necessário analisar se as soluções obtidas através dos arredondamentos estão ou não próximas da solução exata. Para entender e responder esta importante questão, a noção de distância entre matrizes será fundamental.

Por outro lado as funções trigonométricas e polinomiais são de grande importância para aplicações práticas na física e na engenharia. Por exemplo, estudos da propagação de calor, circuitos elétricos, ondas sonoras, sinais musicais e outros precisam da noção de distância para poder falar de aproximações.

O conceito de espaço vetorial é importante, pois permite estudar as matrizes, polinômios e funções dentro de uma única teoria. Logo se queremos definir métricas em cada um destes conjuntos, basta apresentar métricas em espaços vetoriais.

3.1 Espaços vetoriais

Iniciamos lembrando o conceito de espaço vetorial.

Definição 2. Seja V um conjunto no qual duas operações, chamadas *adição* e *multiplicação por escalar*, são definidas. Se u e v estão em V , a soma de u e v é denotada por $u + v$, e se c é uma

escalar, o múltiplo escalar de v por c é denotado por cv . Se os axiomas a seguir são verdadeiros para todo u, v e w em V e para todos os escalares c e d , então V é chamado **espaço vetorial** e seus elementos são chamados de **vetores**

(A1): $u + v$ Fechado sob adição

(A2): $u + v = v + u$ Comutatividade

(A3): $(u + v) + w = u + (v + w)$ Associatividade

(A4): Existe um elemento 0 em V , denotado como *vetor nulo*, tal que $u + 0 = u$

(A5): Para cada u em V , existe um elemento $-u$ em V tal que $U + (-u) = 0$

(A6): cv está em V para todo $v \in V$ e c escalar

(A7): $c(w + v) = cw + cv$ para todo $w, v \in V$ e c escalar

(A8): $(c + d)v = cv + dv$ para todo $v \in V$ e c e d escalares

(A9): $c(dv) = (cd)v$, para todo $v \in V$ e c e d escalares.

(A10): $1v = v$ para todo $v \in V$

Vale lembrar que \mathbb{R}^n , os conjuntos de matrizes $M(m, n)$ de m linhas e n colunas, polinômios $P(n)$ de grau menor ou igual a n e funções f definidas num intervalo são exemplos de espaços vetoriais com operações usuais de adição e multiplicação por um escalar.

Na matemática, sistemas de equações lineares é a base e uma parte fundamental da álgebra linear que tem várias aplicações em diversas ciências como física, química, biologia, engenharia e outras. É fácil encontrar diversos exemplos de problemas que são representados por sistema lineares de equações como circuitos elétricos, balanceamento de equações químicas, controle de fluxo de veículos, cálculo de estruturas de pontes, torres, edifícios, aeronaves e outros. Porém, muitos desses problemas recaem em sistemas lineares com uma quantidade elevada de equações. Com efeito, a resolução desses sistemas não é trivial, então para obter o vetor solução, serão utilizados os métodos iterativos,¹ os quais geram uma procedimento em que gera uma sequência de soluções aproximadas, que vão melhorando conforme as iterações são executadas. Mas para discutir o método em que a sequência de vetores converge para uma solução do sistema, deve-se mensurar a distância entre esses vetores, pois para efeito de aproximação é fundamental utilizar o conceito de distância. Portanto, para definir essa distância em espaços vetoriais será utilizada a noção de norma.

Definição 3. Seja V um espaço vetorial. Uma **norma** em V é uma aplicação $\| \cdot \| : V \rightarrow \mathbb{R}$, que associa a cada vetor $u \in V$ um escalar em \mathbb{R} chamado norma de u e denotado por $\|u\|$, de modo que as seguintes propriedades sejam satisfeitas.

¹ Métodos iterativos(SPERANDIO; MENDES; SILVA, 2003)

(N1) Se $u \neq 0$ então $\|u\| > 0$; para todo $u \in V$

(N2) $\|\lambda \cdot u\| = |\lambda| \cdot \|u\|$; para todo $\lambda \in \mathbb{R}$ e $u \in V$

(N3) $\|u + v\| \leq \|u\| + \|v\|$. para todo $u, v \in V$

Um espaço vetorial, V que admite uma norma $\|\cdot\|$, é chamado de espaço vetorial normado e é denotado pelo par $(V, \|\cdot\|)$. Veja alguns exemplos.

Norma da soma em \mathbb{R}^n

Defina $\|\cdot\|_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ como segue para cada $v = (v_1, v_2, \dots, v_n)$,

$$\|v\|_1 = |v_1| + |v_2| + \dots + |v_n|$$

em outras palavras, $\|u\|_1$ é a soma dos valores absolutos de suas componentes.

Norma do máximo em \mathbb{R}^n

A função $\|\cdot\|_{\max} : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\|v\|_{\max} = \max\{|v_1|, \dots, |v_n|\}$$

para cada $v = (v_1, v_2, \dots, v_n)$.

A norma do máximo $\|v\|_{\max}$ é o maior valor entre os valores absolutos de suas componentes.

Norma Euclidiana em \mathbb{R}^n

A função $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por

$$\|v\| = \sqrt{|v_1|^2 + |v_2|^2 + \dots + |v_n|^2}$$

para cada $v = (v_1, v_2, \dots, v_n)$ é chamada de norma Euclidiana.

Mas geralmente para cada p natural positivo, define-se a norma

$$\|v\|_p = (|v_1|^p + \dots + |v_n|^p)^{\frac{1}{p}}$$

Em particular, para $p = 2$, temos a norma Euclidiana.

$$\|v\|_2 = (|v_1|^2 + \dots + |v_n|^2)^{\frac{1}{2}}$$

Vejamos um exemplo em \mathbb{R}^3 . Se $v = (-1, 2, 3)$ então

$$\begin{aligned}\|v\|_1 &= |-1| + |2| + |3| = 6 \\ \|v\|_{max} &= \max\{|-1|, |2|, |3|\} = 3 \\ \|v\|_2 &= \sqrt{|(-1)|^2 + |2|^2 + |3|^2} = \sqrt{14}\end{aligned}$$

3.2 Bolas e esferas

Na seção anterior foi lembrado o conceito de norma, agora veremos um assunto, no mínimo curioso, é o de bolas e esferas.

Após compreender sobre normas no espaço vetorial normado, o próximo conteúdo é, de fato, no mínimo curioso, são os estudos de bolas e esferas. Pela nomenclatura imagina-se algo com um formato de uma bola de futebol, uma laranja ou outros objetos que se assemelham a uma esfera. Porém, não necessariamente terá esse tipo de formato, mas tudo depende da métrica utilizada.

Dados um elemento a no espaço normado vetorial M e um número real $r > 0$. Define-se uma *bola aberta* de centro a e raio r , como o conjunto $B(a, r)$ de todos os elementos x em M , tais que a norma do vetor $x - a$ seja menor do que r , isto é, $B(a, r) = \{x \in M; \|x - a\| < r\}$

Uma *bola fechada* de centro a e raio r é definido como o conjunto de elementos x menor ou igual a r . Mas precisamente tais que a norma do vetor $x - a$ seja menor que r , $B[a, r] = \{x \in M; \|x - a\| \leq r\}$

Uma *esfera* de centro a e raio r é o conjunto $S(a, r)$ formado pelos elementos x em M tais que a norma do vetor $x - a$ seja sempre r . Isto é, $S(a, r) = \{x \in M; \|x - a\| = r\}$

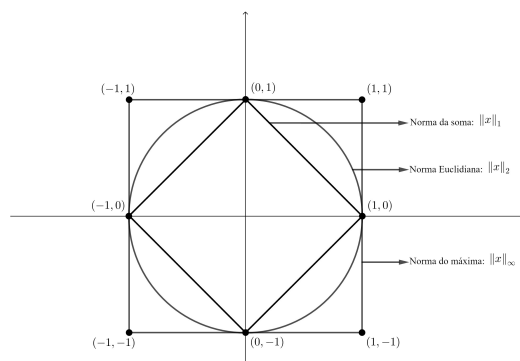


Figura 8 – Bolas unitárias $B(0, 1)$ com as normas soma, máxima e euclidiana

3.3 Distância entre elementos de um espaço vetorial normado

Seja $(V, \|\cdot\|)$ um espaço vetorial normado. Para cada $u, v \in V$ sabe-se que $u - v \in V$ logo podemos definir uma função distância em V como segue, induzida pela norma $\|\cdot\|$. Seja $d : V \times V \rightarrow \mathbb{R}$, definida por:

$$d(u, v) = \|u - v\|$$

Vejamos que a função d satisfaz as condições de métrica. Sejam $u, v, w \in V$, elementos em V . Claramente as condições M_1, M_2, M_3 e M_4 são satisfeitas. Pois $d(u, u) = \|u - u\| = 0$ a condição M_1 , na condição M_2 temos $d(u, v) = \|u - v\| > 0$ para $u \neq v$, já na condição M_3 , obtemos $d(u, v) = \|u - v\| = \|(-1)(v - u)\| = |-1| \|v - u\| = d(v, u)$ e a M_4 temos a condição de $d(u, v) = \|u - v\| = \|u - w + w - v\| \leq \|u - w\| + \|w - v\|$ portanto, $d(u, v) \leq d(u, w) + d(w, v)$. Provaremos assim que $d : V \times V \rightarrow \mathbb{R}$ é uma métrica em V , e dizemos que ela é induzida pela norma $\|\cdot\|$.

Observa-se que, com esta métrica, a definição de bolas e esferas pode ser escrita como:

$$B(a, r) = \{x \in M; d(x, a) < r\}$$

$$B[a, r] = \{x \in M; d(x, a) \leq r\}$$

$$S(a, r) = \{x \in M; d(x, a) = r\}$$

Vejamos alguns exemplos de normas.

As normas definidas anteriormente nos permite definir as métricas d_1, d_∞ e d_2 em \mathbb{R}^n como segue. Dados $u = (u_1, u_2, \dots, u_n)$ e $v = (v_1, v_2, \dots, v_n)$ em \mathbb{R}^n

$$d_1(u, v) = \|u - v\|_1 = \sum_{i=1}^n |u_i - v_i|$$

$$d_\infty(u, v) = \|u - v\|_\infty = \max_{1 \leq i \leq n} |u_i - v_i|$$

$$d_2(u, v) = \|u - v\|_2 = \sqrt{\sum_{i=1}^n (u_i - v_i)^2}$$

No exemplo a seguir, escolhemos um sistema linear de três equações e três incógnitas para ilustrar como o conceito de norma, logo distância, permite formalizar a ideia de aproximação de uma solução aproximada da solução exata. Outros exemplos podem ser encontrados no livro [Burden e Faires \(2008\)](#).

Vejamos o exemplo a seguir. Considere um sistema de equações lineares e iremos calcular as distâncias entre a solução exata e as soluções obtidas por um método iterativo. Desta forma podemos analisar as componentes dos vetores e verificar se as aproximações, de algum modo, são boas ou ruins.

$$\begin{cases} 2.91819v_1 + 1.19987v_2 + 9.91299v_3 = 14.03105 \\ 1.10217v_1 + 7.25539v_2 + 1.8922181v_3 = 10.24978 \\ 5.60919v_1 + 3.11101v_2 + 9.11399v_3 = 17.83419 \end{cases}$$

É fácil verificar que uma solução exata do exemplo é $u = (u_1, u_2, u_3) = (1, 1, 1)$. Resolvendo pelo método iterativo Eliminação de Gauss com arredondamento de cinco algarismos, obtém a seguinte solução numérica

$$v = (v_1, v_2, v_3) = (1.50746, 0.86674, 0.99816)^t$$

Verificaremos a distância entre os vetores solução com a normas apresentadas.

$$\|u - v\|_1 = |1 - 1.50746| + |1 - 0.86674| + |1 - 0.99816| = 0.64256$$

$$\|u - v\|_\infty = \max\{|1 - 1.50746|, |1 - 0.86674|, |1 - 0.99816|\} = 0.50746$$

$$\|u - v\|_2 = \sqrt{(|1 - 1.50746|^2 + |1 - 0.86674|^2 + |1 - 0.99816|^2)} = 0.52466$$

Note-se que ao analisar as componentes, podemos dizer que v_3 está bastante próximo de u_3 , mas as componentes v_1 e v_2 não estão tão próximas de u_1 e u_2 , respectivamente.

Para entender melhor a importância e os problemas de aproximação ou arredondamento num sistema linear de equações, precisamos falar um pouco de distância entre matrizes.

3.4 Distância entre matrizes

Com relação ao exemplo anterior veremos agora como a distância entre matrizes permite entender e analisar como pequenas perturbações nos coeficientes de um dado sistema linear de equações, influencia na solução do sistema. Esta discussão se faz necessário tendo em vista que vários problemas matemáticos são modelados por sistemas de equações lineares. Sabe-se que $M(m, n)$, o conjunto de todas matrizes de linha m com n colunas, é um espaço vetorial com as operações usuais de soma e multiplicação por escalar. Veremos agora que o espaço das matrizes também é normado, logo, podemos dizer que a *distância entre as matrizes* A e B , induzida pelas norma do espaço, é $\|A - B\|$. Outras distâncias entre matrizes podem ser encontradas em [Deza Michel Marie \(2009\)](#).

3.4.1 Normas nos espaço das matrizes

Seja A uma matriz $m \times n$, com entradas em \mathbb{R} , sabe-se que podemos considerar a transformação linear $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida por

$$u \rightarrow Au$$

A norma matricial, induzida pelos espaços normados \mathbb{R}^n e \mathbb{R}^m permite definir uma norma no espaço $M(m, n)$. Mas precisamente:

$$\|A\| = \max_{u \in \mathbb{R}^n} \frac{\|Au\|}{\|u\|}$$

Norma Coluna: Define-se $\|\cdot\|_1: M(m, n) \rightarrow \mathbb{R}$ como a aplicação que associa a cada matriz $A \in M(m, n)$ o número $\|A\|_1$ é definido como sendo o valor máximo da soma dos valores absolutos das colunas de A , ou seja, Se $A = [a_{ij}]$, então

$$\|A\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^m |a_{ij}|$$

Norma Linha: Define-se $\|\cdot\|_\infty: M(m, n) \rightarrow \mathbb{R}$ como sendo a aplicação que associa a cada matriz $A \in M(m, n)$ o número, denotado por soma dos valores absolutos das linhas de A , ou seja, se $A = [a_{ij}]$, então

$$\|A\|_\infty = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|$$

Norma de Frobenius: a norma de Frobenius é definida como o operador $\|\cdot\|_F: M(m, n) \rightarrow \mathbb{R}$ tal que para cada $A = [a_{ij}] \in M(m, n)$ nos dá

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}$$

Vejam um exemplo. Seja

$$\mathbf{A} = \begin{pmatrix} 1 & 4 & 5 \\ -3 & 2 & 7 \\ 2 & -2 & 1 \end{pmatrix}$$

para esta matriz

$$\sum_{i=1}^3 |a_{i1}| = |1| + |-3| + |2| = 6$$

$$\sum_{i=1}^3 |a_{i2}| = |4| + |2| + |-2| = 8$$

$$\sum_{i=1}^3 |a_{i3}| = |5| + |7| + |1| = 13$$

Então, $\|A\|_1 = \max\{\sum_{i=1}^3 |a_{i1}|, \sum_{i=1}^3 |a_{i2}|, \sum_{i=1}^3 |a_{i3}|\} = \max\{6, 8, 13\} = 13$

de modo similar

$$\sum_{j=1}^3 |a_{1j}| = |1| + |4| + |5| = 10$$

$$\sum_{j=1}^3 |a_{2j}| = |-3| + |2| + |7| = 12$$

$$\sum_{j=1}^3 |a_{3j}| = |2| + |-2| + |1| = 5$$

implica que, $\|A\|_{\infty} = \max\{\sum_{i=1}^3 |a_{i1}|, \sum_{i=1}^3 |a_{i2}|, \sum_{i=1}^3 |a_{i3}|\} = \max\{10, 12, 5\} = 12$

A norma de Frobenius será:

$$\|A\|_F = \left[\sum_{i=1}^3 \sum_{j=1}^3 |a_{ij}|^2 \right]^{\frac{1}{2}} = [1^2 + 4^2 + 5^2 + (-3)^2 + 2^2 + 7^2 + 2^2 + (-2)^2 + 1^2]^{\frac{1}{2}} = 10,63$$

As demonstrações de que são normas no sentido da definição 3 pode ser encontrada no livro [Burden e Faires \(2008\)](#)

3.5 Condicionamento de sistemas lineares

Para resolver um sistema linear de equações $Ax = b$, em que A é uma matriz $n \times n$ não singular e $b \in \mathbb{R}^n$, para n grande, implica a necessidade de usar métodos numéricos, logo arredondamento. Na prática o método numérico resolve um "outro" sistema linear de equações. Como garantir que a solução deste outro sistema esteja próximo da solução do sistema inicial? Veremos agora que o conceito de distância entre matrizes permite dar uma resposta a esta questão.

3.5.1 Erro absoluto e erro relativo

Denotamos por x a solução exata e \bar{x} a solução numérica do sistema $Ax = b$. Considerando estas condições, define-se o erro absoluto E_a e o erro relativo E_r respectivamente como segue

$$E_a = \|x - \bar{x}\| \quad \text{e} \quad E_r = \frac{\|x - \bar{x}\|}{\|x\|}$$

De modo, similar, define-se o resíduo absoluto R_a e o resíduo relativo R_r como sendo

$$R_a = \|b - A\bar{x}\| \quad \text{e} \quad R_r = \frac{\|b - A\bar{x}\|}{\|b\|}$$

respectivamente.

Em algumas situações, será conveniente definir os vetores

$$e = x - \bar{x} \quad \text{e} \quad r = b - A\bar{x}$$

como sendo o vetor erro e vetor resíduo.

Vejam os exemplos. Considere o sistema linear $Ax = b$ como

$$\begin{bmatrix} 1 & 1 \\ 1 & 1,00004 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3,00004 \end{bmatrix}$$

é fácil verificar que $x_1 = 1$ e $x_2 = 2$ é a solução exata. Por outro lado, se considerarmos $\bar{x} = (3, 0)$ como uma solução aproximada, então o vetor residual será

$$r = b - A\bar{x} = \begin{bmatrix} 3 \\ 3,00004 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 1,00004 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0,00004 \end{bmatrix}$$

Dessa forma, o vetor residual tem como norma $\| \cdot \|_{\infty}$ o valor $\|r\|_{\infty} = 0,00004$. Nota que apesar dele ter um valor pequeno, a solução aproximada $\bar{x} = (3, 0)$ é, muito ruim, pois $\|x - \bar{x}\|_{\infty} = 2$.

Vejam uma interpretação geométrica. A solução do sistema pode ser interpretada como a intersecção das retas.

$$\begin{cases} r_1 : x_1 + x_2 = 3 \\ r_2 : x_1 + 1,00004x_2 = 3,00004 \end{cases}$$

Usando o método de substituição, obtemos a solução $x_1 = 1$ e $x_2 = 2$. Já a solução $\bar{x} = (3, 0)$ é solução do sistema

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 + x_2 = 3 \end{cases}$$

obtido do primeiro sistema ao arredondar 1.00004 por 1 e 3.00004 por 3.

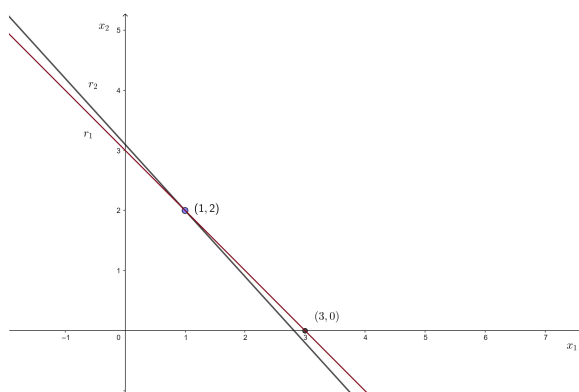


Figura 9 – Gráfico da análise das soluções encontradas com as retas r_1 e r_2

Contudo, ao representar geometricamente o resultado do primeiro sistema é o ponto de intersecção entre as retas de cada equação, e quando efetuar o arredondamento nos coeficientes da equação, as retas ficarão paralelas.

O exemplo dado tem sua importância pois mostra as dificuldades, que podem aparecer. Se não fossem quase coincidentes as retas, intuitivamente espera-se que uma alteração mínima em A ou em b em soluções próximas. Na situação geral, nem sempre se pode recorrer a geometria para descobrir-se os problemas aparecerão. O seguinte teorema usa a distância entre matrizes para tentar apresentar um certo controle sob este tipo de problemas.

Teorema 1. Seja A uma matriz quadrada que possui inversa. Suponha-se que \bar{x} seja uma aproximação da solução do sistema $Ax = b$ e que r seja o vetor residual de \bar{x} . Para qualquer norma, tem-se que

$$\|x - \bar{x}\| \leq \|A^{-1}\| \cdot \|r\|$$

e se $x \neq 0$ e $b \neq 0$

$$\frac{\|x - \bar{x}\|}{\|x\|} \leq \|A^{-1}\| \|A\| \frac{\|r\|}{\|b\|}$$

Demonstração. Dados $r = b - A\bar{x}$ e $Ax = b$, tem-se que $r = Ax - A\bar{x}$ e como a matriz A admite inversa, então, $x - \bar{x} = A^{-1}r$. Pela propriedade da norma de matriz temos:

$$\|x - \bar{x}\| = \|A^{-1}r\| \leq \|A^{-1}\| \cdot \|r\|$$

Provando assim a primeira desigualdade. Por outro lado como, $b = Ax$ então pela propriedade de norma de matriz $\|b\| \leq \|A\| \cdot \|x\|$ de onde juntando esta última desigualdade com a já provada obtemos

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

de onde

$$\frac{\|x - \bar{x}\|}{\|x\|} \leq \|A^{-1}\| \|A\| \frac{\|r\|}{\|b\|}$$

□

O teorema anterior nos diz que $\|A^{-1}\|$ e $\|A\| \cdot \|A^{-1}\|$ podem ser consideradas como indicadores na conexão entre o vetor residual e a precisão na aproximação. Via de regra, o erro relativo $\frac{\|x - \bar{x}\|}{\|x\|}$ é de grande interesse, e a expressão $\|A\| \cdot \|A^{-1}\|$ fornece uma estimativa para este valor.

3.5.2 Condicionamento de sistemas lineares

Para a construção de uma ponte, edifício ou até mesmo um avião, os projetos utilizam vários cálculos matemáticos e muitos deles recaem em sistemas de equações lineares com bastante equações e incógnitas. Para solucionar a questão, esses sistemas utilizam alguns métodos numéricos, que no final fornecem como saber se um dado sistema linear de equações admite soluções numéricas próximas das soluções exatas?

Para esses questionamentos, existe uma teoria matemática que usa o conceito de distância de matrizes para tentar responder essa questão.

Considere um sistema em que x é uma solução exata do sistema linear de $Ax = b$ e que \bar{x} uma solução numérica. Denote por ΔA como sendo a variação $\Delta A = A - \bar{A}$, onde \bar{A} uma matriz, obtida de A após uma perturbação (arredondamento por exemplo). Seja $\Delta x = x - \bar{x}$ o erro na aproximação das soluções de $Ax = b$.

Uma matriz A é chamada de *mal condicionada* se pequenas perturbações nos valores dos seus elementos, podem ocasionar mudanças significativas na solução do sistema $Ax = b$. Se pequenas alterações nos elementos da matriz A resultam em pequenas mudanças na solução $Ax = b$, então será chamada de matriz *bem condicionada*. Vejamos um exemplo.

$$\text{Considere o sistema } \begin{bmatrix} 5 & 4 \\ \lambda & 7 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}$$

Se tomarmos $\lambda = 8,75$ o determinante da matriz A é igual a zero, com isso a matriz A não é invertível, logo precisa adotar $\lambda \neq 8,75$. Para garantir que o sistema tem solução única. Suponhamos as seguintes três situações

$$a) \begin{bmatrix} 5 & 4 \\ 8 & 7 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}$$

$$b) \begin{bmatrix} 5 & 4 \\ 9 & 7 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}$$

$$c) \begin{bmatrix} 5 & 4 \\ 9 & 7 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6,2 \\ 10,1 \end{bmatrix}$$

Não é difícil verificar que

$(-0,67, 2,33)$, $(2, -1)$ e $(-3,5, 3)$ são soluções dos sistemas (a), (b) e (c) respectivamente

O exemplo, mostra que pequenas variações nos coeficientes das matrizes permitem que as soluções fiquem bastante diferentes, isto é, alterações nos dados de entrada acarretam grandes variações na solução do sistema. Logo, conclui-se que o sistema é mal condicionado.

Agora pretendemos dar uma definição precisa de matriz mal e bem condicionada. Isto será feito pelo usando o condicionamento numérico de uma matriz.

3.5.3 Condicionamento numérico de uma Matriz

Com as normas de matriz é possível dar uma forma mais precisa de quando uma a matriz será mal ou bem condicionada.

O número de condicionamento de uma matriz relaciona a forma como os erros se propagam nos dados de entrada e saída. Mas precisamente, dado um sistema linear $Ax = b$ e com A matriz quadrada invertível, então $x = A^{-1}b$, lembrando que $\Delta A = A - \bar{A}$ e $\Delta x = x - \bar{x}$ onde $Ax = b$ e $\bar{A}\bar{x} = b$ então esta última igualdade nos dá

$$(A + \Delta A)(x + \Delta x) = b$$

ou equivalentemente

$$Ax + A\Delta x + \Delta Ax + \Delta A \cdot \Delta x = b$$

Logo,

$$A\Delta x + \Delta Ax + \Delta A \cdot \Delta x = 0$$

ou

$$A\Delta x = -\Delta A(x + \Delta x)$$

Como estamos supondo que o sistema $Ax = b$ tem solução, pode-se dizer que existe a matriz inversa de A . Dessa maneira, obtemos:

$$\Delta x = -A^{-1}\Delta A[x + (\Delta x)] = -A^{-1}\Delta A\bar{x}$$

Usando normas e suas propriedades, temos o seguinte resultado

$$\|\Delta x\| = \|-A^{-1}\Delta A\bar{x}\| = \|A^{-1}\Delta A\bar{x}\|$$

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta A\bar{x}\|$$

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta A\| \|\bar{x}\|$$

Agora dividindo por $\|\bar{x}\|$ em ambos os lados da desigualdade, vale lembrar que $\|\bar{x}\| > 0$, obtemos

$$\frac{\|\Delta x\|}{\|\bar{x}\|} \leq \|A^{-1}\| \|\Delta A\|$$

$$\frac{\|\Delta x\|}{\|\bar{x}\|} \leq \|A^{-1}\| \|A\| \frac{\|\Delta A\|}{\|A\|} \quad (3.1)$$

A expressão $\|A\| \|A^{-1}\|$ é chamada de *número condicional* de A . Nota-se que a razão $\frac{\|\Delta A\|}{\|A\|}$ é uma medida da variação relativa da matriz A , em que assumimos ser pequena. Da mesma forma, a razão $\frac{\|\Delta x\|}{\|\bar{x}\|}$ é uma medida do erro relativo gerado na solução para $Ax = b$ e que neste

caso, o erro é medido em relação ao resultado da solução \bar{x} , e não ao valor inicial de x . Contudo, a desigualdade apresenta uma cota superior em relação a quão grande o erro relativo na solução pode ser em relação ao erro relativo no elemento da matriz. Logo, quanto maior o número condicional, mais a matriz é mal condicionada, pois terá mais "espaço" para que o erro seja grande em relação à solução.

Além das perturbações em A , que refletiram em x , agora considere que a perturbação acontece em b da forma $\bar{b} = b + \Delta b$. Isto provoca um erro em x , novamente $\bar{x} = x + \Delta x$ tal que $A\bar{x} = \bar{b}$ considerando as notações obtemos a igualdade

$$A(x + \Delta x) = b + \Delta b$$

de onde

$$Ax + A\Delta x = b + \Delta b$$

logo

$$A\Delta x = \Delta b$$

Como $Ax = b$ tem uma única solução, A^{-1} existe e

$$A^{-1}A\Delta x = A^{-1}\Delta b$$

logo

$$\Delta x = A^{-1}\Delta b$$

Agora, aplicando a norma e utilizando a propriedade de norma, obtém-se:

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\| \tag{3.2}$$

Faça o mesmo com $Ax = b$

$$\|b\| = \|Ax\| \leq \|A\| \|x\|$$

$$\|b\| \leq \|A\| \|x\| \tag{3.3}$$

Agora, tem que multiplicar membro a membro as desigualdades 3.2 e 3.3

$$\|\Delta x\| \|b\| \leq \|A\| \|A^{-1}\| \|\Delta b\| \|x\|$$

de onde

$$\frac{\|\Delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|} \quad (3.4)$$

Portanto, a perturbação em x está relacionada com a perturbação em b , pela constante $\|A\| \|A^{-1}\|$ e todo esse processo é válido para qualquer norma conveniente.

As desigualdades em 3.1 e 3.4 motivam a seguinte definição.

Número de condicionamento: O número de condicionamento da matriz não singular A em relação dada norma $\|\cdot\|$ é definida como:

$$K(A) = \|A\| \|A^{-1}\|$$

Com essa notação, tem-se que a desigualdade do teorema 1 pode ser reescrita da forma

$$\frac{\|x - \bar{x}\|}{\|x\|} \leq K(A) \frac{\|b - \bar{b}\|}{\|b\|}$$

por outro lado, dada uma matriz qualquer A não singular é imediato que

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = K(A)$$

Com essa nova nomenclatura podemos reescrever os casos de perturbação:

$$\frac{\|\Delta x\|}{\|\bar{x}\|} \leq K(A) \frac{\|\Delta A\|}{\|A\|}$$

e

$$\frac{\|\Delta x\|}{\|x\|} \leq K(A) \frac{\|\Delta b\|}{\|b\|}$$

Com as análises de perturbações e do número condicionado, pode-se dizer que a matriz A é bem condicionada se seu número de condicionamento $K(A)$ é próximo de 1, por outro lado, será *mal condicionada* se $K(A)$ for significativamente maior que 1. Este conceito definido usando a noção de norma (distância) permite entender melhor quando uma solução numérica é boa. Mas precisamente, o condicionamento numérico nos dá a segurança diante uma solução numérica.

3.6 Distância e o método dos mínimos quadrados

Imagina-se que ao ser eleita a presidência do Brasil, precisa viabilizar uma renovação nas leis da previdência referentes a aposentadoria e que uma dessas mudanças tem como efeito o

aumento no tempo de contribuição. Porém, essas alterações devem ter como base a evolução da expectativa de vida dos brasileiros. Então a questão é, como saber a expectativa de vida nos próximos anos? Isto é, seria possível prever essas expectativas de vida para anos futuros? Como será a expectativa dos brasileiros em 2030? Como a matemática e o conceito de distância pode ajudar a responder esta questão? As respostas serão dadas no decorrer desse capítulo, utilizando o método de mínimos quadráticos. Veremos que essa técnica utiliza um conceito de distância entre de um ponto ao um subespaço vetorial.

Como exemplo ilustrativo, iniciaremos analisando a tabela com as expectativas de vida dos brasileiros de 1940 até 2016.

Ano	Parametrização	Expectativa de vida (anos)
1940	1	45,5
1950	2	48
1960	3	52,5
1970	4	57,6
1980	5	62,5
1991	6.1	66,9
2000	7	69,8
2010	8	73,9
2016	8.6	75,8

Fonte: IBGE

Tabela 1 – Expectativa de vida dos brasileiros de 1940 até 2016

A [Figura 10](#) mostra os dados plotados considerando a parametrização dos anos.

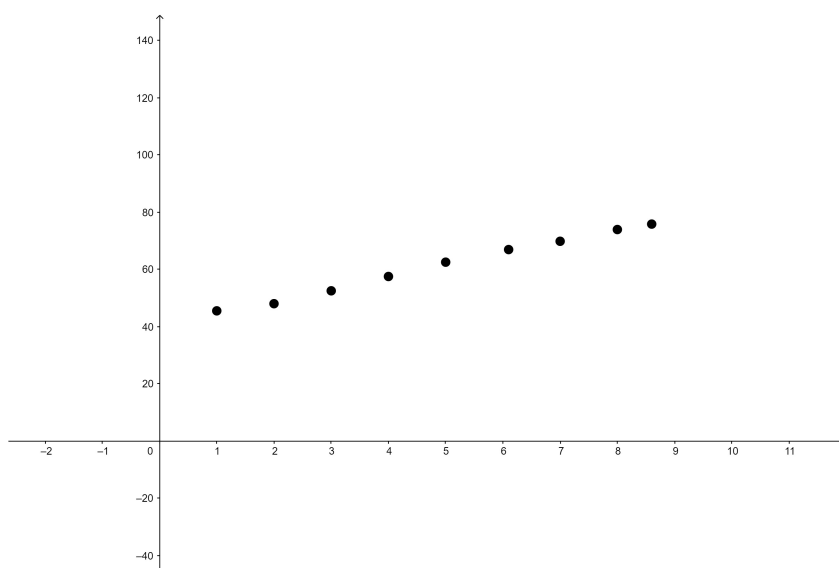


Figura 10 – Pontos plotados da tabela

Vendo a [Figura 10](#), parece razoável supor que os dados da tabela ao serem plotados num plano cartesiano aparentam estar próximos de uma linha reta. Com isso, uma estratégia seria

obter a equação da reta que melhor se ajusta aos dados mostrados, para logo então encontrar o valor correspondente a $x = 10$ (parâmetro que corresponde ao ano de 2030)

A questão agora é, qual seria a reta que melhor se ajusta aos dados?

O método dos mínimos quadrados tem como objetivo encontrar uma curva que "melhor se ajuste" a um conjunto de dados experimentais.

Vejam como o método funciona e como o conceito de distância é importante. Suponhamos que temos um conjunto de dados obtidos numa medição ou levantamento de dados, como no exemplo que vimos no início da seção. No exemplo os pontos na primeira corresponde aos anos e na segunda corresponde a expectativa de vida. Ao plotar este par de pontos num sistema de coordenadas cartesianas, temos uma noção do tipo de curva que aproxima os dados.

Suponha que seja uma reta $y = mx + q$. A questão é, qual é a reta que melhor aproxima dos dados, mas precisamente, qual é o significado que devemos dar quando dizemos que "melhor aproxima". Repare que a escolha por uma reta é motivada pela distribuição dos dados no sistema de coordenadas. Na [Figura 11](#) temos uma outra distribuição que nos induz a pensar numa função exponencial.

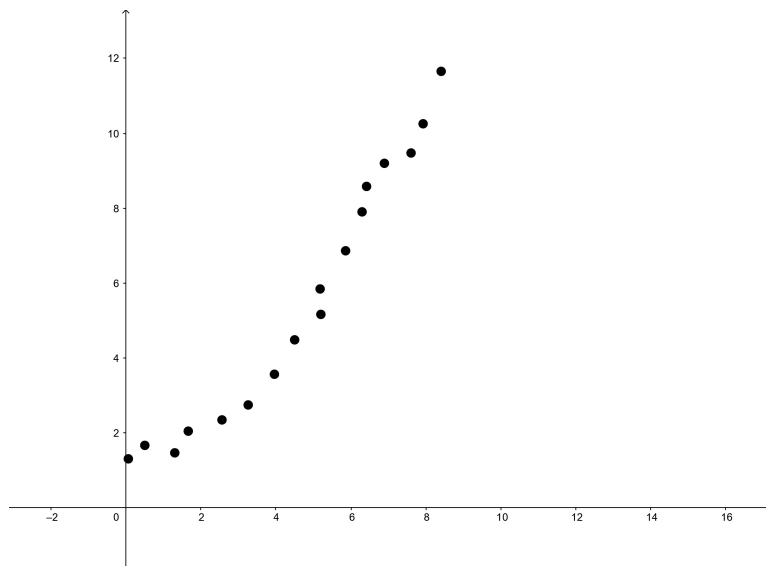


Figura 11 – Pontos plotados com característica de uma exponencial

Independentemente do tipo de curva, precisamos primeiro esclarecer o significado de melhor aproximação.

3.6.1 Teorema da melhor aproximação

Na introdução do métodos mínimos quadrados foi feito um questionamento de qual seria a curva que melhor aproxima os dados. Mas como devemos entender a palavra melhor aproxima? Para responder essa questão precisamos da seguinte condição

Definição 4. Seja W um subespaço de um espaço vetorial normado V . Se v é um vetor em V , então define-se a *melhor aproximação para v em W* como sendo o vetor \bar{v} em W tal que

$$\|v - \bar{v}\| \leq \|v - w\|$$

para todo vetor w em W diferente de \bar{v} .

Em \mathbb{R}^2 ou \mathbb{R}^3 temos por intuição que a menor distância de um ponto a um plano corresponde a distância do ponto ao pé da perpendicular traçada por ela e perpendicular ao plano. Mas, geralmente, num espaço vetorial, a menor distância está associada com a noção de projeção ortogonal: se W é um subespaço de \mathbb{R}^n e v é um vetor em \mathbb{R}^n , então denota-se por $proj_W(v)$ o vetor em W tal que $v - proj_W(v)$ é ortogonal a $proj_W(v)$. Com a projeção ortogonal do vetor v no subespaço vetorial W o vetor $proj_W(v)$ representa a ideia do vetor em W mais próximo de v no sentido de que $v - proj_W(v)$ é vetor de menor norma, comparativamente com $v - w$ para todo $w \in W$

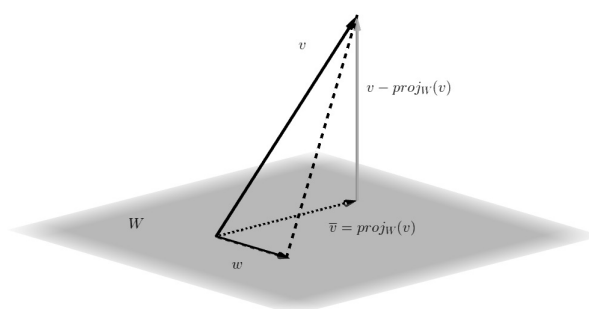


Figura 12 – A menor distância entre v e $proj_W(v)$: $v - proj_W(v)$

Teorema 2. Teorema da melhor aproximação. Dado W um subespaço com dimensão finita de um espaço vetorial V com produto interno e que v é um vetor em V , então $proj_W(v)$ é a melhor aproximação de v em W .

Demonstração. Seja $w \in W$ com $w \neq proj_W(v)$. Claramente $proj_W(v) - w$ pertence a W e $v - proj_W(v)$ é ortogonal a $proj_W(v) - w$. Pelo teorema de Pitágoras, obtém-se

$$\begin{aligned} \|v - proj_W(v)\|^2 + \|proj_W(v) - w\|^2 &= \|(v - proj_W(v)) + (proj_W(v) - w)\|^2 \\ &= \|v - w\|^2 \end{aligned}$$

Como $\|proj_W(v) - w\|^2 > 0$ e por hipótese $w \neq proj_W(v)$, então

$$\|v - proj_W(v)\|^2 < \|v - w\|^2$$

logo

$$\|v - \text{proj}_W(v)\| < \|v - w\|$$

para todo $w \in W$. □

Vejamos um exemplo. Seja $w_1 = (4, 2, -1)$, $w_2 = (1, -2, 1)$ e $v = (2, -1, 3)$. Determinar a melhor aproximação de v no plano W gerado pelos vetores w_1 e w_2 . Em seguida calcular a distância euclidiana de v até W

Para ter a melhor aproximação de v em W deve-se obter a $\text{proj}_W(v)$. Para obter a projeção de um vetor v sobre um subespaço W gerado por vetores w_1 e w_2 é dado por

$$\text{proj}_W(v) = \left(\frac{w_1 \cdot v}{w_1 \cdot w_1} \right) \cdot w_1 + \left(\frac{w_2 \cdot v}{w_2 \cdot w_2} \right) \cdot w_2 = -\frac{1}{7}(4, 2, -1) + \frac{7}{6}(1, -2, 1) = \left(\frac{73}{42}, \frac{-43}{21}, \frac{43}{42} \right)$$

O leitor pode consultar a demonstração da projeção ortogonal de v em W na referência [Poole \(2004\)](#).

Lembrando que a distância de v até W será a distância de v até o ponto mais próxima de W , mas precisamente essa distância será $\|v - \text{proj}_W(v)\|$.

Vejamos no exemplo

$$v - \text{proj}_W(v) = (2, -1, 3) - \left(\frac{73}{42}, \frac{-43}{21}, \frac{43}{42} \right) = \left(\frac{11}{42}, \frac{22}{21}, \frac{83}{42} \right)$$

$$\|v - \text{proj}_W(v)\| = \sqrt{\left(\frac{11}{42} \right)^2 + \left(\frac{22}{21} \right)^2 + \left(\frac{83}{42} \right)^2} = 2.25$$

3.6.2 Método de aproximação por mínimos quadrados

Após ter apresentado o significado de melhor aproximação, será a vez de mostrar como encontrar a curva que "melhor se ajuste" em relação a uma banco de dados obtido. Mas antes será preciso definir o significado "melhor ajuste". Para isso apresentaremos como forma de exemplo e consideremos os dados da [Tabela 1](#). Iniciaremos a parametrização dos pontos representados com relação aos anos, temos que 1940 será parametrizado por 1, 1950 por 2, ..., 1991 por 6.1, 2000 por 7 e 2010 por 8 e com isso temos as coordenadas $(1, 45.5)$, $(2, 48)$ e $(3, 52.5)$. Como foi comentado anteriormente, o plotamento destes pares ordenados [Figura 10](#) sugere a procura de uma reta $y = mx + q$, como a curva adequada. Note-se que se dados os pares ordenados estiverem alinhados, então todos estes pares deverão satisfazer a equação da reta

mas precisamente teremos

$$45.5 = m + q$$

Anos		Ponto	Expectativa de vida
1940	1	1	45.5
1950	2	2	48
1960	3	3	52.5
1970	4	4	57.6
1980	5	5	62.5
1991	6.1	6.1	66.9
2000	7	7	73.9
2010	8	8	73.9
2016	8.6	8.6	75.8

Fonte: IBGE

Figura 13 – Tabela de parametrização dos anos e expectativa de vida

$$48 = 2m + q$$

$$52.5 = 3m + q$$

$$57.6 = 4m + q$$

$$62.5 = 5m + q$$

$$66.9 = 6.1m + q$$

$$69.8 = 7m + q$$

$$73.9 = 8m + q$$

$$75.8 = 8.6m + q$$

escrevendo na forma matricial

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6.1 \\ 1 & 7 \\ 1 & 8 \\ 1 & 8.6 \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} = \begin{bmatrix} 45.5 \\ 48 \\ 52.5 \\ 57.6 \\ 62.5 \\ 66.9 \\ 73.9 \\ 73.9 \\ 75.8 \end{bmatrix}$$

Note-se que como se trata de um sistema com mais equações que variáveis, o sistema de equação linear é impossível quando os pontos não são colineares como no exemplo([Figura 10](#)). Logo a estratégia é encontrar a equação da reta mais próxima possível desses pontos.

Vejamos como fazer isto. Par cada reta candidata será necessário medir a distância vertical de cada ponto do banco de dados até a reta dada. Com estes valores se construí um "vetor erro"e se procura a reta que nos de o vetor erro com menor norma.

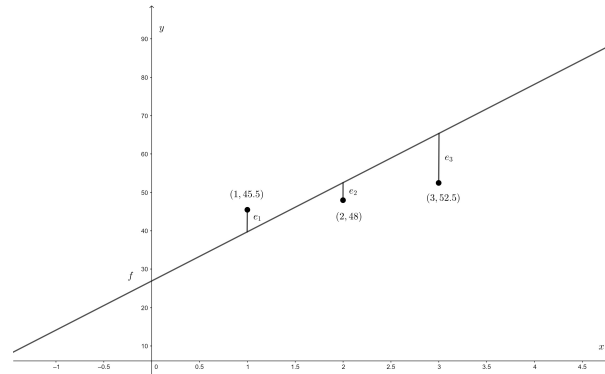


Figura 14 – Reta que minimiza os erros

Mas precisamente se $e = (e_1, e_2, \dots, e_9)$ denota o vetor erro (vide a Figura 14), procuraremos encontrar o vetor e tal que $e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2 + e_7^2 + e_8^2 + e_9^2$ seja o menor possível.

Note-se que estamos utilizando a norma

$$\|e\| = \sqrt{e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2 + e_7^2 + e_8^2 + e_9^2}$$

Note-se também que minimizar $\|e\|$ é equivalente a minimizar

$$\|e\|^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2 + e_7^2 + e_8^2 + e_9^2,$$

expressão que dá origem do nome (mínimos quadráticos).

No exemplo considerado temos

$$e_1 = 45.5 - (q + m)$$

$$e_2 = 48 - (q + 2m)$$

$$e_3 = 52.5 - (q + 3m)$$

$$e_4 = 57.6 - (q + 4m)$$

$$e_5 = 62.5 - (q + 5m)$$

$$e_6 = 66.9 - (q + 6.1m)$$

$$e_7 = 69.8 - (q + 7m)$$

$$e_8 = 73.9 - (q + 8m)$$

e

$$e_9 = 75.8 - (q + 8.6m)$$

Como o modo ilustrativo, consideraremos três retas e para cada uma delas encontraremos o erro quadrático na aproximação dos pontos considerados $(1, 45.5)$, $(2, 48)$, $(3, 52.5)$, $(4, 57.6)$, $(5, 62.5)$, $(6.1, 66.9)$, $(7, 69.8)$, $(8, 73.9)$ e $(8.6, 75.8)$.

$$r_1 : y = 2.5x + 43$$

$$r_2 : y = 3.5x + 42$$

e

$$r_3 : y = 4.1x + 41.1$$

Com a reta $r_1 : y = 2.5x + 43$ temos $q = 43$ e $m = 2.5$, com isso

$$e_1 = 45.5 - (43 + 2.5) = 0$$

$$e_2 = 48 - (43 + 2 \cdot 2.5) = 0$$

$$e_3 = 52.5 - (43 + 3 \cdot 2.5) = 2$$

$$e_4 = 57.6 - (43 + 4 \cdot 2.5) = 4.6$$

$$e_5 = 62.5 - (43 + 5 \cdot 2.5) = 7$$

$$e_6 = 66.9 - (43 + 6.1 \cdot 2.5) = 8.65$$

$$e_7 = 69.8 - (43 + 7 \cdot 2.5) = 9.3$$

$$e_8 = 73.9 - (43 + 8 \cdot 2.5) = 10.9$$

e

$$e_9 = 75.8 - (43 + 8.6 \cdot 2.5) = 11.3$$

Obtemos então os erros $e_1 = 0$, $e_2 = 0$, $e_3 = 2$, $e_4 = 4.6$, $e_5 = 7$, $e_6 = 8.65$, $e_7 = 9.3$, $e_8 = 10.9$ e $e_9 = 11.3$, logo

$$\|e\| = \sqrt{(0)^2 + 0^2 + 2^2 + 4.6^2 + 7^2 + 8.65^2 + 9.3^2 + 10.9^2 + 11.3^2}$$

$$\|e\| = \sqrt{481.972}$$

$$\|e\| = 21.953$$

Similarmente, $r_2 : y = 3.5x + 42$, temos $q = 42$ e $b = 3.5$ de onde

$$e_1 = 45.5 - (42 + 3.5) = 0$$

$$e_2 = 48 - (42 + 2 \cdot 3.5) = -1$$

$$e_3 = 52.5 - (42 + 4 \cdot 3.5) = 0$$

$$e_4 = 57.6 - (42 + 4 \cdot 3.5) = 1.6$$

$$e_5 = 62.5 - (42 + 5 \cdot 3.5) = 3$$

$$e_6 = 66.9 - (42 + 6.1 \cdot 3.5) = 3.55$$

$$e_7 = 69.8 - (42 + 7 \cdot 3.5) = 3.3$$

$$e_8 = 73.9 - (42 + 8 \cdot 3.5) = 3.9$$

e

$$e_9 = 75.8 - (42 + 8.6 \cdot 3.5) = 3.7$$

Assim, obtemos os erros $e_1 = 0$, $e_2 = -1$ e $e_3 = 0$, $e_4 = 1.6$, $e_5 = 3$, $e_6 = 3.55$, $e_7 = 3.3$, $e_8 = 3.9$ e $e_9 = 3.7$, logo

$$\|e\| = \sqrt{(0)^2 + (-1)^2 + 0^2 + 1.6^2 + 3^2 + 3.55^2 + 3.3^2 + 3.9^2 + 3.7^2}$$

$$\|e\| = \sqrt{64.952}$$

$$\|e\| = 8.059$$

Na reta $r_3 : y = 4.1x + 41.1$, temos $q = 41.1$ e $b = 4.1$, logo

$$e_1 = 45.5 - (41.1 + 1 \cdot 4.1) = 0.3$$

$$e_2 = 48 - (41.1 + 2 \cdot 4.1) = -1.3$$

$$e_3 = 52.5 - (41.1 + 3 \cdot 4.1) = -0.9$$

$$e_4 = 57.6 - (41.1 + 4 \cdot 4.1) = 0.1$$

$$e_5 = 62.5 - (41.1 + 5 \cdot 4.1) = 0.9$$

$$e_6 = 66.9 - (41.1 + 6.1 \cdot 4.1) = 0.79$$

$$e_7 = 69.8 - (41.1 + 7 \cdot 4.1) = 0$$

$$e_8 = 73.9 - (41.1 + 8 \cdot 4.1) = 0$$

e

$$e_9 = 75.8 - (41.1 + 8.6 \cdot 4.1) = -0.46$$

logo,

$$\|e\| = \sqrt{(0.3)^2 + (-1.3)^2 + (-0.9)^2 + (0.1)^2 + 0.9^2 + 0.79^2 + 0^2 + 0^2 + (-0.46)^2}$$

$$\|e\| = \sqrt{4.246}$$

$$\|e\| = 2.06$$

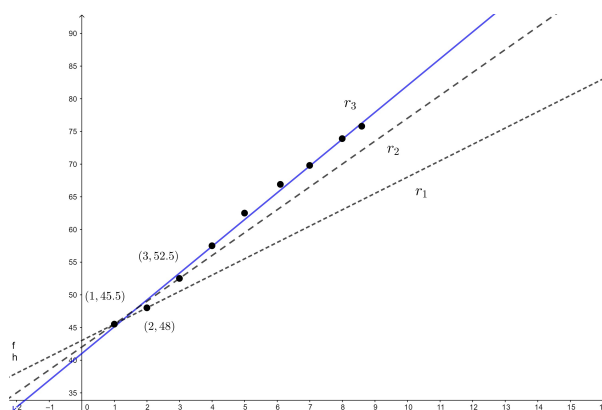


Figura 15 – O menor erro quadrático é a reta r_3

Comparando os erros das retas, conclui-se que a reta que possui o menor erro quadrático entre elas é a reta r_3 . Veja a figura a seguir.

A seguir será descrito o método que me permite deduzir se existe de fato uma reta (curva em geral) que minimize o erro existe e caso sim, como achá-la.

Dados os n pontos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ (banco de dados, medições feitas e etc), ao plotar os dados num sistema de coordenadas cartesiano, os dados estão distribuídos aparentemente ao longo de uma reta. Seja a reta $y = mx + q$ e considere o vetor erro

$$e = (e_1, e_2, \dots, e_n)$$

definida por ela, onde $e_i = y_i - (m + qx_i)$. Lembrando que a $y = mx + q$ que minimiza

$e_1^2 + e_2^2 + \dots + e_n^2$ é chamada de *reta de mínimos quadrados ou reta de melhor ajuste* para os pontos (x_i, y_i) dados. Expressando o problema em forma de matriz teríamos as equações lineares.

$$mx_1 + q = y_1$$

$$mx_2 + q = y_2$$

$$\vdots$$

$$mx_n + q = y_n$$

quando os pontos estiverem todos alinhados. Mas em geral não é difícil imaginar que os pontos (x_i, y_i) não são colineares, assim o sistema torna-se indeterminado. As igualdades anteriores escrito na forma matricial será:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Se por comodidade escrevemos $b = (y_1 y_2 \dots y_n)$, então a igualdade anterior se escreve

como $Ax = b$ onde $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ e $x = \begin{bmatrix} m \\ q \end{bmatrix}$ é a matriz das incógnitas. Note-se também que o

vetor erro é $e = b - Ax$, e queremos encontrar o valor de x (isto é, m e q) para minimizar $\|e\|^2$. A seguinte definição, que usa o conceito de distância entre matrizes será relevante para determinar se existe a incógnita x com a condição desejada.

Definição 5. Seja A uma matriz de ordem $m \times n$ e $b \in \mathbb{R}^m$. Uma solução por mínimos quadrados do sistema linear $Ax = b$ é um vetor $\bar{x} \in \mathbb{R}^n$, tal que

$$\|b - A\bar{x}\| \leq \|b - Ax\|$$

para todo $x \in \mathbb{R}^n$.

Note-se que a desigualdade anterior está comparando os erros para todas as "curvas" que aproximem os dados. Por outro lado, é fácil verificar que qualquer vetor da forma Ax pertence ao espaço coluna da matriz A . Se denotarmos o espaço coluna A por $col(A)$, então encontra uma solução por mínimos quadrados de $Ax = b$ equivale a encontrar um vetor \bar{y} em $col(A)$, tal que

$$\|b - \bar{y}\| \leq \|b - y\|,$$

para todo $y \in col(A)$

Segue da definição de melhor aproximação, que procura-se a projeção ortogonal de b no espaço coluna de A , mas precisamente se \bar{x} é uma solução, então

$$A\bar{x} = proj_{col(A)}(b)$$

Vejamos agora como achar o \bar{x} .

Da desigualdade anterior podemos escrever

$$b - A\bar{x} = b - proj_{col(A)}(b)$$

e concluir que $b - A\bar{x}$ será ortogonal ao espaço $col(A)$. Logo, se a_i denota uma coluna da matriz A , então lembrando o produto interno podemos afirmar que

$$a_i \cdot (b - A\bar{x}) = 0$$

então

$$A^T(b - A\bar{x}) = 0$$

onde A^T denota a transposta de A

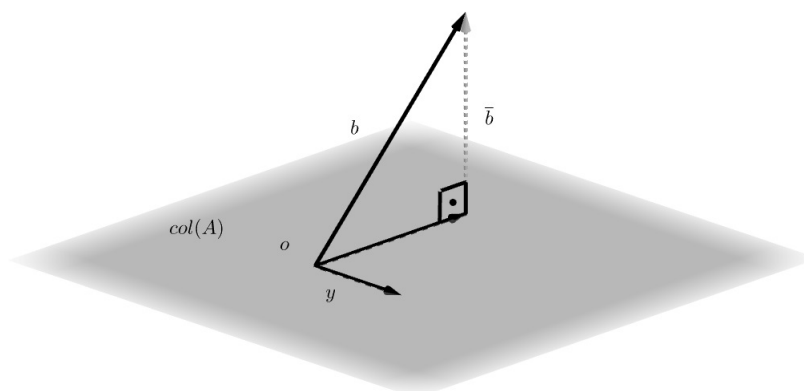


Figura 16 – A projeção ortogonal de b no espaço coluna de A

De onde temos,

$$A^T b - A^T A \bar{x} = 0$$

logo,

$$A^T A \bar{x} = A^T b$$

Note-se que a última igualdade nada mais é que um sistema linear de equações em que o número de equações é igual ao número de incógnitas ($A^T \cdot A$ é uma matriz quadrada) e o sistema possui uma única solução se $A^T A$ tiver inversa. Temos então o teorema a seguir.

Teorema 3. (Teorema dos Mínimos Quadrados)

Seja A uma matriz de ordem $m \times n$ e $b \in \mathbb{R}^m$. Então, $Ax = b$ tem sempre pelo menos uma solução por mínimos quadrados \bar{x} . Além disso,

i. Se \bar{x} é uma solução por mínimos quadrados de $Ax = b$ se, e somente se, \bar{x} é uma solução da equação $A^T A \bar{x} = A^T b$.

ii. A matriz A possui colunas linearmente independentes se, e somente se, $A^T A$ é invertível. Então, neste caso, a solução por mínimos quadrados de $Ax = b$ é única e é dada por

$$\bar{x} = (A^T A)^{-1} \cdot A^T b$$

Demonstração: A demonstração pode ser verificada na referência [Poole \(2004\)](#).

Aplicamos o teorema ao banco de dados do estudo do problema sobre o presidente obter uma estimativa da expectativa de vida para uma pessoa em 2030. Visto na [Tabela 1](#), temos a tabulação dos pontos $(1, 45.5), (2, 48), (3, 52.5), (4, 57.6), (5, 62.5), (6, 66.9), (7, 69.8), (8, 73.9), (8.6, 75.8)$. Com os dados tabulados, devem ser representados na equação $Ax = b$, com isso deve-se encontrar a solução por mínimos quadrados.

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6.1 \\ 1 & 7 \\ 1 & 8 \\ 1 & 8.6 \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} = \begin{bmatrix} 45.5 \\ 48 \\ 52.5 \\ 57.6 \\ 62.5 \\ 66.9 \\ 69.8 \\ 73.9 \\ 75.8 \end{bmatrix}$$

é fácil verificar que as colunas da matriz A são linearmente independentes, logo segue do teorema anterior a solução por mínimos quadráticos existe e é única, mas ainda

$$\bar{x} = (A^T A)^{-1} \cdot A^T b$$

Efetuando os cálculos, temos $\bar{x} = (4.156, 40.745)$, com isso, os parâmetros da reta são $m = 4.156$ e $q = 40.745$, o resultando na reta de mínimos quadrados $y = 4.156x + 40.745$

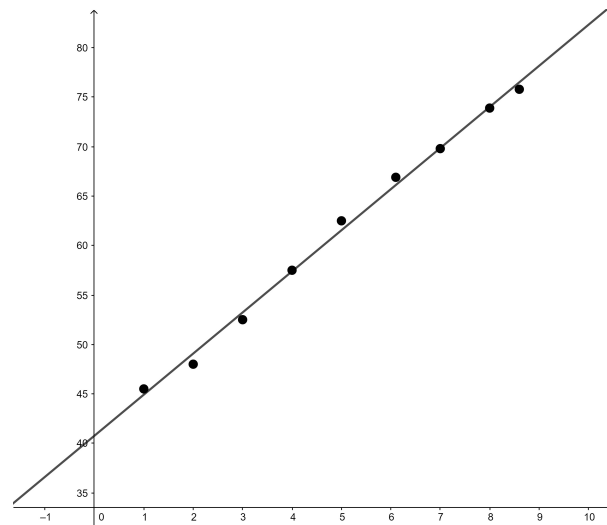


Figura 17 – Expectativa de vida

Portanto, com a equação da reta pode-se estimar a expectativa de vida para os próximos anos. Então, o resultado da expectativa de vida em 2030 é de $y = 4.156 \cdot 10 + 40.745 = 82.305$ anos.

Depois de calcular a expectativa de vida, agora apresentaremos um outro exemplo de mínimos quadrados usando uma função exponencial, para analisar o crescimento populacional brasileiro. Sabe-se que o crescimento populacional tem um comportamento aproximado de uma curva exponencial. Veja na [Figura 18](#). Portanto será utilizada a equação $p(t) = ce^{kt}$, em que representaremos $p(t)$ como o tamanho da população em tempo (anos) t e onde c e k são constantes.

De acordo com os dados do IBGE, como mostra a tabela abaixo, tem-se a representação da população brasileira de 1940 a 2010. Com essas informações pretendemos usar o método dos mínimos quadrados para obter a equação de uma curva exponencial que nos permita estimar a população do Brasil para o ano de 2030.

Ano (t)	População (p)
1940	41.1
1950	51.9
1960	70
1970	93.1
1980	119
1991	146.8
2000	169.8
2010	190.7

Fonte: Censo 2010/IBGE.

Tabela 2 – População brasileira em milhões de pessoas

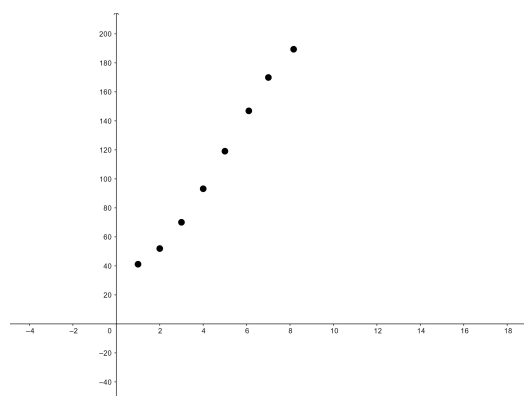


Figura 18 – Pontos plotados em relação aos dados da população brasileira de 1940 até 2010, considerando uma parametrização nos anos.

Efetuada a parametrização, do tempo t , de tal modo que $t = 1$ é 1940, $t = 2$ é 1950, \dots , $t = 6.1$ é 1991, $t = 7$ é 2000 e $t = 8$ é 2010. A Figura 18 "sugere" procurar uma curva do tipo $p(t) = ce^{kt}$ onde c e k serão as constantes que precisamos achar.

Para poder utilizar o método dos mínimos quadrados, será preciso calcular o logaritmo neperiano de ambos os lados para achar a matriz

$$\ln(p(t)) = \ln(ce^{kt})$$

$$\ln(p(t)) = \ln c + \ln e^{kt}$$

$$\ln(p(t)) = \ln c + kt$$

Substitui-se os valores de t e p e introduzindo a variável auxiliar $\ln(c) = w$ temos

$$\ln p = w + kt$$

e

$$3.716 = w + k$$

$$3.949 = w + 2k$$

$$4.25 = w + 3k$$

$$4.533 = w + 4k$$

$$4.779 = w + 5k$$

$$4.989 = w + 6.1k$$

$$5.134 = w + 7k$$

$$5.25 = w + 8k$$

Assim, teremos a correspondência com $Ax = b$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6.1 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad x = \begin{bmatrix} w \\ k \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 3.716 \\ 3.949 \\ 4.25 \\ 4.533 \\ 4.779 \\ 4.989 \\ 5.134 \\ 5.25 \end{bmatrix}$$

Verifica-se parcialmente que a matriz A possui colunas linearmente independentes logo segue do teorema 3 que

$$\bar{x} = (A^T A)^{-1} \cdot A^T b$$

Efetuada os cálculos, temos $\bar{x} = (w, k) (3.55, 0.23)$ com isso a solução por mínimos quadrados tem a forma $w = 3.55$ então $\ln(c) = 3.55$ logo $c = 34.81$. Portanto, a função exponencial que melhor aproxima os dados da [Figura 18](#) é $p(t) = 34.81 \cdot e^{0.23t}$

Veja a figura.

Para conclusão, deve-se determinar a população estimada para 2030.

Como $t = 9$ para a população no ano de 2030. Logo,

$$p(9) = 34.81 \cdot e^{0.23 \cdot 9}$$

$$p(9) = 275.86$$

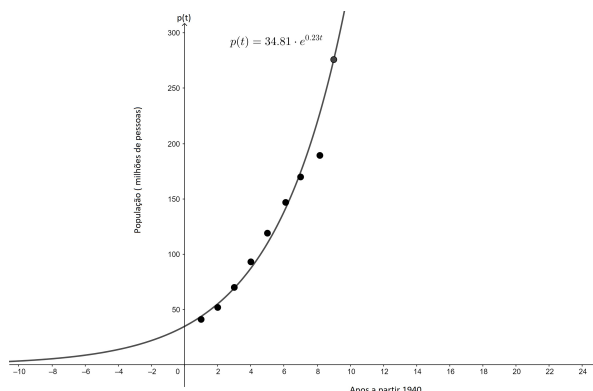


Figura 19 – População brasileira de 1940 até 2010

Portanto, em 2030 a estimativa é que o Brasil tenha aproximadamente 275860 milhões de habitantes.

Nos exemplos apresentados mostramos como o método dos mínimos quadrados pode ser útil quando precisasse criar uma estimativa como foi o caso do prefeito ou governantes que precisava a tomar uma determinada decisão.

3.7 Distância entre funções

Por que é necessário criar uma distância entre funções? Para responder esta questão pense num impulso elétrico ou na figura dada por um eletrocardiograma como é mostrado na [Figura 20](#). Não é difícil imaginar que tais figuras poderiam ser obtidas por meio de funções periódicas. O matemático francês Jean Baptiste Joseph Fourier (1768 – 1830) estudou a decomposição de funções periódicas em séries trigonométricas convergentes, chamadas de “Séries de Fourier” $P_n(x) = a_0 + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)]$. Por meio dessa teoria podemos dizer que funções podem ser aproximadas por somas adequadas de funções senos e cossenos. A aproximação melhora a medida que os termos da soma aumenta, isto é mostrado de modo em outros casos, onde isso pode ser verificado na engenharia eletrônica, estudos de ondas sonoras, sinais musicais, impulsos elétricos ou movimentação de um sistema vibrante como nos eletrocardiogramas [Figura 20](#).



Figura 20 – Eletrocardiograma

Será relevante o conceito de distância entre funções, pois isto nos permitirá analisar o quão próximas as funções estão umas das outras.

Vejamos um exemplo das quatro primeiras iterações de uma função trigonométrica através da série de Fourier.

Por exemplo, considere a função $f(x) = x$ com $x \in [-\pi, \pi]$ e sua série de Fourier

$$P_n(x) = a_0 + \sum_{k=1}^n [a_k \cos(kx) + b_k \sin(kx)]$$

para cada $n \in \mathbb{N}$, com $a_0 = 0$, $a_k = 0$ para todo k e $b_k = (-1)^{k+1} \frac{2}{k}$ para todo k

As seguintes figuras tentam mostrar a aproximação de $P_n(x)$ para $f(x)$ quando o valor de n aumenta.

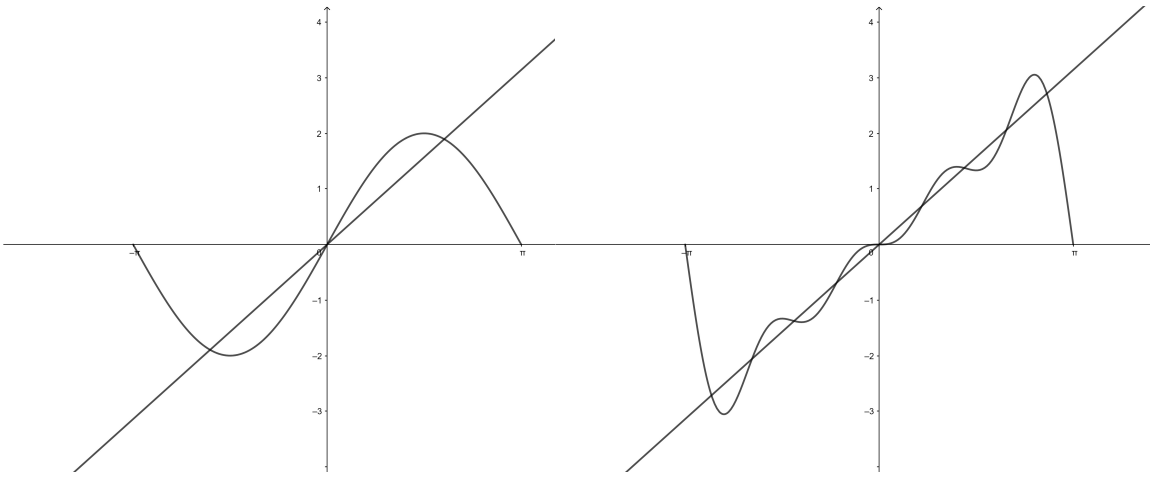


Figura 21 – Primeira aproximação $P_1(x)$

Figura 22 – Quarta aproximação $P_4(x)$

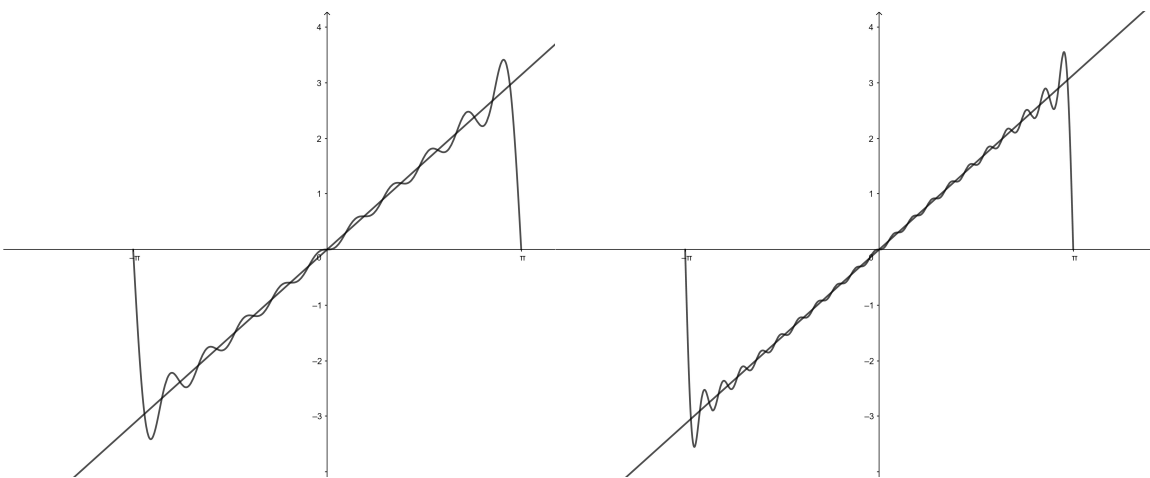


Figura 23 – Décima aproximação $P_{10}(x)$

Figura 24 – Vigésima aproximação $P_{20}(x)$

Como estamos falando de um processo de aproximação, o conceito de distância entre funções é necessário.

Dado um conjunto arbitrário X , considere $\mathbb{B}(X, \mathbb{R})$ o conjunto de todas as funções limitadas de X em \mathbb{R} . Diz-se que $\mathbb{B}(X, \mathbb{R})$ é um espaço métrico com a métrica d definida por:

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|$$

De fato, vejamos que as condições são satisfeitas.

Dado $f \in \mathbb{B}(X, \mathbb{R})$ é imediato a condição M_1 pois

$$d(f, f) = \sup_{x \in X} |f(x) - f(x)| = 0.$$

Vejamos agora a condição M_2 . Sejam $f, g \in \mathbb{B}(X, \mathbb{R})$ com $f \neq g$ este último garante que existirá pelo mesmo um ponto $x_0 \in X$, tal que $f(x_0) \neq g(x_0)$, com isso, $|f(x_0) - g(x_0)| > 0$. Assim pela definição, obtemos

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)| \geq |f(x_0) - g(x_0)| > 0.$$

Portanto M_2 é válido.

Vejamos agora que a condição M_3 é satisfeita. Dadas as funções $f, g \in \mathbb{B}(X, \mathbb{R})$ pela propriedade do valor absoluto

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)| = \sup_{x \in X} |g(x) - f(x)|$$

Logo, $d(f, g) = d(g, f)$. Isto é, M_3 é válido.

Para finalizar, vamos provar a condição M_4 . Dadas as funções f, g e $h \in \mathbb{B}(X, \mathbb{R})$ segue da definição de d e que $|f(x) - g(x)| \geq 0$ e $|g(x) - h(x)| \geq 0$, então

$$\sup_{x \in X} |f(x) - h(x)| + \sup_{x \in X} |h(x) - g(x)| = \sup_{x \in X} \{|f(x) - h(x)| + |h(x) - g(x)|\}.$$

E, pela desigualdade triangular, para todo $x \in X$, obtém-se

$$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)|$$

Assim,

$$\sup_{x \in X} |f(x) - g(x)| \leq \sup_{x \in X} \{|f(x) - h(x)| + |h(x) - g(x)|\}$$

ou seja,

$$\sup_{x \in X} |f(x) - g(x)| \leq \sup_{x \in X} |f(x) - h(x)| + \sup_{x \in X} |h(x) - g(x)|.$$

Portanto, $d(f, g) \leq d(f, h) + d(h, g)$.

Provando assim que $d(f, g)$ definida acima é uma métrica. Esta aplicação é bastante importante no contexto teórico mas não ajuda do ponto de vista prático ou computacional

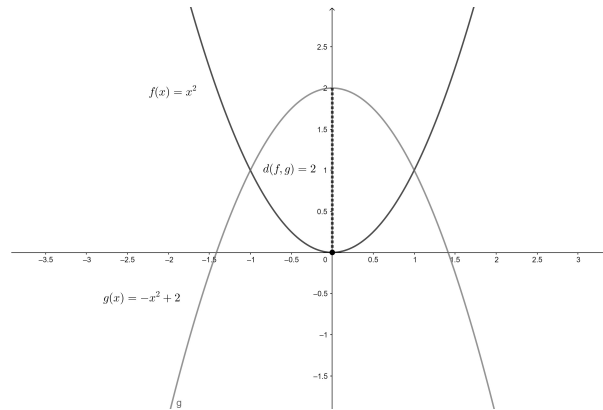


Figura 25 – Distância entre as funções f e g : $d(f, g) = 2$

Vejamos um exemplo considerando as funções quadráticas $f(x) = x^2$ e $g(x) = -x^2 + 2$ no intervalo fechado $X = [-1, 1]$. Considerando f e g como acima, então $d(f, g) = 2$.

Agora apresentaremos um outro exemplo de métrica entre funções que do ponto de vista computacional é mais usado.

Dado $(C[a, b], d)$ o conjunto das funções reais contínuas definidas no intervalo fechado $[a, b]$, ou seja

$$C[a, b] = \{f : [a, b] \rightarrow \mathbb{R} / f \text{ cont\u00ednua}\}$$

defina a aplicação $d : C[a, b] \times C[a, b] \rightarrow \mathbb{R}$, por

$$d(f, g) = \int_a^b |f(x) - g(x)| dx$$

é possível provar que esta aplicação é uma métrica em $C[a, b]$.

A verificação de ser métrica, segue no livro [Lima \(1993\)](#).

3.8 Distância nos códigos

Na História da humanidade as pessoas sempre transmitiram informações por meio de códigos, dentre eles, o mais conhecido é o *Código Morse*, que foi criado por volta do século XIX em que palavras são representadas por pontos e traços.

O código teve um papel muito importante no início das comunicações dos rádios e durante a Segunda Guerra Mundial, onde eram utilizados para a transmissão de mensagens entre navios de guerra e bases navais de diversos países. Outro tipo de código bem familiar é o *Código Binário*, usado na linguagem computacional, que descreve dados em termos de 0s e 1s (que podem ser interpretados como desligado/ligado, fechado/aberto, falso/ verdadeiro ou não/sim), existem também os *códigos lineares*, muito usado pela Nasa para transmitir fotos do espaço. Em 1972, a sonda *Mariner 9* utilizou o código *Reed - Muller*, um tipo de código linear, pelo qual,

enviavam imagens em branco e preto de Marte. Já em 1979 e 1981, com as sondas *Voyager 1 e 2*, utilizaram o *código Golay*, para o envio de fotos, porém dessa vez eram coloridas.

Com o avanço tecnológico, novos códigos foram criados no cotidiano: os de comunicações via satélite, CD, códigos Universais de produtos (UPC - Universal Product Codes) também conhecidos como o código de barras encontrados em embalagens. Além disso, há os códigos de identificação, bem como os números de contas bancárias e de C.P.F. (certificado de pessoa física), em que se fornecem alguns dígitos adicionais para detectar eventuais erros, e o código Padrão Internacional de Numeração de Livros (ISBN - International Standard Book Numbers) com a finalidade de organizar livros publicados e utilizados na organização de bibliotecas. Alguns códigos citados anteriormente, principalmente quando enviados de modo eletrônico, ao serem transmitidos ou enviados a um receptor, podem sofrer algumas alterações provocadas por ruídos, perturbações ou vibrações. Apenas alguns desses códigos têm a capacidade de detectar esses erros e corrigi-los. Existem também códigos que só conseguem detectar os erros e não corrigi-los. Por exemplo, os códigos de barras, CPF e contas bancárias, ao serem digitados erroneamente para a identificação do usuário ou produto, o sistema informa apenas que o código não corresponde ao cadastrado, alertando para realizar novamente o processo para ser redigitado, isso mostra que o código detecta erros, porém não consegue corrigir. Nas últimas décadas, as pessoas começaram a utilizar os aplicativos de mensagens de texto nos smartphones para se comunicarem e, dessa forma, vieram as evoluções. Atualmente, algumas ferramentas em aplicativos celulares possuem corretores de textos ou também conhecidos como corretores ortográficos, que têm a finalidade de corrigir o usuário, quando o mesmo digita alguma palavra de forma errada, com isso o corretor faz algumas sugestões de palavras, como por exemplo: ao digitar a palavra "cachorro" veja na [Figura 26](#). Contudo, o software sugere algumas palavras similares, detectando algum tipo de erro, porém não é eficaz ao sugerir palavras sem o significado desejado ou sem contexto. Portanto, esses softwares detectam erros, sugerem uma correção, porém podem não são tão eficazes. Existem códigos que detectam erros e conseguem corrigir com maior precisão. Veremos neste capítulo as teorias envolvidas e como o conceito de distância ajuda nessa teoria.

Para ilustrar um pouco as ideias vamos considerar o seguinte exemplo: suponha-se que ao enviar uma mensagem, como por exemplo "cachorro", por um acidente foi digitada a palavra "catchorro". Sabe-se que essa palavra não faz parte da língua portuguesa e com isso percebe-se um erro, e nesse caso, é possível fazer uma correção, pois a palavra que mais se assemelha "catchorro" é cachorro. Mas se por exemplo a palavra desejada enviada é "gato", caso tenha uma alteração, pode-se ter mais de uma opção: "rato", "pato", "galo", "gata", "bato" ou "mato", assim torna-se difícil corrigir o erro. O motivo de pouca eficiência deste código é que as palavras são muito próximas (olha aqui a ideia de distância) uma das outras.

O próximo exemplo encontra-se no livro ([HEFEZ; VILLELA, 2008](#)) página 2 ou no livro de ([POOLE, 2004](#)) o qual mostra de forma clara a teoria envolvida.

Numa exploração interplanetária, um robô se move como estivesse num tabuleiro quadri-

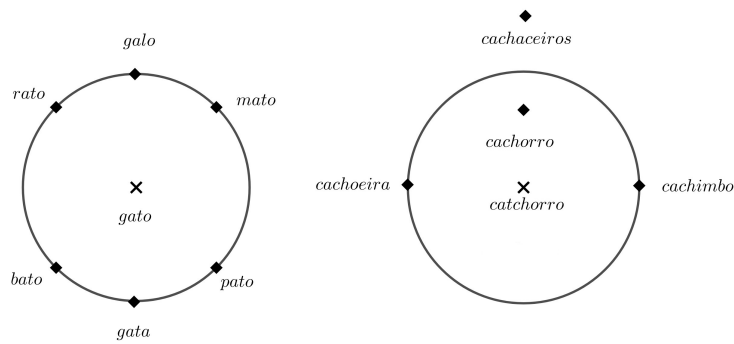


Figura 26 – "Distância" das palavras com relação as palavras gato e cachorro

culado. Ao enviar os comandos de deslocamento Norte, Sul, Leste ou Oeste, esses comandos serão codificados como elementos de $\{0, 1\} \times \{0, 1\}$. Segundo a [Tabela 3](#)

Direção	Código
Leste	00
Oeste	01
Norte	10
Sul	11

Tabela 3 – Códigos Fonte

Esses códigos são chamados de *código fonte*. Ao transmitir a mensagem através de uma transmissão via rádio, ou por outro meio, esses sinais podem sofrer algumas interferências.

Por exemplo o caso do robô, o operador ao enviar a mensagem 00(deslocar para leste), o robô pode receber 01, fazendo com que o mesmo se mova para oeste e não para leste. Para lidar com o problema, recodifica-se as palavras para introduzir redundância para "absorver" os ruídos. O novo código será chamado de *código de canal*, por exemplo

Direção	Código do canal
Leste	00000
Oeste	01011
Norte	10110
Sul	11101

Tabela 4 – Códigos de Canal

Com esta nova codificação, se o operador do robô, por exemplo, enviar a ordem (palavra) 10110 e o receptor do robô recebe 11110 poderá perceber ao comparar com as palavras do código que houve uma alteração (ruído) no canal de transmissão. A questão é, como o robô compara as palavras?

Portanto, o objetivo da teoria é usar um código adequado que saiba lidar com o canal de transmissão e, dependendo do objetivo, detectar e corrigir erros. Naturalmente algumas condições serão necessárias para que a teoria funcione. Basicamente o canal de transmissão deverá satisfazer o seguinte: (P_1): Todos os símbolos transmitidos possuem a mesma probabilidade (pequena) de serem recebidos errados.

(P_2): Se um símbolo é recebido errado, a probabilidade de ser qualquer um dos outros é a mesma.

3.8.1 Métrica de Hamming

Visto acima, será preciso uma noção de distância entre palavras. O objetivo desta seção é apresentar a métrica de Hamming.

Seja A um conjunto finito denominado de *alfabeto* e por comodidade o número de elementos desse conjunto representado pela notação $|A|$.

Define-se um código como sendo qualquer subconjunto próprio do conjunto $A^n = \{(a_1, a_2, \dots, a_n); a_i \in A\}$

Definição 6. Dado $x, y \in A^n$, define-se a distância de Hamming entre x e y como:

$$d_H(x, y) = |\{m; x_m \neq y_m, 1 \leq m \leq n\}|.$$

Isto é, a distância de Hamming entre as palavras x e y será o número de letras distintas comparadas componente a componente.

Verifiquemos que $d_H(x, y)$ satisfaz as condições da [Definição 1](#) para garantir que d_H seja uma métrica.

Ficará a cargo do leitor verificar as três primeiras propriedades, pois são triviais. Será verificada apenas a desigualdade triangular.

M_4 : Dados os vetores palavras x e y , note-se que a contribuição das i -ésimas coordenadas x e y , para $d_H(x, y)$ será zero quando $x_i = y_i$ e um quando $x_i \neq y_i$.

No caso em que a contribuição é zero, tem-se a certeza que a contribuição das i -ésimas coordenadas a $d_H(x, y)$ é menor ou igual a das i -ésimas coordenadas a $d_H(x, z) + d_H(z, y)$.

Com os vetores palavras x, y e z , note-se que no caso em que $x_i \neq y_i$, e, portanto não podemos ter $x_i = z_i$ e $z_i = y_i$ simultaneamente. Como consequência, a contribuição das i -ésimas coordenadas a $d_H(x, z) + d_H(z, y)$ com certeza é maior ou igual a 1, que é a contribuição das i -ésimas coordenadas de $d_H(x, y)$. Portanto, $d_H(x, y) \leq d_H(x, z) + d_H(z, y)$

Provando, assim, que d_H é uma métrica.

Vejamos o exemplo. Seja $A = \{0, 1\}$ e considere o código $C = \{x, y, z\} \subset A^3$ onde

$$x = \{000\} \quad y = \{101\} \quad z = \{111\}$$

Então temos,

$$d_H(x, y) = d_H(000, 101) = 2$$

$$d_H(x, z) = d_H(000, 111) = 3$$

$$d_H(y, z) = d_H(101, 111) = 1$$

O conceito de distância de código, terá uma importância fundamental no próximo teorema. Lembremos que dependendo do contexto o problema o código precisa detectar erros, como por exemplo o leitor de código de barras, o qual percebe o erro, mas não corrige. Já para problemas como da expedição espacial é de grande interesse a detecção e a correção de erros, afinal os custos da operação são caros e demorados para serem enviados toda hora que for detectado o erro. As seguintes definições permitirão apresentar o teorema 4.

Disco e Esfera

Dados um elemento $c \in A^n$ e um número real \mathbb{R}^+ , definimos o *disco e esfera* com centro em c e raio r , como sendo respectivamente os conjuntos seguintes:

$$D(c, r) = \{x \in A^n; d_H(c, x) \leq r\}$$

$$S(c, r) = \{x \in A^n; d_H(c, x) = r\}$$

Seja C um código. A *distância mínima* de C é a menor distância de Hamming entre todos os elementos distintos de C , mas precisamente

$$d(C) = \min\{d_H(x, y); x, y \in C, x \neq y\}$$

Exemplo 1. Visto no exemplo anterior a distância de Hamming entre $x = (0, 0, 0)$, $y = (1, 0, 1)$ e $z = (1, 1, 1)$ são:

$$d_H(x, y) = d_H(000, 101) = 2$$

$$d_H(x, z) = d_H(000, 111) = 3$$

$$d_H(y, z) = d_H(101, 111) = 1$$

Portanto, a distância mínima do código $C = \{x, y, z\}$ é $d(C) = 1$

Lema 1. Seja C um código com distância mínima d . Se a e b são palavras de C , então $D(a, k) \cap D(b, k) = \emptyset$ onde $k = \lfloor \frac{d-1}{2} \rfloor$ e $\lfloor m \rfloor$ representa a parte inteira de um número real m

A prova desse lema segue na referência no livro [Hefez e Villela \(2008\)](#)

Veremos agora a importância da distância mínima d de um dado código.

Teorema 4. Seja C um código com a distância mínima d . Então, C pode *corrigir* até $k = \lfloor \frac{d-1}{2} \rfloor$ erros e *detectar* até $d-1$ erros.

Demonstração. Se ao enviarmos uma palavra a do código cometemos m erros com $m \leq k$, recebendo a palavra b , então $d(a, b) = m \leq k$, enquanto pelo [Lema 1](#), a distância de b a qualquer outra palavra do código é maior do que k . Com isso pode-se determinar a univocamente a partir de b . Por outro lado, dada uma palavra do código, pode nela introduzir até $d - 1$ erros sem encontrar outra palavra do código. Então a detecção do erro será possível. \square

Vejamos um exemplo

Exemplo 2. Consideremos a situação do robô do início do capítulo. Vejamos como o teorema acima analisado até quantos erros o código corrige e detecta. Lembremos que o código do robô é Na ocasião tem as palavras:

Direção	Código
Leste	00000
Oeste	01011
Norte	10110
Sul	11101

Tabela 5 – Códigos do robô

A distância entre as palavras do código é

$$\begin{aligned} d_H(L, O) &= 3 & d_H(O, N) &= 4 \\ d_H(L, N) &= 3 & d_H(O, S) &= 3 \\ d_H(L, S) &= 4 & d_H(N, S) &= 3 \end{aligned}$$

Logo, a distância mínima do código do robô é 3. Tendo o conhecimento da distância mínima, pode-se calcular quantos erros o código do robô consegue corrigir e detectar como $k = \lfloor \frac{3-1}{2} \rfloor = 1$ e $d - 1 = 3 - 1 = 2$, então o código do robô pode corrigido até um erro e detectar até dois erros no envio de palavras.

Veja um exemplo em que o emissor envia a palavra 10110(norte) e o sistema do robô recebe a palavra com apenas uma alteração 1110. O receptor do robô detecta que a palavra não faz parte do código(detecta o erro) e assumindo que os ruídos só trocam uma componente, o robô poderá corrigir a palavra pelo vizinho mais próximo 10110 no caso.

3.8.2 Representação da distância mínima e geometricamente

Em alguns casos é possível representar a distância mínima de um código geometricamente, isso irá facilitar ao leitor ampliar a compreensão do conteúdo.

Veja um exemplo. Considere o código $C = \{c_0, c_1\}$ onde $c_0 = (0, 0, 0)$ e $c_1 = (1, 1, 1)$ posicionando estas palavras nos vértices de um cubo unitário na forma de vetores, como mostra a [Figura 27](#).

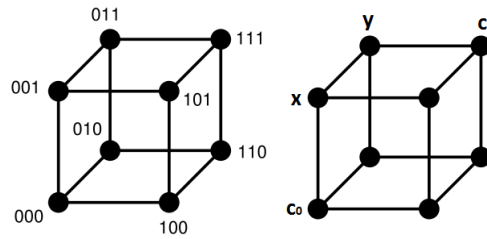


Figura 27 – Representação da distância mínima e corretores de erros através do cubo unitário

Note-se que $C \subset \{0, 1\}^3$. Pela teoria descrita acima pode-se verificar o quanto o código C pode corrigir e até um erro e detectar até dois erros. Isto pois efetuando os cálculos temos $d(c_0, c_1) = d(000, 111) = 3$, logo a menor distância será 3. Sabendo a $d(C)$, providenciaremos os estudos dos erros e correções que o código C consegue efetuar. Vejamos $k = \lfloor \frac{3-1}{2} \rfloor k = 1$ e $d - 1 = 3 - 1 = 2$.

Por exemplo, suponha que o emissor envia a palavra $c_1 = (1, 1, 1)$, porém o receptor recebe a palavra $y = (0, 1, 1)$, pode-se observar que houve apenas uma troca e que em uma das componentes da palavra, essa troca pode ter sido provocada por algum ruído durante a transmissão. Pelo teorema o código consegue detectar até duas trocas, neste caso perceberá o erro. Com isso ao perceber que a palavra que recebeu não faz parte do código ($(0, 1) \notin C$), então o próprio código conseguirá corrigir, afinal foi visto que o código consegue corrigir até uma troca nas componentes das palavras. Como a palavra mais próxima de y é $c_1 = (1, 1, 1)$, então o receptor irá compreender que mesmo recebendo a palavra y , entende que o emissor quis comunicá-lo com a palavra c_1 .

Contudo, nesta situação pode-se representar geometricamente através de um cubo conforme a Figura 27, em que as distâncias são representadas pelas arestas e os vértices os códigos (palavras).

Note-se que ao criar um código deve ser feita uma análise com as possíveis alterações que o código pode sofrer durante o envio e essas alterações são provocadas por ruídos ou interferência. Por isso, temos algumas palavras nos vértices do cubo que não fazem parte do código, porém são palavras que podemos aparecer ao receptor devido esses fenômenos que podem ocorrer ao enviar a mensagem ao receptor.

Vejamos um exemplo. Considere o código $C = c_1, c_2, c_3, c_4$

$$\begin{aligned} c_1 &= 10010 & c_3 &= 00101 \\ c_2 &= 11001 & c_4 &= 01110 \end{aligned}$$

Suponha que o receptor recebe as palavras $u = [11000]$, $v = [01000]$ e $w = [00001]$.

Vamos decodificar as palavras u, v e w pelo vizinho mais próximo. Para isso veremos a quantidade de todas as combinações dos vetores tomados 2 a 2.

$$\begin{aligned}d_H(c_1, c_4) &= 3 & d_H(c_2, c_3) &= 3 \\d_H(c_1, c_3) &= 4 & d_H(c_2, c_4) &= 4 \\d_H(c_1, c_2) &= 3 & d_H(c_3, c_4) &= 3\end{aligned}$$

Portanto, a distância mínima $d(C) = 3$. Pelo teorema 4 podemos analisar o quanto este código detecta erros e o quanto pode corrigir. Para esse código temos $d(C) - 1 = 2$, com isso o código C detecta até dois erros e corrige $k = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$ erro pelo vizinho mais próximo. Agora verificaremos se é possível decodificar os vetores u, v e w , calculando todas as suas distâncias.

$$\begin{aligned}d_H(u, c_1) &= 2 & d_H(v, c_1) &= 3 & d_H(w, c_1) &= 3 \\d_H(u, c_2) &= 1 & d_H(v, c_2) &= 2 & d_H(w, c_2) &= 2 \\d_H(u, c_3) &= 4 & d_H(v, c_3) &= 3 & d_H(w, c_3) &= 1 \\d_H(u, c_4) &= 3 & d_H(v, c_4) &= 2 & d_H(w, c_4) &= 4\end{aligned}$$

Portanto pode-se analisar que mesmo que o receptor entenda a palavra u consegue compreender que o emissor enviou a palavra c_2 , o mesmo acontece com a palavra w que decodifica a palavra c_3 , pelo fato de que o código consegue corrigir uma troca na componente da palavra. Já não consegue corrigir a palavra v tendo em vista que $d_H(v, c_2) = d_H(v, c_4) = 2$

3.9 Espaço de Hausdorff

Para a maioria dos seres humanos fazer o reconhecimento facial de uma pessoa é de fato simples, basta comparar com uma foto ou mesmo recorrer a nossa memória. Mas realizar um reconhecimento facial para uma máquina não é trivial, então as máquinas fazem o reconhecimento através dos seguintes processos: captam a imagem, depois passa por um algoritmo para identificar centenas de pontos únicos, através desse processo pode calcular a distância entre o nariz e os olhos, boca e queixo, orelhas e olhos e outras partes do rosto. Portanto, a distância entre conjuntos será relevante pois são feitas medidas entre as partes do rosto como conjunto de pontos, mas precisamente analisamos o quão perto um conjunto está do outro utilizando o conceito de distância, para isso a definição de distância precisa ser adequada. Nesses cálculos devemos tomar cuidado, pois podem existir alguns conjuntos que se interceptam, e quando isso ocorrer a distância entre eles será nula, de fato não podemos definir como uma métrica. Perante esse fato, será visto mais adiante do capítulo a importância de os conjuntos estudados sejam fechados e limitados.

Diante dessas particularidades, o matemático Felix Hausdorff apresentou um estudo da métrica entre subconjuntos de um dado no espaço métrico. A métrica foi introduzida pela primeira vez por Hausdorff em seu livro *Grundzüge der Mengenlehre*, Fundamentos da teoria dos conjuntos, publicado em 1914. Em poucas palavras, Hausdorff propõe calcular a menor distância entre elementos de A com todos os elementos de B , reciprocamente, a menor distância entre

elementos de B com A e assim considerar o maior valor entre essas distâncias. Então, Hausdorff definiu que a maior destas distâncias seria a distâncias entre A e B , com isso será possível medir as distâncias entre vários subconjuntos.

A métrica de Hausdorff existem inúmeras aplicações, por exemplo, na área computacional, a métrica de Hausdorff é bastante utilizada, pois as imagens são apresentadas como nuvens de pontos em que representam os conjuntos, fazendo possível a medição da distância entre essas imagens. Além da computação, é possível encontrarmos estudos com a métrica de Hausdorff nas representações cartográficas, usando as imagens de satélites como conjuntos de pontos.

Vale lembrar que se deve medir a distância entre conjuntos quaisquer, e não entre conjuntos representados por elementos numéricos.

Agora analisaremos uma situação problema em que a definição da distância mínima entre os conjuntos não será a melhor opção, pois muitos utilizam do argumento que a distância entre conjunto seja a menor possível, mas nem sempre é válido esse tipo de argumento. Veja a seguir o motivo pela qual essa técnica não é a mais adequada.

Lembramos da definição usual de distância definida num espaço métrico

$$d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}$$

Se existir uma intersecção entre os conjuntos e os mesmos forem distintos ($A \neq B$) então $d(A, B) = 0$, o qual não condiz com a definição de métrica em um espaço métrico.

Note-se que dentro do contexto considerado as condições dos conjuntos serão limitados e fechados, também é conhecido como *conjuntos compactos*.

O conjunto deve ser limitado, pois não é possível medir a distância de um ponto de um conjunto com todos os outros pontos de um conjunto com infinitos pontos. Agora veremos no exemplo a importância de os conjuntos serem fechados.

Vejamos um exemplo. Consideremos os conjuntos $A = [1, 2] \times [1, 2]$ e $B =]1, 2[\times]1, 2[$. Temos que $d(A, B) = d(B, A) = 0$, lembramos que os conjuntos são distintos, com isso não é possível definir uma métrica, pois pela propriedade do espaço métrico quando a distância for zero os elementos tem que ser iguais, o que contradiz com a situação do exemplo.

3.9.1 Distância de Hausdorff

De uma maneira informal, se diz que a distância de Hausdorff mede o quão perto os dois conjuntos estão entre eles, num espaço métrico. Na situação problema a seguir, podemos analisar o raciocínio feito pelo Hausdorff.

Dados dois conjuntos compactos A e B num espaço métrico X . Seja os elementos x e y dos conjuntos A e B respectivamente, podemos dizer que a distância de x até o conjunto B é

dado pelo $\inf_{x \in A} d(x, B)$ e de y até A $\inf_{y \in B} d(y, A)$, em que representa a menor distância de um elemento até o conjunto. Portanto temos

$$H(A, B) = \sup\{d(x, B); x \in A\}$$

Como consequência tem-se a definição

$$H(B, A) = \sup\{d(y, A); y \in B\}$$

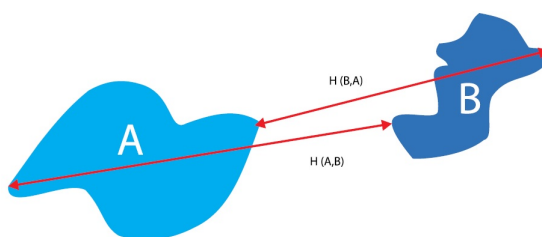


Figura 28 – A distância entre o conjunto A e B é distinto de B e A : $H(A, B) \neq H(B, A)$

Fonte – Santos (2016)

Note-se que $H(A, B) \neq H(B, A)$, vide o [Seção 3.9](#), Hausdorff irá definir a métrica entre conjuntos, como a maior distância entre x ao B e de y ao A , assim a maior dessas distâncias será definida como a distância de Hausdorff. Vale ressaltar que a distância entre os conjuntos tem como métrica presente em A e B . De maneira formal definimos a métrica de Hausdorff.

Definição 7. Dado um espaço métrico (X, d) para cada subconjunto compacto A e B em X ; define-se a distância de Hausdorff

$$d_H(A, B) = \max\{H(A, B); H(B, A)\}$$

A seguir segue um exemplo inspirado Santos (2016)

Dados os conjuntos em \mathbb{R}^2 , $A = [0, 5] \times [0, 9]$ e $B = [12, 14] \times [0, 4]$

Primeiramente iremos calcular a distância usual.

$$d(A, B) = \inf\{d(x, y) / x \in A, y \in B\}$$

$d(A, B) = 7$ veja na figura [Figura 29](#).

Por outro lado, a distância de Hausdorff.

$$d_H(A, B) = \max\{\sup d(x, B), \sup d(y, A)\}$$

Como $\sup_{x \in A} d(x, B) = 13$ e o $\sup_{y \in B} d(y, A) = 7$, então a $d_H(A, B) = 13$

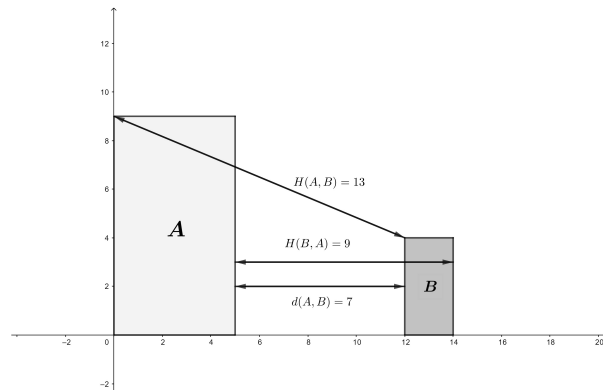


Figura 29 – Distância usual vale 7 e distância de Hausdorff vale 13

Vejamos outro exemplo. Sejam os conjuntos

$$A = \{(2, 1), (3, 1), (3, 2), (2, 2), (1, 3), (4, 3), (2, 4), (3, 4), (2, 3), (3, 5), (2.5, 5)\}$$

e

$$B = \{(6, 5), (6, 6), (7, 5), (7, 6)\}.$$

Veja no plano cartesiano esses conjuntos

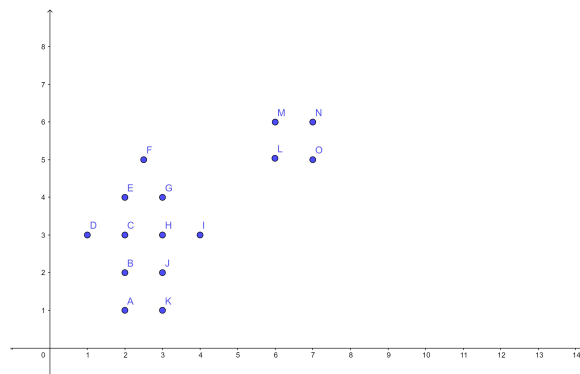


Figura 30 – Distância de Hausdorff é de 5.66

$$H(A, B) = \sup_{x \in A} d(x, B) = 5.66$$

Por outro lado

$$H(B, A) = \sup_{y \in B} d(y, A) = 5$$

a $H(B, A)$, portanto, a $d_H(A, B) = \max \{5.66, 5\} = 5,66$

Agora verificaremos as condições da métrica de Hausdorff são satisfeitas.

Teorema 5. Seja (\mathbb{M}, d) um espaço métrico. Dados dois conjuntos compactos $A, B \in \mathbb{M}$ define-se a distância de Hausdorff entre A e B como:

$$d_H(A, B) = \max\{\sup_{x \in A} d(x, B); \sup_{y \in B} d(y, A)\}$$

ou escrito de outro modo

$$d_H(A, B) = \max\{H(A, B), H(B, A)\}$$

Verificaremos as condições de ser métrica.

M_1 : Para todo $A \in \mathbb{M}$, temos pela definição $H(A, A) = \sup_{x \in A} d(x, A)$, como d é uma métrica em \mathbb{M} , pode-se concluir que a distância entre o mesmo ponto é zero. Logo $H(A, A) = 0$, portanto $d_H(A, A) = 0$

M_2 : Dados $A, B \in \mathbb{M}$ com $A \neq B$ e $H(A, B) = \sup_{x \in A} d(x, B)$ e $H(B, A) = \sup_{y \in B} d(y, A)$, deve-se mostrar $d_H(A, B) > 0$.

Como os conjuntos A e B são compactos, então existe $x_o \in A$ e $y_o \in B$, tais que $H(A, B) = d(x_o, y_o)$, como são fechados, limitados e os conjuntos são distintos, então $d(x_o, y_o) > 0$. Isto, pois Se $A \neq B$, então existe $x \in A$ e $x \notin B$, assim $H(A, B) \neq 0$, pois A e B são fechados, portanto $d_H(A, B) \neq 0$.

M_3 : Dados $A, B \in \mathbb{M}$, deve-se mostrar a simetria $d_H(A, B) = d_H(B, A)$. Essa demonstração é imediata.

M_4 : Dados $A, B, C \in \mathbb{M}$, precisa-se mostrar $d_H(A, B) \leq d_H(A, C) + d_H(C, B)$.

Iniciaremos a verificação da desigualdade.

$$H(A, B) \leq H(A, C) + H(C, B)$$

Sabendo que $a \in A$ e $b \in B$, temos que:

$$\text{Para qualquer } c \in C \quad d(a, b) \leq d(a, c) + d(c, b)$$

Então, para todo $c \in C$,

$$\min_{b \in B} d(a, b) \leq \min_{b \in B} \{d(a, c) + d(c, b)\} = d(a, c) + \min_{b \in B} (d(c, b))$$

, e para todo $c \in C$, temos

$$d(a, B) \leq d(a, c) + d(c, B)$$

Como $d(c, B) \leq H(C, B)$, em que não depende de c , com isso pode-se concluir que:

$$d(a, B) \leq \min_{c \in C} d(a, c) + H(C, B)$$

No que implica em

$$d(a, B) \leq d(a, C) + H(C, B) \leq H(A, C) + H(C, B).$$

Concluí-se que:

$$H(A, B) = \max_{a \in A} d(a, B) \leq H(A, C) + H(C, B)$$

Analogamente, prova-se que

$$H(B, A) \leq H(B, C) + H(C, A)$$

Portanto:

$$\begin{aligned} d_H(A, B) &= \max\{H(A, B), H(B, A)\} \leq \\ &\max\{H(A, C) + H(C, B), H(B, C) + H(C, A)\} \end{aligned}$$

Usando a propriedade de máximo, obtém:

$$\begin{aligned} &\leq \max\{H(A, C), H(C, A)\} + \max\{H(B, C), H(C, B)\} \\ &= d_H(A, C) + d_H(C, B). \end{aligned}$$

Contudo, prova-se que $d_H(A, B) \leq d_H(A, C) + d_H(C, B)$.

No livro [Lima \(1993\)](#) temos a distância de um ponto a um conjunto e distância entre conjuntos.

GRAFOS

4.1 Grafos

Neste capítulo apresentaremos uma breve passagem contanto sobre o surgimento da teoria de grafos e a relevância do conceito de distância que pode ser aplicado em grafos. A teoria se originou com a modelagem de um problema e a discussão sobre a resolução dessa problemática, conhecido como o famoso "Problema das pontes de Königsberg". Contam que no século *XVIII*, na cidade Königsberg (atualmente Kaliningrado) uma cidade dentro de um enclave russo entre a Polônia e a Lituânia, a cidade possuía duas ilhas separadas pelo Rio Pregel, em que existiam seis pontes que as ligavam às margens da cidade e uma outra fazia a ligação de duas ilhas entre si (veja na [Figura 31](#)). Na época, a população se questionava se era possível atravessar as sete pontes, sem passar duas vezes pela mesma, retornando ao ponto de partida. Esse questionamento persistiu por muitos anos por não ter ninguém que chegasse a uma resolução.

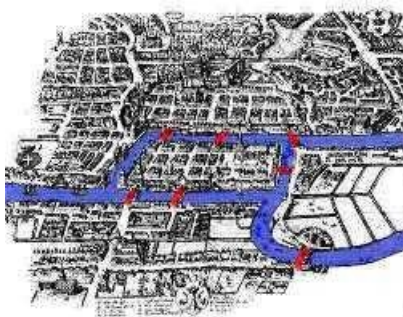


Figura 31 – As Setes pontes de Konigsberg

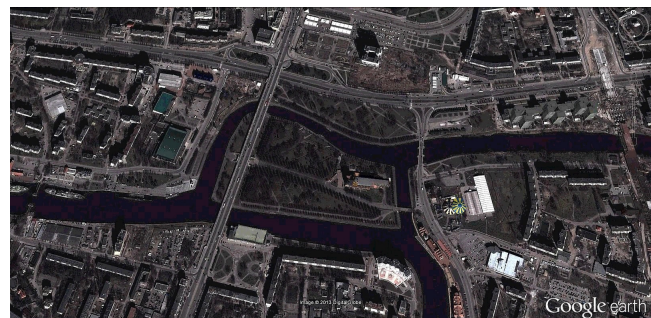


Figura 32 – Vista da atual Kaliningrado

Em meados de 1735, o matemático suíço Leonhard Euler demonstrou que o problema tinha resposta negativa. Para isso utilizou o raciocínio de transformar os caminhos em retas e as intersecções dessas retas em pontos, além disso, desconsiderou algumas coisas irrelevantes ao problema, como por exemplo a largura e o formato das pontes, com isso criou possivelmente o primeiro grafo. Portanto, Euler conseguiu elaborar uma teoria em que pode-se aplicar em vários

problemas semelhantes ao apresentado em Königsberg, com isso a modelagem desses problemas passou a ser conhecida como *grafos Eulerianos*.

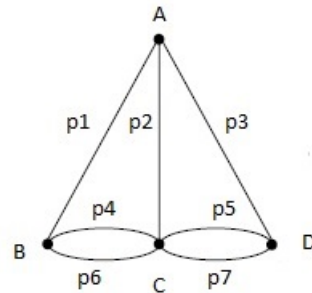


Figura 33 – A, B, C e D representam as intersecções entre as pontes e os P_1, P_2, \dots, P_7 as pontes.

Contudo, Euler conseguiu elaborar uma estratégia para estudar o problema da travessia das pontes, da seguinte maneira: ao caminhar de qualquer ponto A, B, C ou D são utilizadas duas pontes, uma para chegar de um ponto e outra para sair. Então, concluiu que cada ponto deve possuir uma quantidade par de linhas (lembrando que a condição é de cruzar uma única vez cada ponte). Porém, no grafo das pontes de Königsberg há uma quantidade de linhas ímpares, isso permitiu concluir que a resposta do problema é negativa.

Atualmente não existem mais as sete pontes originais, pois uma foi demolida e reconstruída em 1935, outras duas foram destruídas durante a Segunda Guerra Mundial e duas foram demolidas para dar lugar a uma única via expressa. Veja na [Figura 32](#)

Curiosamente, no Brasil na região central de Recife, em Pernambuco, tem-se uma situação semelhante quanto a questão das travessias das pontes em Königsbergem. Veja o mapa e o grafo da situação em Recife.



Figura 34 – Mapa das pontes de Recife

Os grafos são modelos matemáticos que auxiliam problemas práticos em diversas áreas, como na indústria, informática ou na engenharia, bem como: grafos das possibilidades de modelar relações de amizades, na coloração de mapas, sistemas rodoviários, ferroviários e aéreos, além dos tráfegos urbanos, jogos de campeonatos, problemas de roteamento, fluxos de rede e sistema de transmissão e outras aplicações. A seguir mostraremos uma situação real, em que será apresentado um modelo de grafo e a importância de definir uma distância para essa situação real.

Custo mínimo de gerenciamento de uma transportadora

Em qualquer empresa a palavra "economia" soa bem para qualquer empresário, afinal, tentar otimizar e reduzir o custo de uma empresa certamente pode aumentar seu lucro. Veja o exemplo de uma transportadora, que deve distribuir suas mercadorias pela cidade. Mas de que forma um grafo pode ajudar essa empresa a ter menos gastos? Qual a importância de definir um conceito de distância num grafo e sua relação com o menor caminho?

Pode-se modelar em um grafo com todos os caminhos possíveis que o caminhão da transportadora possa percorrer, indo da transportadora até o local determinado para a entrega. Contudo, deve-se obter o "menor caminho" para que possam economizar com o combustível e desgaste do caminhão por exemplo.

Nesta seção apresentaremos brevemente o algoritmo de Dijkstra, que poderá auxiliar e mostrar o menor caminho entre a transportadora e o destino das encomendas, assim com essas informações em mãos o motorista terá o menor caminho entre dois locais, no qual, um é a transportadora e ou outro o destino da entrega. Logo, o conceito de distância é relevante para determinarmos esse menor percurso e como consequência a diminuição do custo da entrega.

A seguir será apresentado a definição de grafo e em seguida a importância de sua utilidade.

Define-se um grafo como um par $G = (V, A)$, onde $V = \{v_1, v_2, \dots, v_n\}$ é um conjunto finito, cujos elementos são chamados de *vértices*, e $A \subseteq \{(v_i, v_j); v_i, v_j \in V\}$ é um conjunto de pares não ordenados de vértices, cujos elementos são chamados de *arestas* de G .

Dependendo da natureza do problema, os vértices podem ser representados por pessoas, casas, cruzamento de rodovias, linhas de transmissões e outros, e as arestas também podem ser representadas como uma relação de amizade nas redes sociais, caminhos entre duas casas, uma ligação entre duas ruas e etc. Para uma melhor compreensão e por comodidade representaremos uma aresta (u, v) por uv . Com isso podemos ilustrar os vértices de um grafo como ponto e as arestas como segmento de reta unindo esses pontos. Em caso de existir uma aresta ligando dois vértices quaisquer, diremos que os vértices são adjacentes.

O grafo G é dito *ponderados* se as arestas ou vértices têm uma representação tanto numéricas como alfabéticas, no caso das arestas tiverem uma representação numérica esses valores serão chamados de *pesos*. Esses pesos podem caracterizar uma distância entre duas cidades, o custo do deslocamento, quantidade de fios nas redes de transmissões, sinais entre roteadores e outras atribuições. Define-se um *passeio* de um grafo $G = (V, A)$ como uma sequência finita de vértices e arestas $v_1, v_1v_2, v_2, v_2v_3, v_3 \dots, v_{n-1}, v_{n-1}v_n, v_n$ de G tal que $v_i \neq v_j$ para $i \neq j$ com $i, j \in \{1, \dots, n\}$, iniciando em v_1 terminando em v_n . No caso em que uma aresta "liga" num mesmo vértice é chamada de *laço*.

Temos uma *cadeia* que representa um passeio sem repetição de arestas. Já o *caminho* é uma cadeia sem repetição de vértices. Num grafo G é definido ser *conexo* se todo o par de

vértices v_i e v_j existem pelo menos um caminho entre v_i e v_j . No grafo ponderado o *comprimento* de um caminho é a soma dos pesos das arestas desse caminho. Os grafos apresentam vários tipos de caminhos, porém para nosso estudo será relevante o conhecimento sobre caminho mais curto em grafos conexos, pois com isso vamos analisar e obter o menor custo para a transportadora. Portanto um *caminho mais curto* entre dois vértices v_i e v_j de um grafo $G = (V, A)$ ponderado de pesos positivos é aquele caminho cuja soma dos pesos das arestas tem o menor valor possível dentre todos os caminhos que existem entre v_i e v_j .

4.1.1 Distância entre vértices

Seja $G = (V, A)$ um grafo conexo e ponderado com os pesos positivos, defina-se $d_G : V \times V \rightarrow \mathbb{R}$ por $d_G(v_i, v_j)$ como sendo a soma dos pesos dos vértices do menor caminho que une v_i e v_j e quando $v_i \neq v_j$. E define-se $d_G(v_i, v_j) = 0$ quando $v_i = v_j$.

Seja $G = (V, A)$ um grafo conexo e ponderado com valores positivos. Provaremos que a função $d_G : V \times V \rightarrow \mathbb{R}$, definida acima por $d_G(v_i, v_j)$ é uma métrica. Bastará verificar as condições da [Definição 1](#) para garantir que d_G seja uma métrica. As condições M_1 e M_2 são imediatas pois, $d_G(v_i, v_i) = 0$ para todo vértice $v_i \in V$, dado o menor caminho entre v_i e v_j , com $v_i \neq v_j$ pela condição no grafo em que os pesos são positivos, tem-se então $d_G(v_i, v_j) > 0$

Agora verificaremos as condições M_3 e M_4

Na condição M_3 , seja v_i e v_j dois vértices arbitrários do conjunto V . Veremos agora que $d_G(v_i, v_j) = d_G(v_j, v_i)$. Suponhamos que $d_G(v_i, v_j) \neq d_G(v_j, v_i)$. Temos as seguintes possibilidades:

$$d_G(v_i, v_j) > d_G(v_j, v_i) \text{ ou } d_G(v_i, v_j) < d_G(v_j, v_i)$$

Suponhamos que a primeira desigualdade é válida (o argumento na outra possibilidade é similar). Lembramos que $d_G(u, v)$ é a soma dos pesos do menor caminho entre u e v . Seja $v_j, v_j v_1, v_1, v_1 v_2, v_2, \dots, v_{k-1}, v_{k-1} v_k, v_i$ o menor caminho entre v_j e v_i , logo $d_G(v_j, v_i)$ é a soma dos pesos determinados por este caminho. Por outro lado se tomarmos o caminho,

$$v_i, v_k v_{k-1}, \dots, v_2, v_2 v_1, v_1, v_1 v_j, v_j$$

que une os vértices v_i e v_j temos que a soma dos pesos deste novo caminho é também $d_G(v_j, v_i)$, mas estamos supondo que $d_G(v_i, v_j) > d_G(v_j, v_i)$, logo temos um caminho entre v_i e v_j cujo a soma dos pesos é menor que $d_G(v_i, v_j)$ o qual é um absurdo. Logo, necessariamente $d_G(v_i, v_j) = d_G(v_j, v_i)$.

Para finalizar, vejamos agora a condição M_4 é satisfeita.

Seja v_i e v_j dois vértices quaisquer de V e seja $v_i, v_1 v_2, v_2, v_2 v_3, v_3, \dots, v_{k-1}, v_{k-1} v_k, v_k, v_k v_j, v_j$ o menor caminho entre v_i e v_j , logo $d_G(v_i, v_j)$ é a soma dos pesos deste caminho. Deve-se provar

a desigualdade triangular a seguir.

$$d_G(v_i, v_j) \leq d_G(v_i, v_p) + d_G(v_p, v_j)$$

Onde v_p um vértice arbitrário de V . Denote se o menor caminho entre v_i e v_p por

$$v_i, v_1 v_2, v_2, v_2 v_3, v_3 \cdots, v_{p-1}, v_{p-1} v_p, v_p$$

, logo $d_G(v_i, v_p)$ é a soma dos pesos deste caminho. Seja $v_p, v_p v_q, v_q, v_q v_{q+1}, v_{q+1} v_{q+2}, v_{q+2}, \cdots, v_{j-1}, v_{j-1} v_j, v_j$ o menor caminho entre v_p e v_j , então $d_G(v_p, v_j)$ é a soma dos pesos deste caminho. Sabe-se que $d_G(v_i, v_j)$ é a menor soma dos pesos entre v_i e v_j , então a soma entre $d_G(v_i, v_p) + d_G(v_p, v_j)$ é superior a de $d_G(v_i, v_j)$, pelo fato de $d_G(v_i, v_j)$ ser a menor soma dos pesos e de que

$$v_i, v_1 v_2, v_2, v_2 v_3, v_3 \cdots, v_{p-1}, v_{p-1} v_p, v_p, v_{j-1}, v_{j-1} v_j, v_j$$

é um caminho entre v_i e v_j , cuja a soma dos pesos é $d_G(v_i, v_p) + d_G(v_p, v_j)$. Portanto, $d_G(v_i, v_j) \leq d_G(v_i, v_p) + d_G(v_p, v_j)$ como queríamos provar.

4.1.2 Algoritmos para achar o caminho mais curto

O algoritmo de Edsger Wybe Dijkstra (1952) tem como objetivo determinar o caminho mais curto entre vértices de um grafo ponderado com pesos positivos. Com a escolha de um vértice inicial, o algoritmo calcula o caminho mais curto deste vértice inicial v_i para todos os outros vértices do grafo. Note-se que para aplicar o algoritmo de Dijkstra o grafo deve ser conexo. O algoritmo inicia com uma estimativa inicial para o caminho mais curto e vai progressivamente ajustando esta estimativa. Quando for obtido o caminho mais curto em relação ao vértice adotado como inicial até ele, denotaremos como vértice *fechado*, caso contrário o vértice será representado como *aberto*.

A seguir apresentaremos aplicação do algoritmo na situação problema da transportadora, em que queremos percorrer o menor caminho para ter o custo mínimo de suas viagens.

Uma transportadora quer determinar o menor caminho conseqüentemente o custo mínimo, para ir da empresa até o lugar de entrega das mercadorias. Esses lugares são representados por vértices e as arestas os caminhos que interligam os estabelecimentos. Já os pesos serão representados pelos custos de cada viagem. O grafo mostrado na [Figura 35](#) deve ser analisado pela empresa para decidir qual percurso seguir pelo caminhão para as entregas.

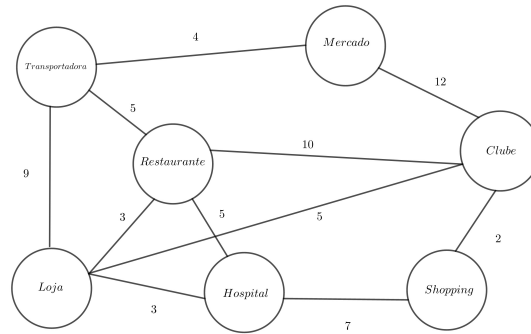


Figura 35 – Grafo representando os pontos de entrega da transportadora e os caminhos com os "custo"(pesos) de cada viagem

Para um melhor entendimento do algoritmo de Dijkstra, faremos uma descrição de forma comentada. Iniciaremos colocando o símbolo ∞ para interpretar os vértices ainda não foi analisados pelo processo.

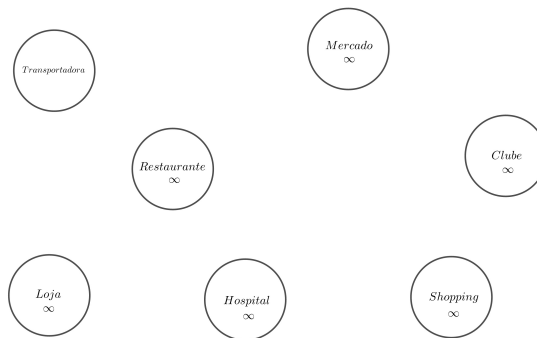


Figura 36 – Iniciando o algoritmo de Dijkstra

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
	M - mercado	∞	aberto
	R - restaurante	∞	aberto
	L - loja	∞	aberto
	C - Clube	∞	aberto
	H - hospital	∞	aberto
	S - shopping	∞	aberto

Tabela 6 – Início do algoritmo

Primeiramente devemos obter o menor percurso. Para iniciar, temos o ponto de partida representado pela transportadora, sendo assim o custo 0. Com isso devemos agora verificar os vértices adjacentes à transportadora, assim temos as seguintes opções: ir ao mercado, com custo 4, ou ao restaurante com o custo 5 ou a loja com o custo 9. Portanto o menor custo neste caso é ir até o mercado, assim dizemos o precedente do mercado é a transportadora.

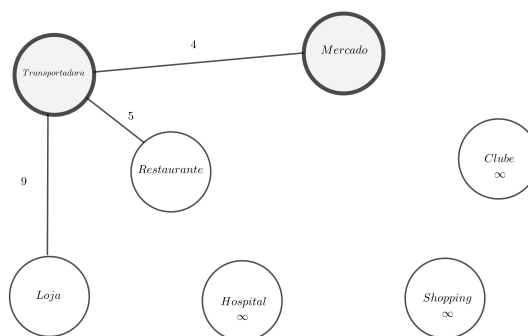


Figura 37 – Analisando o caminho até o mercado e fechando o vértice do mercado que precede da transportadora

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
	R - restaurante	∞	aberto
	L - loja	∞	aberto
	C - Clube	∞	aberto
	H - hospital	∞	aberto
	S - shopping	∞	aberto

Tabela 7 – Fechando o caminho até o mercado

Lembrando que o caminho até o mercado não deve ser alterado, pois qualquer outra decisão de caminho será maior que 4, então dizemos que o vértice mercado é fechado. Agora a partir do mercado temos apenas uma alternativa para chegar ao clube precedente do mercado, com isso o total de custo até o clube será de $4 + 12 = 16$, logo obtém um caminho, da transportadora até o clube passando pelo mercado, porém ainda não consideramos o vértice representado pelo clube como fechado, pois ainda devemos analisar outras opções de caminhos que chegam a ele. Agora verificaremos o caminho da transportadora até o restaurante será direto, pois qualquer outro caminho terá um custo maior, com isso podemos torná-lo um vértice fechado.

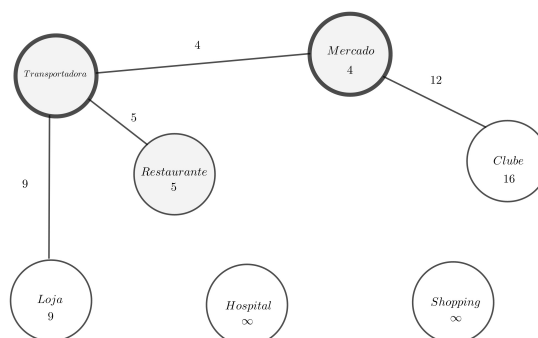


Figura 38 – Analisando o caminho até o mercado e fechando o vértice do mercado que precede da transportadora

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
Fechado	R - restaurante	5	T-transportadora
	C - Clube	16	aberto
	L - loja	∞	aberto
	H - hospital	∞	aberto
	S - shopping	∞	aberto

Tabela 8 – Fechando o caminho até o restaurante

Dando procedência na investigação, iremos analisar a partir do restaurante, em que podemos chegar em três lugares, como loja, hospital e clube. Veja a seguir todas essas possibilidades.

Loja: Temos que $5 + 3 = 8$ o caminho entre a transportadora e a loja passando pelo restaurante. Note-se que ao comparar a ligação da transportadora direto com a loja obtemos $8 < 9$, assim o menor custo até a loja é indo pelo restaurante, portanto fechamos o vértice da loja que precede do restaurante.

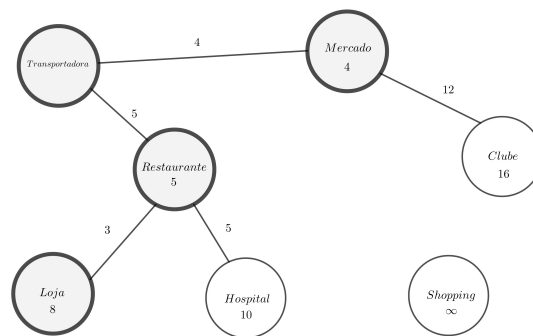


Figura 39 – Analisando os caminhos até a loja

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
Fechado	R - restaurante	5	T-transportadora
Fechado	L - loja	8	R - restaurante
	C - Clube	16	M - mercado
	H - hospital	10	aberto
	S - shopping	∞	aberto

Tabela 9 – Fechando o caminho até a loja

Hospital: Temos que $5 + 5 = 10$ o caminho do restaurante até o hospital. Outra opção de caminho será do restaurante passando pela loja ($5 + 3 + 3 = 11$), logo, podemos concluir que o menor caminho da transportadora até o hospital será passando pelo restaurante e chegando direto ao hospital. Então, fechamos o vértice do hospital que precede do restaurante.

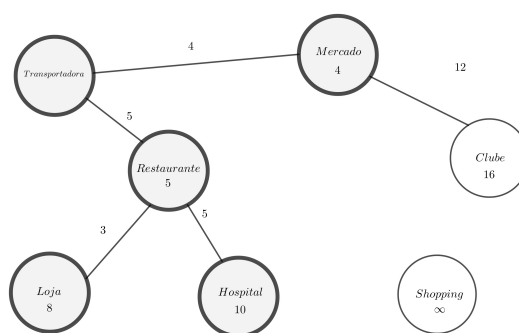


Figura 40 – Analisando o caminho até hospital

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
Fechado	R - restaurante	5	T-transportadora
Fechado	L - loja	8	R - restaurante
	C - Clube	16	L - loja
Fechado	H - hospital	10	R - restaurante
	S - shopping	∞	aberto

Tabela 10 – Fechando o caminho até hospital

Clube: Temos que $5 + 3 + 5 = 13$ é o custo do caminho do restaurante, passando pela loja e assim chegando ao clube. No caso de ir do mercado ao clube, obtemos $4 + 12 = 16$ e outra opção é ir do restaurante direto ao clube, obtendo $5 + 10 = 15$. Portanto, o melhor caminho da transportadora até o clube é saindo da dela, passando pelo restaurante, depois pela loja e por fim até o clube, logo podemos fechar o vértice do clube que precede da loja.

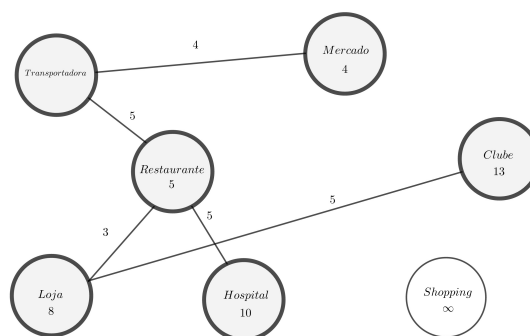


Figura 41 – Analisando o caminho até o clube

Contudo falta o último vértice a ser fechado, teremos duas opções para alcançar o shopping, uma será passando pelo hospital e a outra pelo clube.

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
Fechado	R - restaurante	5	T-transportadora
Fechado	L - loja	8	R - restaurante
Fechado	C - Clube	13	L - loja
Fechado	H - hospital	10	R - restaurante
	S - shopping	∞	aberto

Tabela 11 – Fechando o caminho até o clube

Se optarmos pelo hospital temos o custo de $5 + 5 + 7 = 17$, no caso de irmos pelo clube obtemos $5 + 3 + 5 + 2 = 15$, portanto o menor caminho é seguir pelo clube, logo, fechamos o último vértice do shopping que precede do clube.

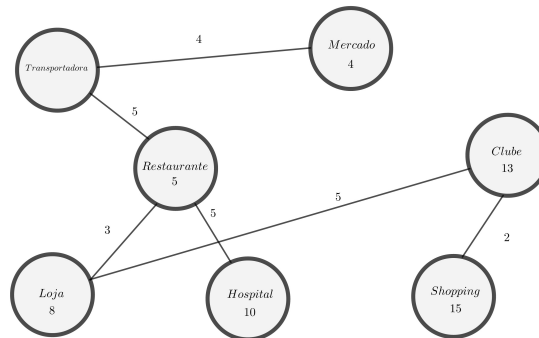


Figura 42 – Analisando o caminho até o shopping

Fechado	Ir até o ponto de entrega	Custo	Precedente
	T - transportadora	0	—
Fechado	M - mercado	4	T-transportadora
Fechado	R - restaurante	5	T-transportadora
Fechado	L - loja	8	R - restaurante
Fechado	C - Clube	13	L - loja
Fechado	H - hospital	10	R - restaurante
Fechado	S - shopping	15	H - hospital

Tabela 12 – Fechando o caminho até o shopping

Portanto, vimos a importância de obtermos a menor distância (custo) num grafo convexo e ponderado. Com o algoritmo de Dijkstra analisamos os caminhos do grafo e obtivemos o menor custo entre a transportadora até um ponto de entrega, veja [Figura 42](#)

Observação 1. O algoritmo na forma computacional encontrar-se na referência [Goldbarg e Goldbarg \(2012\)](#).

DISTÂNCIA ENTRE GENES

5.1 Distância entre genes

O propósito do capítulo é discutir um pouco a distância entre genes no campo da genética, principalmente discutir se o conceito de distância entre genes pode ser caracterizado como uma métrica. A seguir, serão lembrados alguns conceitos básicos de genética, afinal, trata-se de um assunto cujas as definições não são familiares na matemática.

5.2 Definições

Na Biologia, a genética é a área responsável pelo estudo da herança biológica, que identifica e analisa as características que são transmitidas de pais para filhos. Na realidade o que herdamos dos nossos pais não são suas características, e sim as informações que produzem essas características. Essas informações estão relacionadas a segmentos da molécula de DNA denominados *genes*, que denota-se como G (um conjunto de genes representados por letras do alfabeto grego), que são transmitidos de pais para filhos através de seus gametas (*espermatozoide ou óvulo*). O DNA é uma molécula que está no núcleo das células de qualquer ser vivo e que carrega todas as informações genéticas de um organismo. Os *cromossomos* são estruturas compostas de DNA que possuem os genes. Já os *cromossomos homólogos* são cromossomos pareados entre si (lado a lado), semelhantes e que juntos formam um par, formados no processo de fecundação. Os genes alelos são aqueles que ocupam o mesmo locus (posição que um determinado gene que ocupa um cromossomo) em cromossomos homólogos e estão envolvidos na determinação das características. Alguns genes são dominantes, e serão representados pelas letras maiúsculas do alfabeto. As características dominantes necessitam de no mínimo um dos alelos do par de homólogos para se expressar. As características podem ser homozigotas formadas por alelos iguais, ou heterozigotas, formadas por um alelo dominante e um recessivo. Já os genes recessivos serão representados pelas letras minúsculas do alfabeto. Estes necessitam

estar em homozigose (homozigoto) no par de cromossomos homólogos para se expressarem. Veja na [Figura 43](#) a localização do DNA e dos genes em um cromossomo e onde se encontra o cromossomo na célula.

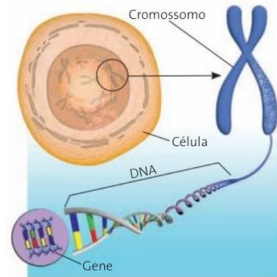


Figura 43 – Localização do DNA e dos genes no cromossomo

Fonte – [Usberco et al. \(2012\)](#)

O estudo da distância entre genes iniciou-se com o geneticista Thomas Hunt Morgan (1866 – 1945) com moscas *Drosófilas melanogaster*, ele criou centenas dessas moscas e com isso obteve várias gerações para seu experimento. Com uma grande variação de gerações, as características foram passando de pais para filhos e com isso surgiram várias suposições e conclusões. Diante dessas observações, conseguiu desenvolver, juntamente com sua equipe de estudo, o primeiro mapa de posição de genes sobre um cromossomo (mapa genético). Com a descoberta de Morgan, vários cientistas começaram a se aprofundar nesta área da genética e em 1987 surgiu o estudo do primeiro projeto do genoma humano (*P.H.G.*), este projeto teve como base o mapeamento genético realizado por Morgan.

O mapeamento genético criado por Morgan através das *Drosófilas* teve como base a *taxa de recombinação entre genes*. Mas para compreender sobre essa taxa, deve-se antes compreender o que é ligação gênica ou *linkage*.

5.3 Ligação gênica ou linkage

Sabe-se que o número de genes é muito superior ao número de cromossomos em um organismo. Por essa razão, cada gene não tem seu próprio cromossomo, portanto, alguns ficam localizados no mesmo, e com isso muitos acabam alinhados em sequência em cada cromossomo e uma parte acaba ficando próxima das outras. Quando estão situados no mesmo cromossomos se diz que estão em *linkage* ou em ligação. Existem várias nomenclaturas para representar genes em *linkage*, veja a representação de genes em linkage na figura [Figura 44](#).

Quando esses genes estão muitos distantes no mesmo cromossomo, ou em cromossomos separados, os mesmos segregam independentemente. Com isto, genes que irão para os gametas, o alelo recebido para um gene não afeta o alelo recebido de outro, com isso reforça a segregação independente.

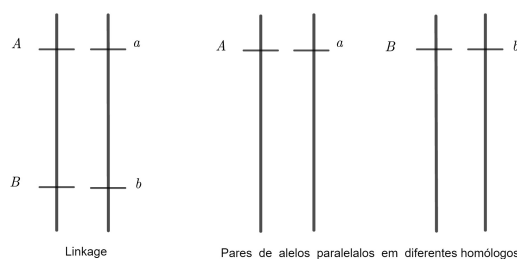


Figura 44 – Representação de genes em linkage

Nas pesquisas realizadas por Morgan, foram obtidos resultados em que novas características foram observadas, diferente do esperado - esperava-se encontrar a proporção 1 : 1 : 1 : 1, característica da segregação independente. Morgan esperava encontrar características que estivessem presentes em cromossomos distintos, mas na realidade eles estavam lidando com genes localizados no mesmo cromossomo (linkage).

Dessa forma, ele começou a investigar o motivo dessa diferença na proporção. E percebeu que no decorrer da divisão celular, especificamente na meiose (processo de formação de gametas), os genes ligados tendem a permanecer juntos. Porém, no início da divisão meiótica ocorre a troca de pedaços entre cromátides homólogas (Figura 45). Esse processo é conhecido como *permutação* ou *crossing-over*.

Na Figura 45 a seguir dois esquemas, meiose sem crossing-over ou com crossing-over são apresentados

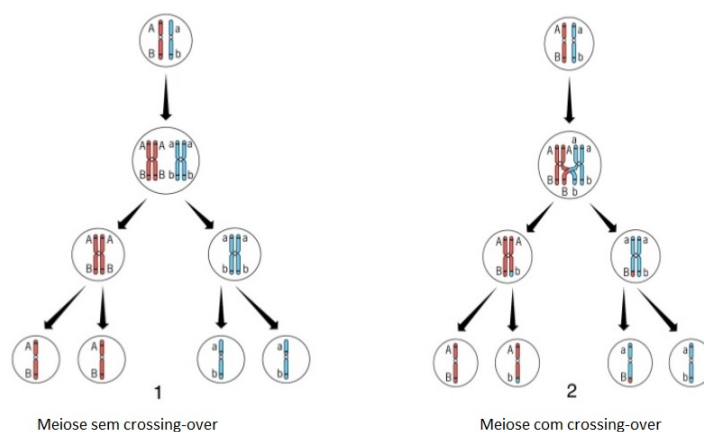


Figura 45 – Meiose sem crossing-over (1) ou com crossing-over(2)

No processo de meiose sem crossing-over, 100% dos gametas mantêm a ligação gênica igual a célula que iniciou o processo de divisão celular. Já no processo de meiose com crossing-over, 50% das células serão *recombinantes*.

5.4 Determinação de taxa de recombinação

Suponha-se que num cruzamento ilustrativo entre 1200 moscas *Drosófilas melanogaster* há o interesse em analisar se dois genes das moscas estão ligados um ao outro e, em caso afirmativo, o quão forte ligados estão. Para obter a frequência de recombinação, utilizará o cruzamento ilustrativo (veja Figura 46). Mas antes existem algumas restrições a serem consideradas.

Será preciso uma mosca que não seja apenas heterozigota para ambos os genes, além disso será preciso saber quais genes estão juntos no cromossomo. O cruzamento das moscas com características distintas, como corpo cinzas e asas longas e moscas com corpo preto e asas curtas, geraram descendentes que apresentaram corpos cinzas e asas longas. Com isso, conclui-se que o gene (P) que representa o corpo cinza era gene dominante e o (p) que representa o corpo preto o gene recessivo, além do gene (V) que representa a asa longa, dominante sobre o (v) recessivo que representa asa curta, assim tendo formado a geração $F1$.

Em um novo cruzamento foi realizado entre indivíduos, moscas machos duplos recessivos $ppvv$ de corpo preto e asas curtas e as fêmeas da geração $F1$ de corpo cinza e asas normais $PpVv$. A finalidade de usar uma mosca macho duplos recessivos $ppvv$ no cruzamento, é garantir que os alelos fornecidos pela genitora duplo heterozigoto possam determinar totalmente o fenótipo, ou aparência, da descendência. No cruzamento da mosca da geração $F1$ e o macho $ppvv$, pode-se "ler" diretamente o genótipo de cada gameta a partir da aparência física da descendência.

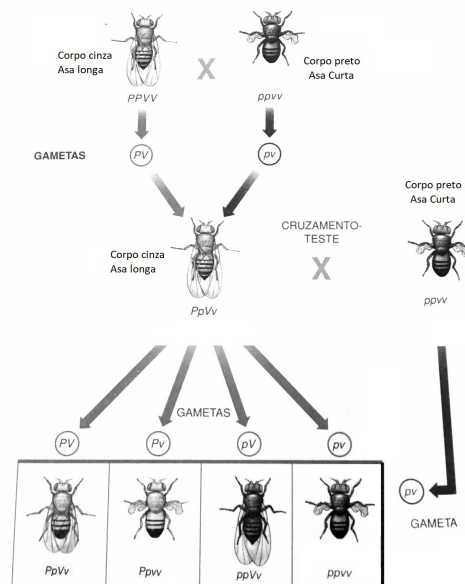


Figura 46 – Cruzamentos

Fonte – Amabis e Martho (2004)

PpVv - 532 indivíduos (parental)
 ppvv - 464 indivíduos (parental)
 Ppvv - 104 indivíduos (recombinante)
 ppVv - 100 indivíduos (recombinante)

Na [Figura 46](#) é mostrado o resultado da quantidade de moscas que foram geradas. Se observa que as quantidades são geradas de maneiras não uniformes e de característica distintas, o que mostra que os genes do corpo preto e asas curtas estão ligados. Como é esperado para os genes ligados entre si, as configurações do cromossomo parental tem uma maior representatividade e os cromossomo recombinantes uma representatividade menor, em relação toda a prole. Para medir a quantidade da ligação dos genes, pode-se calcular a *taxa de recombinação* (tx) entre o corpo preto e asas curtas, da seguinte forma:

$$\text{taxa de recombinação } (tx) = \frac{\text{recombinantes}}{\text{total de indivíduos}} \cdot 100\%$$

Vale lembrar que as moscas descendentes recombinantes são de corpo preto e asas longas, e corpo cinzas e asas curtas. Para poder identificar quais as moscas são recombinantes temos dois caminhos: a primeira é de conhecimento que a partir dos cruzamentos realizados, as moscas tem herdado um cromossomo da mãe que provém de um evento de recombinação; segunda modo é ver se a quantidade de descendentes recombinantes tem uma quantidade de descendentes menor que as parentais. Calculando a taxa de recombinação do experimento com as moscas temos.

$$tx = \frac{100 + 104}{532 + 464 + 100 + 104} = \frac{204}{1200} = 0,17 = 17\%$$

Então, a taxa de recombinação entre os genes de corpo preto e e asas curta é de 17%

Com os cálculos da taxa de recombinação foi possível a construção de mapas genéticos, e com o estudo de linkage os cientistas conseguiram concluir que a distribuição dos genes nos cromossomos é do tipo lineares e que cada gene tem o seu lugar específico no cromossomo.

As taxas de recombinação não é uma medida física entre os genes que estão no cromossomo, porém, dão uma estimativa ou uma aproximação da distância física no qual, quanto menor a frequência da recombinação entre os genes, mais fortemente estarão ligados, já com uma frequência maior, estarão ligados com menor intensidade. Em resumo, quanto maior a distância entre os genes, maior será a a probabilidade de ocorrer um crossing-over. Quanto menor a distância entre os genes, menor a probabilidade de ocorrer o crossing-over entre eles. Assim, a taxa de recombinação (tx) está relacionada com a distância entre genes em *linkage*. O valor da taxa de recombinação dependerá do número de células que realizam o crossing-over. Vale lembrar, que mesmo nessas células, apenas metade dos gametas produzidos serão recombinantes, pois a troca acontece apenas com as cromátides que se situam no meio do par de homólogos emparelhados.

5.4.1 Construção dos mapas gênicos

Os primeiros cientistas que estudaram o mapa gênico foram Morgan e Alfred Sturtevent, eles afirmaram que era possível construir o mapa gênico baseado nas taxas de recombinações. Na verdade, os estudos sobre *linkage* ajudou a estabelecer que os cromossomos eram, na verdade "lineares" e que cada gene teria um local específico (posição) no cromossomo.

Vale lembrar que a taxa de recombinação não é uma medida direta em que os genes estão fisicamente distantes entre eles, mas fornecem uma estimativa ou aproximação da distância física entre genes. Logo, pode-se dizer que um par de genes com uma taxa de recombinação maior estão provavelmente mais distante e analogamente com a taxa menor, estão provavelmente mais próximos.

Sabe-se que a taxa de recombinação tem como sua maior frequência de 50%. Portanto, caso queira descobrir a distância entre genes que estão mais afastados do que isso, deve-se adicionar as frequências de vários pares de genes, contudo "construindo" um mapa que se estende entre os dois genes distantes. A comparação das taxas de recombinação também é utilizada para descobrir a ordem dos genes no cromossomo.

No estudo que Morgan e Sturtevent realizaram envolviam três locus gênicos da mosca drosófila: *yellow* (*y*), *vermilion* (*v*) e *miniature* (*m*). No experimento visto livro [Amabis e Martho \(2004\)](#) realizado com essas moscas, concluíram que a taxa de recombinação entre *y* e *v* era de 32,2% e de *y* e *m* era de 35,5%. Pelo resultado apresentado tem-se que *y* está mais perto de *v* do que *m*, entretanto com apenas essas informações não conseguiriam saber em que ordem os esses genes poderiam entrar, isto é, *v* está entre *y* e *m*, ou se *y* está entre *v* e *m*. Então, Sturtevent precisou descobrir a taxa de recombinação entre *v* e *m*, pela hipótese essa taxa poderia ser de 67,7% pelo fato da soma de 32,2% + 35,5% ou de ser 3,3% pelo motivo da subtração de 35,5% e 32,2%. Assim, Alfred focou o estudo no cruzamento de uma fêmea duplo heterozigoto $VvMm$ com um macho duplo recessivo e com isso obteve a taxa de recombinação entre os locus *v* e *m* em 3%. Após esse resultado, conseguiu concretizar o posicionamento de *m*, garantindo que a partir da sua hipótese era possível construir um mapa gênico para as moscas drosófilas e para qualquer outra espécie.

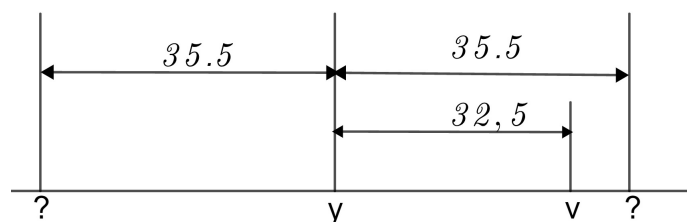


Figura 47 – Posicionamento dos genes *y*, *v* e *m*

Portanto, as taxas de recombinações estão relacionadas com a distância entre genes no

mesmo cromossomo e essa distância tem sua unidade de medida conhecida como unidade de recombinação *UR* ou *centimorgan*, uma homenagem a Thomas Morgan. E existe uma relação direta entre esses valores: uma taxa de recombinação de $1\% = 1UR$ é equivalente a 1 *centimorgan* ou 1 *unidade de mapa*.

5.4.2 Sob o conceito de espaço métrico e o mapeamento genético

Sabe-se que para definir uma métrica tem que satisfazer as condições da [Definição 1](#). A importância dessa verificação é para analisar se a ideia de distância apresentada acima entre genes se encaixa nas condições de métricas definidas na matemática.

Dado o conjunto G com dois ou mais genes em linkage (mesmo cromossomo) defina-se a função $d : G \times G \rightarrow \mathbb{R}$, em que A e B são genes do conjunto G por:

$$d(A, B) = \text{taxa de recombinação entre } A \text{ e } B$$

Verificaremos a veracidade de cada propriedade.

M_1 : A taxa de recombinação nula só acontece se o gene estiver no mesmo locus gênico. Logo temos que $d(A, A) = 0$.

M_2 : Seja dois ou mais genes heterozigotos em G . Na divisão celular quando ocorrer o crossing-over, temos pela definição uma porcentagem da taxa de recombinação, logo $d(A, B) > 0$

M_3 : Dados dois ou mais genes heterozigotos em G . Com o crossing-over temos uma taxa de recombinação entre dois genes A e B . Esses genes estão em linkage. Dado como referência o gene A o gene B pode estar localizado tão pela direita ou pela esquerda. Logo, podemos concluir que $d(A, B) = d(B, A)$

M_4 : Na desigualdade triangular vale a desigualdade do maior ou igual, porém na taxa de recombinação teremos apenas a igualdade entre genes de G . Pois os genes devem estar em linkage, ocupando posições distintas em uma estrutura linear. Dessa forma, a somatória das taxas entre os genes internos será "igual" a soma das taxas dos genes nas extremidades e podemos considerar que

$$d(A, B) = d(A, C) + d(C, B)$$

Outras distâncias que são estudadas na biologia podem ser encontradas no livro [Deza Michel Marie \(2009\)](#).

ATIVIDADE E CONCLUSÃO

6.1 Atividade com alunos do ensino médio

A atividade foi direcionada para alunos do ensino médio, trazendo uma proposta de estimular e mostrar que o conceito de distância, embora seja natural, e necessário realizar uma formalização matemática. Também foi mostrado que esse conceito está presente em ações e momentos cotidianos, além disso foi discutido como eles entendem e definem o que é uma distância.

Antes de iniciar a atividade, os alunos responderam um questionário ([Figura 56](#)), cujo teor das questões estava voltado para uma análise na formação do conceito que eles têm sobre distância. Também foi questionado em que momento do cotidiano fazem o uso dela e colocamos desafiados para discutir como deve ser medida a distância entre palavras, usando o corretor de texto dos aplicativos de smartphones. (questionário parte A). Na segunda etapa, foram apresentados quatro objetos distintos (veja na [Figura 49](#)) para que tentassem determinar as distâncias entre os pontos já pré-determinados neles.



Figura 48 – Distância numa esfera e num objeto plano



Figura 49 – Distância entre pontos em objetos com curvatura negativa

Essas medições primeiramente foram feitas com barbantes com marcações de espaçamento de 10cm e em seguida repetiram as mesmas medições porém com fita métrica. A turma foi separada em grupos de 4 a 6 alunos. Em cada objeto foram marcados dois pontos e cada grupo foi desafiado, sendo que durante a atividade aconteceram alguns fatos interessantes. Em alguns grupos os integrantes discutiam entre eles sobre a estratégia para medir, se o barbante e a fita métrica passavam por dentro ou por fora dos objetos. Já em outro grupo questionavam as medidas, manipulando para obter a menor distância entre os pontos.

Outro momento interessante foram as comparações dos resultados das medições realizadas com barbante e fita métrica, vários grupos chegaram próximo em relação aos valores entre as medidas. Porém, quando questionados qual dos resultados escolheriam como definitivo eles optaram pela fita métrica, mostrando, assim, que é importante saber a ferramenta "métrica" mais adequada para se fazer uma medição com maior precisão dentro do contexto que está trabalhando.

Contudo, com as respostas dos questionários e as discussões, foi possível concluir dois aspectos que foram bem pontuados por eles. A primeira conclusão foi que nem sempre a distância é dada pelo comprimento de um segmento de reta e o segundo aspecto é que, ao efetuar a distância entre elementos, deve-se obter o menor caminho. Mas, o fato que mais chamou atenção para eles foi a medida entre palavras pois foram desafiados a enviar uma frase por meio de um aplicativo de



Figura 50 – Medição com barbante e fita métrica

smartphone. No desafio, um integrante de cada grupo deveria digitar a frase dita pelo professor, porém, tinham que digitar em poucos segundos e quando digitavam alguma palavra errada era preciso escolher a primeira sugestão que o corretor ortográfico fornecia. Veja a seguir algumas frases digitadas pelos alunos e corrigidas pelo corretor ortográfico.

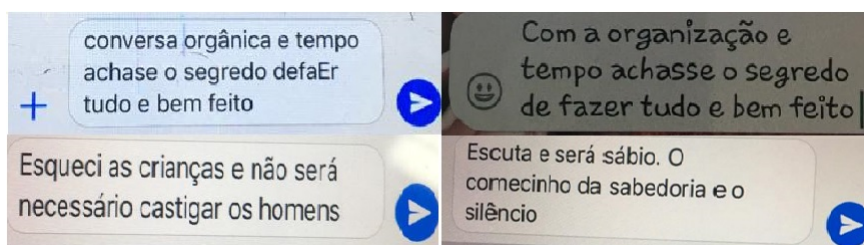


Figura 51 – Frases digitadas pelos alunos e corrigidas com a sugestão do corretor ortográfico

Frases ditadas pelo professor.

- Com organização e tempo, acha-se o segredo de fazer tudo e bem feito - Pitágoras
- Eduquem as crianças, para que não seja necessário punir os adultos - Pitágoras
- Escuta e serás sábio. O começo da sabedoria é o silêncio - Pitágoras

Essa dinâmica mostrou que mesmo digitando uma palavra errada o corretor nem sempre consegue dar a melhor opção ou elas não fazem parte do contexto, fazendo que a frase não tenha sentido. Com isso, foi questionado se existe um conceito de distância entre as palavras digitadas com as sugestões do corretor. Afinal, as palavras sugeridas são semelhantes ou "próximas" da palavra digitada de forma incorreta.

Diante do questionamento de como deveria ser feita a medição entre palavras, muitos não tiveram a menor ideia de como deve ser efetuar essa distância, já alguns arriscaram um palpite. O fato é que ao apresentar um conceito de forma dinâmica e próxima ao cotidiano dos alunos, faz despertar o interesse e mostrar que a matemática vai além de uma ciência que trabalha só com teorias, mas com várias aplicações. Vejamos, a seguir, algumas respostas do questionário.

Nesta questão os alunos mostraram a ideia de distância como sempre um linha reta, não pensaram calcular pela superfície esférica.

3. E a distância entre duas capitais de dois países? Também será o comprimento de uma linha reta unindo as duas capitais?

Do centro de uma capital ao centro da outra. Através do Arcos, através, através do dados geográficos pela superfície, e medir pelo interior da planície através de uma linha reta.

Figura 52 – Questão 3: parte A

4. Atualmente os aplicativos de mensagens dos smartphones possuem corretores automáticos de textos, em que ao digitar uma palavra errada sugerem outras similares. Como o grupo acha que é realizado essas sugestões de palavras? Qual critério de proximidade vocês acham que o corretor utiliza?

As sugestões são feitas segundo um padrão de similaridade (similar) entre duas letras consecutivas, fazendo uma triangulação, entre palavras com pedras precisas.

Figura 53 – Questão 4: parte A

Esse grupo usou a ideia da comparação entre as letras, com isso não achou uma relação de comprimento de um segmento de reta com distância entre palavras.

4. Afim de toda atividade e esclarecimento e discussões. Houve uma mudança na perspectiva do grupo em relação ao conceito de distância? O que puderam aprender com essa atividade?

Aprendemos que o "arbitrário" (uso de barbafeita) não dá a medida precisa; e a dificuldade existente para medir objetos em distâncias que possuem desníveis ou obstáculos

Figura 54 – Questão 4: parte B

Os estudantes deste grupo mostraram que um dos conceitos de distância deve se utilizar a ferramenta mais adequada conforme o meio em que estão sendo feitas as medições para ter uma precisão sobre o valor da distância.

4. Afim de toda atividade e esclarecimento e discussões. Houve uma mudança na perspectiva do grupo em relação ao conceito de distância? O que puderam aprender com essa atividade?

Sim, o grupo percebeu que uma distância não precisa ser necessariamente em linha reta porque há casos em que podem existir impedimentos/barreiras (como, por exemplo, montanhas, rios ou depressões que estão entre duas cidades). O grupo aprendeu que ao medir a distância entre dois objetos quaisquer, deve-se sempre considerar o caminho de menor distância.

Figura 55 – Questão 4: parte B

Já neste os integrantes concluíram que a distância deve ser o menor caminho.

6.2 Considerações finais

Este trabalho contribuiu para meu enriquecimento no conteúdo matemático, principalmente na importância de formalização matemática dos conceitos de distância. Tive a oportunidade de compreender e aprofundar a questão teórica de métrica, além de estudar vários modelos e poder ver aplicações importantes e interessantes que utilizam o conceito de distância. Em relação a atividade, o resultado foi, de modo expressivo, positivo, pois alguns alunos conseguiram argumentar e discutir que distância é muito mais que um comprimento entre dois pontos. Além disso, essa atividade serviu como incentivo para elaborar outras atividades que possam despertar nos estudantes o interesse, não só por distância, e sim pela matemática.

CAPÍTULO
7

ANEXOS

Questionário sobre métricas

Parte A → GRUPO 02

1. Como o grupo calcularia a distância entre dois pontos quaisquer da sala? E quais são esses pontos?

O grupo calcularia utilizando uma trena. Esses pontos poderiam ser dois quaisquer na sala: a distância entre a porta e uma janela, por exemplo.

2. Como o grupo acha que se mede a distância entre duas cidades? É o comprimento uma linha reta unindo esses pontos?

O grupo acha que mede-se considerando o ponto central de ambas as cidades. Desse modo, uma linha reta unindo os pontos centrais se não houvesse impedimentos no caminho (como casas, prédios, construções, etc.). Mas como há uma malha seria medir a distância através de ruas que andam nas estradas.

3. E a distância entre duas capitais de dois países? Também será o comprimento de uma linha reta unindo as duas capitais?

De centro de uma capital ao centro da outra. Através do Aciculato, Usnadax, através do dados geográficos pela superfície, e medir pela interação da planície tornando de uma linha reta.

4. Atualmente os aplicativos de mensagens dos smartphones possuem corretores automáticos de textos, em que ao digitar uma palavra errada sugerem outras similares. Como o grupo acha que é realizado essas sugestões de palavras? Qual critério de proximidade vocês acham que o corretor utiliza?

As sugestões e o critério são feitas a partir de palavras com escrita "próximas" (poucas letras em comum e em uma sequência adequada).

5. Vista na questão anterior em relação ao corretor, podemos afirmar que existe um tipo de distância entre palavras?

Distanciamento de vocábulos, número de letras, significação, depende do critério utilizado.

Figura 56 – Questionário parte A

Questionário sobre métricas

Parte B

1. Quais dos objetos o grupo teve mais dificuldade para medir a distância? O grupo encontrou alguma dificuldade para determinar a distância? Em caso afirmativo, descreva?

O grupo encontrou dificuldade para medir as distâncias nos objetos "bolid" e "urso de madeira" pela forma da superfície deles não ser plana (o que dificultou a medição com o barômetro e, em menor grau, com a fita métrica flexível).

2. Após as medições, o grupo obteve os valores das distâncias entre os pontos marcados. Esses valores foram obtidos com um consenso de todos? Ou existiram diversas sugestões distintas?

Houve sugestões distintas na medição com o barômetro por não apresentar muita precisão. Porém, com a fita métrica, houve o consenso de todos.

3. Quais fatores influenciaram para concluir que a distância entre os pontos marcados nos objetos era a verdadeira distância?

Inicialmente, houve a dúvida coletiva de como fazer a medição porque havia vários caminhos e cada um resultava em um valor. No entanto, tendo por base a lógica, o grupo concluiu que o caminho a ser seguido seria aquele de menor comprimento/distância.

4. Afim de toda atividade e esclarecimento e discussões. Houve uma mudança na perspectiva do grupo em relação ao conceito de distância? O que puderam aprender com essa atividade?

Sim, o grupo percebeu que uma distância não precisa ser necessariamente em linha reta porque há casos em que podem existir impedimentos/barreras (como, por exemplo, montanhas, rios ou depressões que estão entre duas cidades). O grupo aprendeu que ao medir a distância entre dois objetos, qual quer, deve-se sempre considerar o caminho de menor distância.

Figura 57 – Questionário parte B

REFERÊNCIAS

- AMABIS, J. M.; MARTHO, G. R. **Biologia das populações**. 2nd. ed. São Paulo: Moderna, 2004. 57 p. Citado nas páginas 92 e 94.
- BURDEN, R. L.; FAIRES, J. D. **Análise numérica - tradução da 8 edição norte - americana**. 8nd. ed. São Paulo: Cengage Learning, 2008. 400-461 p. Citado nas páginas 39 e 42.
- DEZA MICHEL MARIE, D. E. **Encyclopedia of Distances**. 1nd. ed. New York: Springer, 2009. 17, 354, 388 p. Citado nas páginas 11, 13, 40 e 95.
- GOLDBARG, M.; GOLDBARG, E. **Grafos: Conceitos, algoritmos e aplicações**. 1nd. ed. Rio de Janeiro: Elsevier, 2012. 200 p. Citado na página 88.
- HEFEZ, A.; VILLELA, M. L. T. **Códigos Corretores de Erros**. 2nd. ed. Rio de Janeiro: IMPA, 2008. 2,6 p. Citado nas páginas 67 e 70.
- LIMA, E. L. **Espaços Métricos**. 3nd. ed. Rio de Janeiro: Projeto Euclides, 1993. 16-19 p. Citado nas páginas 66 e 78.
- POOLE, D. **Álgebra Linear**. 2nd. ed. São Paulo: Cengage Learning, 2004. 337,525 p. Citado nas páginas 52, 59 e 67.
- SANTOS, J. B. d. **Distância entre curvas algébricas: Uma discussão para o ensino médio**. 57-58 p. Dissertação (Mestrado) — Unicamp, São Paulo, 2016. Citado na página 75.
- SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. e. **Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos**. 1nd. ed. São Paulo: Pearson Prentice Hall, 2003. Citado na página 36.
- USBERCO, J.; MARTINS, J. M.; SCHECHTMANN, E.; FERRER, L. C.; VELLOSO, H. M. **Companhia das ciências**. 1nd. ed. São Paulo: Saraiva, 2012. 223 p. Citado na página 90.

