

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS BLUMENAU
LICENCIATURA EM MATEMÁTICA

Valdir Wille Neto

Geometria de distâncias: Explorando com álgebra linear

Blumenau
2022

Valdir Wille Neto

Geometria de distâncias: Explorando com álgebra linear

Trabalho de Conclusão de Curso de Graduação em Licenciatura em Matemática do Campus Blumenau da Universidade Federal de Santa Catarina para a obtenção do título de Licenciado(a) em Matemática.
Orientador: Prof. Felipe Delfini Caetano Fidalgo, Dr.

Blumenau
2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Wille Neto, Valdir
Geometria de distâncias : Explorando com álgebra linear
/ Valdir Wille Neto ; orientador, Felipe Delfini Caetano
Fidalgo, 2022.
72 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Campus Blumenau,
Graduação em Matemática, Blumenau, 2022.

Inclui referências.

1. Matemática. 2. Problema de geometria de distâncias
moleculares. 3. Álgebra linear. 4. Decomposição em valores
singulares. I. Fidalgo, Felipe Delfini Caetano. II.
Universidade Federal de Santa Catarina. Graduação em
Matemática. III. Título.

Valdir Wille Neto

Geometria de distâncias: Explorando com álgebra linear

Este Trabalho Conclusão de Curso foi julgado adequado para obtenção do Título de Licenciado(a) em Matemática e aprovado em sua forma final pelo Curso de Licenciatura em Matemática.

Blumenau, 11 de março de 2022.

Prof. Júlio Faria Corrêa, Dr.
Coordenador(a) do Curso

Banca Examinadora:

Prof. Felipe Delfini Caetano Fidalgo, Dr.
Orientador
UFSC- Blumenau

Prof. André Vanderlinde da Silva, Dr.
Avaliador
UFSC- Blumenau

Prof. Luiz Rafael dos Santos, Dr.
Avaliador
UFSC- Blumenau

Dedico este trabalho à minha mãe, que sempre me incentivou e levantou nos momentos mais difíceis, e nunca desistiu. Ao meu pai que me olha e guarda, lá onde brilham as estrelas que nos enchem de esperança. Aos meus avós e meu padrasto por todo o carinho. À minha noiva, que trouxe luz e amor aos meus dias. Ao professor Felipe Fidalgo, por me mostrar o amor pela docência e pela matemática.

AGRADECIMENTOS

O presente trabalho é também a realização de um sonho, a conclusão de uma etapa em minha vida, e não poderia deixar de agradecer especialmente a todos que me apoiaram e incentivaram: minha mãe, Rosângela Senff Wille, meu padrasto Evaldo Braga, meus avós Geni e Romeu Senff, minha noiva, Ariadne Comazzetto, e todos os amigos e familiares que contribuíram com todo tipo de suporte: livros, idas e vindas a Blumenau, cafés e conversas inspiradoras. Sou grato ao meu professor-orientador Felipe Fidalgo, que sempre me incentivou a não desistir e me mostrou o quão vasta e interessante é a matemática aplicada. À Universidade Federal de Santa Catarina, pela grande instituição de ensino que é, e que me proveu toda a educação de qualidade que tenho. Por fim, um agradecimento especial aos meus colegas, supervisores e coordenadores da Azul Linhas Aéreas do Aeroporto de Navegantes, que por diversas vezes permitiram que eu pudesse trabalhar meio período para me dedicar ainda mais aos estudos.

"Cada um, no fundo, é gênio, na medida em que existe uma vez e lança um olhar inteiramente novo sobre as coisas. Multiplica a natureza, cria por este novo olhar. Salvem seu gênio. É o que é preciso gritar para as pessoas. Liberem-no, façam o possível para libertá-lo". Friedrich Nietzsche.

RESUMO

Neste trabalho, analisamos a solução do *Molecular Distance Geometry Problem* (MDGP) com distâncias exatas usando a Decomposição de Valores Singulares (SVD, do inglês *Singular Value Decomposition*). O MDGP consiste em determinar as posições dos átomos de uma molécula, no espaço tridimensional, a partir de um conjunto de distâncias entre eles. Quando todas as distâncias são conhecidas, o problema pode ser resolvido em tempo polinomial. Caso contrário, é um problema NP-difícil.

Palavras-chave: *molecular distance geometry problem*, decomposição em valores singulares, álgebra linear.

ABSTRACT

In this work, we analyse the solution to the Molecular Distance Geometry Problem (MDGP), with exact distances, using the Singular Value Decomposition (SVD). The MDGP consists in estimating the positions of atoms in a molecule, given their pairwise distances. When all such distances are known, the problem can be solved in polynomial time. Otherwise, it is an NP-Hard problem.

Keywords: molecular distance geometry problem, singular value decomposition, linear algebra.

LISTA DE FIGURAS

Figura 1 – Molécula com sete átomos com conjunto completo de distâncias	60
--	----

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVOS	12
1.2	ROTEIRO DO TRABALHO	12
2	EXPLORANDO A ÁLGEBRA LINEAR	14
2.1	ESPAÇOS VETORIAIS	14
2.2	TRANSFORMAÇÕES LINEARES	31
2.3	ESPAÇOS DE PRODUTO INTERNO	39
2.4	DECOMPOSIÇÃO EM VALORES SINGULARES (SVD)	51
3	<i>MOLECULAR DISTANCE GEOMETRY PROBLEM (MDGP) COM SVD</i>	59
3.1	RESOLUÇÃO DO MDGP VIA DECOMPOSIÇÃO EM VALORES SINGULARES	60
4	EXPERIÊNCIA COMPUTACIONAL	66
5	CONCLUSÃO	71
	REFERÊNCIAS	72

1 INTRODUÇÃO

"We are all connected; To each other, biologically. To the Earth, chemically. To the rest of the universe, atomically".

Neil DeGrasse Tyson.

Fazer ciência, pesquisar, buscar respostas, experimentar, refutar afirmações por séculos tomadas como inquestionáveis, colocar-se a prova, aceitar a possibilidade do erro em prol da verdade – acredito que essas são características de todo cientista, ou seja, de todo homem que busca estar além de si mesmo para o benefício de toda a humanidade.

Minha história, como cientista, começou quando eu era ainda muito criança, com o desejo de que minha mãe vivesse para sempre, algo semelhante ao que Nicolau Flammel buscou na mitológica pedra filosofal. Essa foi a primeira chama que aguçou a curiosidade em meu coração, e cresceu vertiginosamente: brincando, explorei a natureza quando menino, dediquei-me a aprender as fórmulas químicas ensinadas em um jogo popular nos anos 2000, e logo as ciências e a matemática se tornaram as minhas matérias favoritas na escola. Quando cresci e precisei escolher uma faculdade, a matemática foi a minha escolha.

Foi uma grata surpresa descobrir que existe um dos muitos ramos da ciência da lógica e dos números que se dedicava ao estudo das proteínas, suas distâncias e interações. No fim, não descobri como fazer minha mãe viver para sempre, mas enveredei pelo caminho onde a matemática é fundamental para os avanços das tecnologias relacionadas as vacinas e medicamentos, ou seja, à manutenção da vida humana.

Assim, é com muito entusiasmo, que começo minha vida acadêmica através desta monografia. O DNA é um conjunto de moléculas que contém as instruções genéticas que coordenam o desenvolvimento

e o funcionamento de todos os seres vivos e de alguns vírus, e transmitem as características hereditárias entre gerações. Desde o descobrimento da sua estrutura química [12], é uma das moléculas mais estudadas do mundo. Assim, o conhecimento sobre sua estrutura e seu sequenciamento, tem fundamental importância não apenas para a filogenia biológica [4], determinando as linhagens entre espécies conhecidas, mas também para a engenharia genética e sua busca para a cura das mais diversas doenças [11].

Uma das intersecções de tais estudos com a matemática aplicada remonta à década de 1960, com o trabalho do pesquisador Christian B. Anfinsen e seus colaboradores [1, 2], que descobriram que a disposição tridimensional das proteínas está diretamente ligada ao sequenciamento dos aminoácidos: era de fundamental importância determinar a estrutura da molécula no \mathbb{R}^3 para o entendimento do sequenciamento das características biológicas dos seres terrestres. Este trabalho foi basilar para o estudo do presente trabalho.

1.1 OBJETIVOS

Nesta monografia iremos explicar sobre o Problema de Geometria de Distâncias e uma maneira de resolvê-lo, com o objetivo de aplicar os conhecimentos básicos da Álgebra Linear presentes (ou não) no curso de Licenciatura em Matemática. Iremos também aplicar este método de resolução através da decomposição em valores singulares utilizando o *software* Julia.

1.2 ROTEIRO DO TRABALHO

Iniciaremos com o estudo das bases teóricas para a decomposição em valores singulares no capítulo dois — para isso, desenvolvemos toda a teoria de álgebra linear necessária. Em seguida, no capítulo

3, enunciaremos o *Molecular Distance Geometry Problem* (MDGP), resolvendo-o posteriormente através da Decomposição em Valores Singulares (SVD). Para finalizar, no capítulo 4, abordaremos a experiência computacional para a resolução de alguns exemplos do MDGP, utilizando o *software* Julia e trataremos das conclusões no capítulo 5.

2 EXPLORANDO A ÁLGEBRA LINEAR

Esse capítulo traz uma revisão da área de Álgebra Linear, seguindo desenvolvimentos em [3, 9, 10, 13, 14]. Primeiro, definimos espaços vetoriais e suas propriedades básicas, focando no exemplo do espaço \mathbb{R}^n . Em seguida, revisamos algumas propriedades de transformações lineares e suas matrizes. Em seguida, introduzimos espaços de produto interno, que nos permitem definir transformações adjuntas. Em particular, o Teorema Espectral para operadores simétricos é demonstrado. Tudo isso culmina no desenvolvimento da decomposição em valores singulares (SVD) para transformações lineares entre espaços de dimensão finita. Alguns dos aspectos interessantes dos desenvolvimentos feitos aqui são que não são usados determinantes no desenvolvimento da teoria e ela é adequada para qualquer produto interno de \mathbb{R}^n e não apenas o usual, válida para noções distintas de ortonormalidade nesse espaço.

2.1 ESPAÇOS VETORIAIS

Um espaço vetorial é uma estrutura algébrica que nos permite somar elementos e multiplicá-los por números escalares provenientes de um corpo. Essa definição é formalizada a seguir.

Definição 1. Um *espaço vetorial* sobre um corpo \mathbb{K} é uma estrutura algébrica $(V, \mathbb{K}, +, \cdot, 0)$, com $0 \in V$, satisfazendo as relações

1. $u + (v + w) = (u + v) + w$;
2. $v + 0 = 0 + v = v$;
3. $\exists(-v) \in V, \quad v + (-v) = (-v) + v = 0$;
4. $u + v = v + u$;

$$5. \alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v;$$

$$6. (\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v;$$

$$7. 1 \cdot v = v,$$

para quaisquer $u, v, w \in V$ e $\alpha, \beta \in \mathbb{K}$, sendo $1 \in \mathbb{K}$ a unidade do corpo. É comum denotarmos $\alpha \cdot v = \alpha v$ por justaposição.

Exemplo 1. O conjunto \mathbb{R}^n é um espaço vetorial sobre o corpo \mathbb{R} dos números reais, com as operações definidas como

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n) \\ \text{e } \alpha \cdot (x_1, \dots, x_n) &= (\alpha x_1, \dots, \alpha x_n),\end{aligned}$$

sendo $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. De forma geral, denotaremos os vetores de \mathbb{R}^n como vetores-coluna:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Exemplo 2. O conjunto $M(\mathbb{K}; m, n)$ de matrizes de dimensões $m \times n$ com coeficientes em \mathbb{K} são naturalmente um espaço vetorial.

Exemplo 3. O conjunto $V = \{0\}$, com $0 + 0 = 0$ e $\alpha \cdot 0 = 0$, para qualquer $\alpha \in \mathbb{K}$, é um espaço vetorial dito *trivial*.

Espaços vetoriais possuem, em geral, uma noção de subestrutura — subconjuntos do espaço que são espaços vetoriais por si mesmos com as mesmas operações do conjunto original.

Definição 2. Se um subconjunto não-vazio $S \subset V$ de um espaço vetorial V satisfaz as propriedades

$$\begin{aligned} 0 &\in S; \\ u, v \in S &\implies u + v \in S; \\ \alpha \in \mathbb{K}, v \in S &\implies \alpha v \in S, \end{aligned}$$

S é dito um *subespaço* de V .

Exemplo 4. Seja $V = \mathbb{R}^2$. Um subespaço desse espaço vetorial é o conjunto

$$S = \{(x, y) \in \mathbb{R}^2 : y = 0\},$$

ou seja, o conjunto dos pares ordenados da forma $(x, 0)$.

Dados subespaços de um espaço vetorial, é possível construir outros subespaços. Dois exemplos, a intersecção e a soma de subespaços, são mostrados a seguir.

Proposição 1. *Seja V um espaço vetorial e S e T subespaços de V . São também subespaços:*

- $S \cap T$;
- $S + T = \{s + t : s \in S, t \in T\}$.

Demonstração. • Temos $0 \in S$ e $0 \in T$, logo $0 \in S \cap T$. Sejam $u, v \in S \cap T$ e $\alpha \in \mathbb{K}$. Então, $u \in S$ e $v \in S$, logo $u + v \in S$, pois S é subespaço. Similarmente, $u + v \in T$, logo $u + v \in S \cap T$. Da mesma forma, $u \in S$ e $u \in T$ significa que $\alpha u \in S$ e $\alpha u \in T$, logo $\alpha u \in S \cap T$.

- Como $0 = 0 + 0$ e $0 \in S, T$, segue que $0 \in S + T$. Sejam $u, v \in S + T$ e $\alpha \in \mathbb{K}$. Temos $u = u_S + u_T$, $v = v_S + v_T$, com $u_S, v_S \in S$ e $u_T, v_T \in T$, assim $u + v = (u_S + v_S) + (u_T + v_T)$ e $\alpha u =$

$\alpha u_S + \alpha u_T$ e, sendo S e T subespaços, vale que $u_S + v_S, \alpha u_S \in S$ e $u_T + v_T, \alpha u_T \in T$, logo $u + v, \alpha u \in S + T$, o que mostra que $S + T$ é subespaço

□

No caso da soma de subespaços, faz-se necessário distinguir um caso especial em que a intersecção é trivial.

Definição 3. Seja V um espaço vetorial e S e T subespaços de V . Se $S \cap T = \{0\}$, dizemos que $S + T$ é uma *soma direta* e denotamos esse subespaço por $S \oplus T$.

Exemplo 5. Seja $V = \mathbb{R}^n$. Nesse espaço, uma reta passando pela origem na direção do vetor $v \in V$ é um conjunto descrito como $R_v = \{x \in \mathbb{R}^n : x = \alpha v, \alpha \in \mathbb{R}\}$. É possível verificar que esses conjuntos são subespaços de V . Se $u, v \in V$ são tais que não existe $\alpha \in \mathbb{R}$ que resulta em $u = \alpha v$, então os conjuntos R_u e R_v descrevem retas não-colineares. Se R_u e R_v são retas não-colineares, então sua soma é uma soma direta:

$$R_u + R_v = R_u \oplus R_v.$$

Abaixo introduzimos uma noção fundamental para espaços vectoriais: a de combinações lineares.

Definição 4. Seja V um espaço vetorial e $v_1, \dots, v_n \in V$ uma coleção de vetores de V . Chamaremos a expressão

$$\alpha_1 v_1 + \dots + \alpha_n v_n$$

de uma *combinação linear* dos vetores v_1, \dots, v_n pelos números escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$.

Usando combinações lineares de um conjunto qualquer (não necessariamente um subespaço), podemos usar a noção de combinações lineares para definir o menor subespaço que contém esse conjunto.

Definição 5. Seja $G \subset V$ um subconjunto não-vazio de um espaço vetorial V . O *subespaço gerado por G* , denotado por $\langle G \rangle$, é o conjunto de todas as combinações lineares de G , isto é, dada qualquer coleção $v_1, \dots, v_n \in G$ e escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, segue que

$$\alpha_1 v_1 + \dots + \alpha_n v_n \in \langle G \rangle$$

e que, além disso, todo elemento de $\langle G \rangle$ é dessa forma. Quando $V = \langle G \rangle$, diz-se que V é gerado por G .

Resta mostrar que $\langle G \rangle$ é, de fato, um subespaço de V .

Proposição 2. Dado um subconjunto não vazio $G \subset V$ de um espaço vetorial V , segue que $\langle G \rangle$ é, de fato, um subespaço de V .

Demonstração. Como $G \neq \emptyset$, existe $u \in G$, portanto $0 = 0 \cdot u \in \langle G \rangle$. Dado que $u, v \in \langle G \rangle$, segue que $u = \alpha_1 u_1 + \dots + \alpha_n u_n$ e $v = \beta_1 v_1 + \dots + \beta_m v_m$, para $u_1, \dots, u_n, v_1, \dots, v_m \in G$ e $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in \mathbb{K}$, logo

$$u + v = \alpha_1 u_1 + \dots + \alpha_n u_n + \beta_1 v_1 + \dots + \beta_m v_m,$$

sendo assim $u + v$ é uma combinação linear de elementos de G e, portanto, $u + v \in \langle G \rangle$. Similarmente, para $\gamma \in \mathbb{K}$, temos

$$\gamma u = (\gamma \alpha_1) u_1 + \dots + (\gamma \alpha_n) u_n$$

e $\gamma u \in \langle G \rangle$. □

Observação 1. Quando G possui apenas um único elemento, ou seja, $G = \{u\}$, denotaremos $\langle G \rangle = \langle u \rangle$

Outra noção relacionada a combinações lineares é a de independência linear.

Definição 6. Um subconjunto não vazio $\mathcal{I} \subset V$ de um espaço vetorial V é dito *linearmente independente* quando, para qualquer coleção $v_1, \dots, v_n \in \mathcal{I}$ satisfazendo

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0$$

para escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, decorre necessariamente que

$$\alpha_1 = \dots = \alpha_n = 0.$$

Quando isso não é válido, o conjunto é dito *linearmente dependente*.

A próxima proposição nos dá exemplos simples de conjuntos linearmente independentes e linearmente dependentes.

Proposição 3. *Dado um espaço vetorial V e um vetor não nulo $v \in V$, $v \neq 0$, segue que $\{v\}$ é linearmente independente. Por outro lado, o conjunto $\{0\}$ é linearmente dependente.*

Demonstração. Quando $v \neq 0$, temos que, se $\alpha v = 0$, para algum $\alpha \in \mathbb{K}$. Se fosse $\alpha \neq 0$, poderíamos multiplicar a expressão por α^{-1} , resultando em

$$0 = \alpha^{-1} \cdot 0 = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1 \cdot v = v \neq 0,$$

o que é uma contradição. Agora, se $v = 0$, temos

$$1 \cdot v = 1 \cdot 0 = 0,$$

que é uma combinação linear com coeficientes não-nulos resultando em 0, mostrando que $\{v\}$ é linearmente dependente. \square

Juntando as noções de conjuntos geradores e linearmente independentes, chegamos à noção de bases de um espaço vetorial.

Definição 7. Um conjunto não vazio $B \subset V$ que é linearmente independente e que gera V — isto é, $V = \langle B \rangle$ — é chamado de uma *base* de V .

Exemplo 6. Uma base para o espaço \mathbb{R}^n , $n > 0$, é a coleção dos vetores

$$\begin{aligned}e_1 &= (1, 0, 0, \dots, 0), \\e_2 &= (0, 1, 0, \dots, 0), \\&\vdots \\e_n &= (0, 0, \dots, 0, 1),\end{aligned}$$

que é chamada de *base canônica* desse espaço.

A seguir, mostramos um teorema bastante geral para espaços vetoriais. Ele nos diz que, em certas condições, podemos completar conjuntos linearmente independentes para uma base e reduzir conjuntos geradores também para uma base. Esse teorema faz uso de um teorema bastante geral — inclusive equivalente ao axioma da escolha nas teorias de conjuntos padrões — chamado de Lema de Zorn.

Lema 1 (de Zorn). *Seja (P, \leq) um conjunto parcialmente ordenado. Uma cadeia em P é um subconjunto $C \subset P$ com a propriedade de que, para quaisquer $a, b \in C$, temos $a \leq b$ ou $b \leq a$. Um máximo para um subconjunto $X \subset P$ em P é um elemento $m \in P$ tal que $a \leq m$ para todo $a \in X$. Similarmente, um elemento mínimo em P é um elemento $m \in P$ com a propriedade de que, para qualquer $a \in P$ com $m \leq a$, temos $a = m$, ou seja, é um elemento que não é estritamente menor do que qualquer outro em P .*

O Lema de Zorn diz que, se toda cadeia em P possui um máximo em P , então P possui um elemento maximal.

Demonstração. Uma referência para essa prova é [15], onde são mostradas, sob a teoria de Zermelo-Fraenkel, as equivalências entre o axioma da escolha, o teorema da boa ordenação e o lema de Zorn. \square

Agora provamos o seguinte resultado.

Teorema 1 (Completamento). *Seja V um espaço vetorial e \mathcal{I}, G subconjuntos não vazios de V com $\mathcal{I} \subset G \subset V$ tais que \mathcal{I} é linearmente independente e G gera V — isto é, $V = \langle G \rangle$. Então existe uma base $B \subset V$ de V tal que $\mathcal{I} \subset B \subset G$.*

Demonstração. Vamos construir o conjunto \mathcal{J} , parcialmente ordenado por inclusão, dos conjuntos linearmente independentes contidos em G que contém \mathcal{I} . Vemos primeiro que $\mathcal{J} \neq \emptyset$, pois $\mathcal{I} \in \mathcal{J}$. Vamos tomar uma cadeia $\mathcal{C} \subset \mathcal{J}$ não vazia em \mathcal{J} . Seja então

$$C = \bigcup \mathcal{C}.$$

Vamos mostrar que C é um conjunto linearmente independente. Suponha que não seja. Então existem $u_1, \dots, u_n \in C$ e $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ com, digamos, $\alpha_1 \neq 0$, tais que

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0.$$

Existem $C_1, \dots, C_n \in \mathcal{C}$ tais que $u_i \in C_i$ para cada $i \in \mathbb{N}$, $1 \leq i \leq n$. Como \mathcal{C} é uma cadeia, segue que

$$C_i = C_1 \cup \dots \cup C_n$$

para algum $i \in \mathbb{N}$, $1 \leq i \leq n$. Mas isso implica que $u_1, \dots, u_n \in C_i$, o que é uma contradição, pois $C_i \in \mathcal{C}$ e C_i deve ser linearmente independente. Segue que C é então um conjunto linearmente independente, logo $C \in \mathcal{J}$, além de ser um máximo para a cadeia \mathcal{C} em \mathcal{J} . Logo podemos aplicar o Lema de Zorn e construir um elemento maximal $B \in \mathcal{J}$.

Como $B \in \mathcal{J}$, segue que B é linearmente independente. Suponha agora que $\langle B \rangle \neq V$, ou seja $V \setminus \langle B \rangle \neq \emptyset$. Como $\langle G \rangle = V$, existe $v \in G$ tal que $v \notin \langle B \rangle$. O conjunto $B \cup \{v\}$ é então linearmente independente, o que contradiz a maximalidade de B . Logo deve ser $\langle B \rangle = V$ e então B é uma base para V . \square

Seu principal corolário é o seguinte.

Corolário 1. *Todo espaço vetorial não trivial V — isto é, $V \neq \{0\}$ — possui uma base.*

Demonstração. Como $V \neq \{0\}$, segue que existe $v \in V$ tal que $v \neq 0$. Da Proposição 3, segue que $\{v\}$ é um conjunto linearmente independente. Além disso, V é um conjunto tal que $\langle V \rangle = V$, logo temos $\{v\} \subset V$ tal que $\{v\}$ é linearmente independente e V gerando V , logo, pelo Teorema 1, existe uma base B de V tal que $\{v\} \subset B \subset V$. \square

Para finalizar a caracterização de bases de espaços vetoriais, pelo menos para espaços finitamente gerados, podemos provar os seguintes resultados. Denote a cardinalidade de um conjunto X por $|X|$.

Teorema 2. *Seja V um espaço vetorial, $G \subset V$ um conjunto finito tal que $V = \langle G \rangle$ e $\mathcal{I} \subset V$ um conjunto linearmente independente e também finito. Segue que $|\mathcal{I}| \leq |G|$.*

Demonstração. Suponha \mathcal{I}, G , como nas hipóteses, com $|\mathcal{I}| = n \in \mathbb{N}$ e $|G| = m \in \mathbb{N}$. Queremos mostrar que $n \leq m$. Vamos escrever os conjuntos como $\mathcal{I} = \{x_1, \dots, x_n\}$ e $G = \{y_1, \dots, y_m\}$. Então olhamos para a coleção $\{x_1\} \cup G = \{x_1, y_1, \dots, y_m\}$. Essa coleção é linearmente dependente, pois G gera V , logo podemos escrever x_1 como uma combinação linear de y_1, \dots, y_m , então dados escalares $\alpha_1, \beta_1, \dots, \beta_m \in \mathbb{K}$

com

$$\alpha_1 x_1 + \beta_1 y_1 + \cdots + \beta_n y_n = 0,$$

segue que pelo menos um dos escalares é não nulo. Se fosse necessariamente $\beta_1 = \cdots = \beta_n = 0$, então teríamos $\alpha_1 x_1 = 0$, resultando em $\alpha_1 = 0$, o que levaria à coleção ser linearmente independente, o que é uma contradição. Então podemos escolher $\beta_{i_1} \neq 0$, para algum $i_1 \in \mathbb{N}$, $1 \leq i_1 \leq m$. Isso significa que, dividindo a expressão por β_{i_1} , podemos escrever y_{i_1} como uma combinação linear dos outros vetores da coleção $\{x_1\} \cup G$. Isso significa que, ao remover y_{i_1} da coleção, resultando em $\{x_1\} \cup (G \setminus \{y_{i_1}\}) = \{x_1, y_1, \dots, y_{i_1-1}, y_{i_1+1}, \dots, y_m\}$, ela ainda gerará V , pois, além de ter cada um de seus elementos, ela é capaz de gerar y_{i_1} , garantindo que seja possível gerar todos os elementos de V .

Agora adicionamos o elemento x_2 , obtendo

$$\{x_1, x_2\} \cup (G \setminus \{y_{i_1}\}) = \{x_1, x_2, y_1, \dots, y_{i_1-1}, y_{i_1+1}, \dots, y_m\},$$

que é linearmente dependente, porque a coleção anterior $\{x_1\} \cup (G \setminus \{y_{i_1}\})$ gerava V . Similarmente, podemos escolher um y_{i_2} , $i_2 \in \mathbb{N}$, $1 \leq i_2 \leq m$, $i_2 \neq i_1$, que é combinação linear dos outros, pois, caso contrário, isso significaria que $\{x_1, x_2\}$ é uma coleção linearmente dependente, uma contradição. Assim mostramos que a coleção

$$\{x_1, x_2\} \cup (G \setminus \{y_{i_1}, y_{i_2}\})$$

gera V .

Repetindo esse processo k vezes, sem exaurir o conjunto \mathcal{I} , conseguimos obter coleções

$$\{x_1, \dots, x_k\} \cup (G \setminus \{y_{i_1}, \dots, y_{i_k}\}),$$

todas elas geradoras de V , com a propriedade de, ao adicionar x_{k+1} , se houver ainda, a coleção resultante é linearmente dependente. Por

absurdo, se fosse $n > m$, poderíamos, na etapa $k = m$, remover todos os elementos de G , obtendo a sequência

$$\{x_1, \dots, x_m\} \cup (G \setminus \{y_{i_1}, \dots, y_{i_m}\}) = \{x_1, \dots, x_m\},$$

com a propriedade de que $\{x_1, \dots, x_m, x_{m+1}\}$ é linearmente dependente (que é possível de construir, pois $n \geq m + 1$). Mas isso é uma contradição, logo segue que $n \leq m$. \square

Corolário 2. *Sejam B_1 e B_2 duas bases de um espaço vetorial não-trivial V . Se B_1 é um conjunto finito, então B_2 também é. Além disso, segue que as cardinalidades de B_1 e B_2 são iguais.*

Demonstração. Suponha que B_2 é infinito. Esse conjunto então possui subconjuntos finitos, todos eles linearmente independentes, de todas as cardinalidades. Mas isso é impossível, pois, sendo $|B_1| = n$, teríamos que B_2 teria um subconjunto linearmente independente de cardinalidade $n + 1$. Mas, como B_1 gera V , o teorema anterior mostra que isso é impossível. Logo B_2 deve ser finito.

Como B_1 e B_2 são finitos, com, em particular, B_1 linearmente independente e B_2 gerando V , segue, também do teorema anterior, que $|B_1| \leq |B_2|$. Similarmente, trocando os papéis de B_1 e B_2 , segue que $|B_2| \leq |B_1|$, ou seja, $|B_1| = |B_2|$. \square

Com isso, podemos definir a noção de dimensão de um espaço vetorial

Definição 8. *Seja V um espaço vetorial. Se $V = \{0\}$, definimos a *dimensão* de V como 0. Se V é não-trivial, definimos sua dimensão como a cardinalidade de uma de suas bases. Em qualquer caso, denotamos a dimensão de V como $\dim V$.*

Definição 9. *Um espaço vetorial V é dito de *dimensão finita* quando $\dim V$ é uma cardinalidade finita.*

Usando bases, podemos definir a noção de coordenadas de vetores.

Proposição 4. *Seja V um espaço de dimensão finita $n > 0$ e $B = \{e_1, \dots, e_n\}$ uma base desse espaço. Para todo vetor $v \in V$, existe uma coleção única de escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tais que*

$$v = \alpha_1 e_1 + \dots + \alpha_n e_n,$$

a menos da ordem dos vetores $e_i \in B$.

Demonstração. Do fato de B gerar V , existe uma coleção de escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tais que

$$v = \alpha_1 e_1 + \dots + \alpha_n e_n.$$

Resta mostrar a unicidade deles. Sejam $\beta_1, \dots, \beta_n \in \mathbb{K}$ tais que

$$v = \beta_1 e_1 + \dots + \beta_n e_n.$$

Igualando as duas expressões de v , temos

$$\begin{aligned} \alpha_1 e_1 + \dots + \alpha_n e_n &= \beta_1 e_1 + \dots + \beta_n e_n \\ (\alpha - \beta_1) e_1 + \dots + (\alpha_n - \beta_n) e_n &= 0, \end{aligned}$$

o que, do fato de B ser um conjunto linearmente independente, implica

$$\alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0,$$

ou seja, $\alpha_i = \beta_i$, para $i = 1, \dots, n$, mostrando sua unicidade. \square

Definição 10. Nas condições da Proposição 4, o vetor-coluna denotado por

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

é chamado do *vetor de coordenadas de v na base B* .

Dadas duas, bases, é possível fazer mudança de coordenadas de um vetor entre elas.

Proposição 5 (Mudança de base). *Seja V um espaço vetorial de dimensão finita $n > 0$, com bases $B_1 = \{e_1, \dots, e_n\}$ e $B_2 = \{f_1, \dots, f_n\}$. Dado que, para cada $j = 1, \dots, n$, temos que as coordenadas dos vetores da base B_2 na base B_1 são*

$$[f_j]_{B_1} = \begin{bmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{bmatrix},$$

podemos definir uma matriz $P_{B_1 \text{ to } B_2}$, chamada da matriz de mudança de base de B_1 para B_2 , de dimensões $n \times n$ como

$$P_{B_1 \rightarrow B_2} = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \dots & \alpha_{nn} \end{bmatrix}$$

tal que $P_{B_1 \rightarrow B_2}$ é invertível e, para todo $v \in V$, temos que

$$[v]_{B_2} = P_{B_1 \rightarrow B_2}^{-1} [v]_{B_1}.$$

Demonstração. Seja $v \in V$ e escrevemos v na base B_1 como

$$v = \sum_{i=1}^n \gamma_i e_i$$

e na base B_2 como

$$v = \sum_{j=1}^n \lambda_j f_j.$$

Sabemos que cada elemento da base B_2 pode ser escrito como

$$f_j = \sum_{i=1}^n \alpha_{ij} e_i,$$

logo, na expressão de v na base B_2 , temos

$$v = \sum_{j=1}^n \lambda_j f_j = \sum_{j=1}^n \lambda_j \sum_{i=1}^n (\alpha_{ij} e_i) = \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} \lambda_j \right) e_i,$$

obtendo uma expressão na base B_1 . Da unicidade de coordenadas, obtemos

$$\gamma_i = \sum_{j=1}^n \alpha_{ij} \lambda_j,$$

ou, matricialmente,

$$[v]_{B_1} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = P_{B_1 \rightarrow B_2} [v]_{B_2}.$$

Resta agora mostrar que a matriz $P_{B_1 \rightarrow B_2}$ é inversível.

Escrevendo agora os vetores de B_1 na base B_2 como

$$[e_j]_{B_2} = \begin{bmatrix} \beta_{1j} \\ \vdots \\ \beta_{nj} \end{bmatrix},$$

obtemos uma matriz de mudança de base

$$P_{B_2 \rightarrow B_1} = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{bmatrix}.$$

Repetindo o processo anterior, obtemos a expressão

$$[v]_{B_2} = P_{B_2 \rightarrow B_1} [v]_{B_1}$$

para qualquer $v \in V$, que, substituindo na expressão anterior, leva a

$$[v]_{B_1} = P_{B_1 \rightarrow B_2} P_{B_2 \rightarrow B_1} [v]_{B_1}$$

e, substituindo as equações na ordem contrária,

$$[v]_{B_2} = P_{B_2 \rightarrow B_1} P_{B_1 \rightarrow B_2} [v]_{B_2}.$$

Como isso é válido para qualquer $v \in V$, podemos usar os vetores das bases B_1 e B_2 , respectivamente na primeira e na segunda equação, o que resulta nas equações matriciais

$$P_{B_1 \rightarrow B_2} P_{B_2 \rightarrow B_1} = I_n \quad \text{e} \quad P_{B_2 \rightarrow B_1} P_{B_1 \rightarrow B_2} = I_n,$$

em que I_n é a matriz identidade de dimensão $n \times n$, logo segue que $P_{B_1 \rightarrow B_2}$ é inversível, com

$$P_{B_1 \rightarrow B_2}^{-1} = P_{B_2 \rightarrow B_1}.$$

□

Para uma última sequência de resultados para essa seção, mostramos algumas propriedades da dimensão da soma de subespaços.

Teorema 3. *Sejam U e W subespaços de dimensão finita de um espaço vetorial V . Tem-se*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Demonstração. Suponha $U \cap W = \{0\}$ e sejam $E = \{e_1, \dots, e_n\}$ e $F = \{f_1, \dots, f_m\}$ bases para U e W , respectivamente. Temos $E \cap F = \emptyset$, pois, se $v \in E \subset U$ e $v \in F \subset W$, então $v \in U \cap W$ e $v = 0$, o que não pode ser, pois E e F são linearmente independentes. Naturalmente $E \cup F$ gera $U + W$. Falta provar que $E \cup F$ é linearmente independente. Sejam $\alpha_i, \beta_j \in \mathbb{K}$ tais que $\sum_{i=1}^n \alpha_i e_i + \sum_{j=1}^m \beta_j f_j = 0$. Podemos reorganizar essa equação como

$$\sum_{i=1}^n \alpha_i e_i = - \sum_{j=1}^m \beta_j f_j,$$

e o lado esquerdo está em U , enquanto o lado direito pertence a W . Mas eles são iguais, logo pertencem a $U \cap W = \{0\}$, ou seja,

$$\sum_{i=1}^n \alpha_i e_i = 0 = - \sum_{j=1}^m \beta_j f_j$$

e de E e F serem linearmente independentes, segue que

$$\alpha_1 = \dots = \alpha_n = 0 = \beta_1 = \dots = \beta_m,$$

logo $E \cup F$ é linearmente independente e, portanto, uma base. Fazendo a contagem, temos $\dim(U+W) = |E \cup F| = n+m = \dim(U) + \dim(W)$ e o resultado segue.

Seja agora $U \cap W \neq \{0\}$. Se for $U \cap W = U$, então $U \subset W$ e $U + W = W$, logo $\dim(U+W) = \dim(W) = \dim(W) + \dim(U) - \dim(U) = \dim(W) + \dim(U) - \dim(U \cap W)$ e o resultado segue. Analogamente para $U \cap W = W$. Suponha então que $U \cap W \subsetneq U$ e $U \cap W \subsetneq W$ e tome uma base G de $U \cap W$. Como $U \cap W \subset U$, segue que $U \cap W$ tem dimensão finita e podemos escrever $G = \{g_1, \dots, g_k\}$, com $\dim(U \cap W) = k$, sendo $0 < k < \dim U$ e $0 < k < \dim W$. Seja $\dim U = n$ e $\dim W = m$. Podemos completar G para uma base de U como $E = G \cup E'$, com $E' = \{e_{k+1}, e_{k+2}, \dots, e_n\}$ e da mesma forma para uma base F de W com $F = G \cup F'$ com $F' = \{f_{k+1}, f_{k+2}, \dots, f_m\}$. Temos primeiro que $G \cap E' = \emptyset = G \cap F'$, pois caso contrário $\dim(U) = |G \cup E| < n = \dim(U)$ – uma contradição – e analogamente para W . Além disso, $E' \cap F' = \emptyset$, pois, caso $v \in E' \cap F'$, teríamos $v \in U \cap W$ e v seria um elemento, digamos, de E' gerado por G e $E = G \cup E'$ não seria linearmente independente. Vamos agora tomar o conjunto $H = G \cup E' \cup F'$, com cada um dos componentes dois a dois, disjuntos. Temos também que $H = E \cup F$, portanto, H gera $U + W$. Resta mostrar que H é linearmente independente. Sejam

$\alpha_i, \beta_j, \gamma_r \in \mathbb{K}$ tais que

$$\sum_{r=1}^k \gamma_r g_r + \sum_{i=k+1}^n \alpha_i e_i + \sum_{i=j=k+1}^m \beta_j f_j = 0.$$

Essa equação implica, por exemplo,

$$\sum_{i=k+1}^n \alpha_i e_i = - \sum_{r=1}^k \gamma_r g_r - \sum_{i=j=k+1}^m \beta_j f_j = 0.$$

portanto $\sum_{i=k+1}^n \alpha_i e_i \in W$. Mas também vale que $\sum_{i=k+1}^n \alpha_i e_i \in U$, logo $\sum_{i=k+1}^n \alpha_i e_i \in U \cap W$ e $\sum_{i=k+1}^n \alpha_i e_i$ deve ser gerado por G , ou seja, esse vetor pode ser escrito como uma combinação linear de elementos de G . Se algum dos α_i for não nulo, poderíamos escrever e_i como uma combinação linear dos outros elementos de E' juntos com os elementos de G , o que seria uma contradição, pois $E' \cup G = E$ é linearmente independente. Logo $\alpha_{k+1} = \dots = \alpha_n = 0$. Analogamente, $\beta_{k+1} = \dots = \beta_m = 0$. Disso, chegamos na igualdade

$$\sum_{r=1}^k \gamma_r g_r = 0,$$

que, sendo G linearmente independente, mostra que $\gamma_1 = \dots = \gamma_k = 0$, ou seja, H é um conjunto linearmente independente, e, portanto, uma base de $U + W$. Contando a quantidade de elementos, temos

$$\begin{aligned} \dim(U + W) &= |H| = |G| + |E'| + |F'| = k + (n - k) + (m - k) \\ &= n + m - k = \dim U + \dim W - \dim(U \cap W). \end{aligned}$$

□

Corolário 3. *Seja V um espaço de dimensão finita e $S \subset V$ e $T \subset V$ subespaços de V tais que $S \cap T = \{0\}$. Então*

$$\dim(S \oplus T) = \dim S + \dim T.$$

2.2 TRANSFORMAÇÕES LINEARES

Agora que sabemos de algumas propriedades básicas de espaços vetoriais, passaremos a estudar transformações entre esses espaços que preservam a estrutura deles.

Definição 11. Dados espaços vetoriais U e V sobre \mathbb{K} , uma função $T: U \rightarrow V$ é chamada de uma *transformação linear* se, para qualquer $u, v \in U$ e $\alpha \in \mathbb{K}$ temos

$$T(u + v) = T(u) + T(v)$$
$$\text{e } T(\alpha u) = \alpha T(u).$$

Decorre da definição acima que $T(0_U) = 0_V$ para qualquer transformação linear $T: U \rightarrow V$.

Exemplo 7. Dado um vetor coluna $x \in \mathbb{R}^n$ e uma matriz $A \in \mathbb{R}^{m \times n}$, a aplicação $x \mapsto Ax$ define uma transformação linear $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Afinal, temos que $A(\alpha x) = \alpha(Ax)$ e $A(x + y) = Ax + Ay$, para quaisquer $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$.

Definição 12. Uma transformação linear $T: V \rightarrow V$ entre um mesmo espaço vetorial V é também chamada de um *operador linear*.

A seguir, definiremos algumas classes especiais de transformações lineares.

Definição 13. Sejam U e V espaços vetoriais e $T: U \rightarrow V$ uma transformação linear.

- T é dita *sobrejetiva* quando $T(U) = V$;
- T é dita *injetiva* quando $T(x) = T(y) \implies x = y$, para quaisquer $x, y \in U$;

- T é dita *bijetiva* quando é injetiva e sobrejetiva.

Quando uma transformação linear T é bijetiva, T é também chamada de um *isomorfismo*.

Isomorfismos, em particular, são importantes, porque definem a seguinte relação entre espaços vetoriais.

Definição 14. Sejam U e V espaços vetoriais. Quando existe um isomorfismo $T: U \rightarrow V$, os espaços U e V são ditos *isomorfos* e denota-se $U \cong V$.

Espaços isomorfos são espaços equivalentes na estrutura de suas operações. Uma das consequências dessa equivalência está na Proposição 9. Mas, antes, vamos definir algumas outras estruturas relacionadas a transformações lineares;

Definição 15. Sejam U e V espaços vetoriais e $T: U \rightarrow V$ uma transformação linear. Definimos a *imagem* de T como o subespaço

$$\text{Im}(T) = \{v \in V : v = T(u), \text{ para algum } u \in U\} = T(U)$$

e o *espaço nulo*, ou *núcleo*, de T como o subespaço

$$N(T) = \{u \in U : T(u) = 0\}.$$

Proposição 6. Nas condições acima, $\text{Im}(T)$ e $N(T)$ são subespaços de V e U , respectivamente.

Demonstração. É imediato que $0 \in N(T)$. Seja $u_1, u_2 \in N(T)$ e $\alpha \in \mathbb{K}$. Disso temos que $T(u_1) = 0$ e $T(u_2) = 0$, logo $T(u_1 + u_2) = T(u_1) + T(u_2) = 0$ e $T(\alpha u_1) = \alpha T(u_1) = 0$, logo $u_1 + u_2 \in N(T)$ e $\alpha u_1 \in N(T)$, o que implica que $N(T)$ é um subespaço de U .

Para a imagem, note que $T(0) = 0$, logo $0 \in \text{Im}(T)$. Seja $v_1, v_2 \in \text{Im}(T)$ e $\alpha \in \mathbb{K}$. Existe $u_1, u_2 \in U$ tais que $v_1 = T(u_1)$ e $v_2 = T(u_2)$.

Logo $T(u_1 + u_2) = T(u_1) + T(u_2) = v_1 + v_2$ e $T(\alpha u_1) = \alpha T(u_1) = \alpha v_1$, logo $v_1 + v_2 \in \text{Im}(T)$ e $\alpha v_1 \in \text{Im}(T)$, mostrando que $\text{Im}(T)$ é um subespaço de V . \square

Esses espaços podem nos dizer quando uma transformação é sobrejetiva ou injetiva, conforme a próxima proposição.

Proposição 7. *Sejam U e V espaços vetoriais e $T: U \rightarrow V$ uma transformação linear.*

1. T é sobrejetiva se, e somente se, $\text{Im}(T) = V$.
2. T é injetiva se, e somente se, $N(T) = \{0\}$.

Demonstração. A primeira afirmação é trivial. Para a segunda, suponha T injetiva e seja $u \in U$ tal que $T(u) = 0$. Como $T(0) = 0$, segue que $u = 0$, logo $N(T) = \{0\}$. Agora, suponha que $N(T) = \{0\}$ e sejam $u, v \in U$ tais que $T(u) = T(v)$. Disso segue que $T(u - v) = 0$ e o espaço nulo trivial garante que $u - v = 0$, isto é, $u = v$, ou seja, T é injetiva. \square

Outra propriedade de transformações injetivas é que o domínio é isomorfo à sua imagem.

Proposição 8. *Sejam U e V espaços vetoriais e $T: U \rightarrow V$ uma transformação injetiva. Então segue que $\text{Im}(T) \cong U$.*

Demonstração. A restrição de contradomínio de T , dada por

$$T|_{\text{Im}(T)}: U \rightarrow \text{Im}(T), \quad u \mapsto T(u)$$

é trivialmente uma transformação linear sobrejetiva. Como T é injetiva, $T|_{\text{Im}(T)}$ também será, logo $T|_{\text{Im}(T)}$ é um isomorfismo entre U e $\text{Im}(T)$. \square

Agora, mostraremos uma importante consequência da existência de isomorfismos entre espaços.

Proposição 9. *Sejam U e V espaços vetoriais. Temos que $U \cong V$ se, e somente se, $\dim U = \dim V$.*

Demonstração. Quando $U \cong V$, existe um isomorfismo $T: U \rightarrow V$. Se um deles é um espaço trivial, o outro também deve ser, pois T é injetiva, logo eles possuem a mesma dimensão nesse caso. Se U é não-trivial, ele possui uma base B . Note que a aplicação T define uma bijeção entre os conjuntos B e $T(B)$, logo $|B| = |T(B)|$. Afirmamos que $T(B)$ é uma base para V . Como ela é injetiva, qualquer combinação linear de elementos de $T(B)$ resultando no vetor nulo resulta em uma combinação linear de elementos de B resultando no vetor nulo, que implica coeficientes nulos, ou seja, $T(B)$ é linearmente independente. Como T é sobrejetiva, para qualquer $v \in V$ existe $u \in U$ tal que $v = T(u)$. Expressando u na base B resulta, por essa relação, em uma expressão de v por elementos de $T(B)$, mostrando que $T(B)$ gera V .

Por outro lado, se a dimensão dos espaços é nula, eles são claramente isomorfos. Se a dimensão não é nula, então tomamos bases B_U e B_V de U e V , respectivamente. Temos por hipótese $|B_U| = |B_V|$, logo existe uma bijeção $T: B_U \rightarrow B_V$. Podemos estender essa função para uma transformação linear em $U \rightarrow V$. A unicidade dos coeficientes da expansão em B_U garante que essa transformação está bem definida e que ela é única. O fato de B_V ser uma base implica, agora, que T é injetivo (por B_V ser linearmente independente) e sobrejetivo (por B_V gerar V), logo essa transformação é um isomorfismo. \square

Para o próximo importante teorema, o Teorema do Posto e Nulidade, que também é conhecido como Teorema do Núcleo e da Imagem, vamos definir as seguintes quantidades.

Definição 16. Sejam U e V espaços vetoriais e $T: U \rightarrow V$ uma transformação linear. Definimos o *posto* de T como a dimensão de $\text{Im}(T)$, denotado por

$$\text{posto}(T) = \dim \text{Im}(T),$$

e definimos a *nulidade* de T como a dimensão de $N(T)$, denotado

$$\text{nul}(T) = \dim N(T).$$

Teorema 4 (Posto-Nulidade). *Sejam U e V espaços vetoriais de dimensão finita e $T: U \rightarrow V$ uma transformação linear. Então*

$$\dim(U) = \text{nul}(T) + \text{posto}(T).$$

Demonstração. Suponha $\dim(U) = n$. Se $T = 0$, então $\text{Im}(T) = \{0\}$ e $N(T) = U$, portanto, a equação é válida. Seja então $T \neq 0$. Se $N(T) = \{0\}$, a transformação é injetiva e $\text{Im}(T) \cong U$, logo a equação é válida. Suponha agora que existe $u \in U$, $u \neq 0$, tal que $T(u) = 0$. Isso significa que $N(T) \neq \{0\}$ e existe uma base E' de $N(T)$. Como $\{0\} \subsetneq N(T) \subsetneq U$, $N(T)$ é um subespaço próprio de U , com $0 < \dim(N(T)) < n$. Seja então $\dim(N(T)) = n - r$, para $r \geq 1$. Podemos ordenar a base de $N(T)$ como $E' = \{e_{r+1}, e_{r+2}, \dots, e_{n-1}, e_n\}$. Vamos completar essa base para uma base $E \supset E'$ de U , escrevendo $E = \{e_1, \dots, e_r\} \cup \{e_{r+1}, \dots, e_n\}$. Vemos que, com essa ordenação, de fato E possui n elementos e E' possui $n - r$ elementos, com $E' \subset E$. Agora, Seja $v \in \text{Im}(T)$. Existe $u \in U$ tal que $T(u) = v$. Escrevendo u na base E , temos $u = \sum_{i=1}^n \alpha_i e_i$, para coeficientes únicos $\alpha_i \in \mathbb{K}$. Podemos pegar outro vetor $u' = \sum_{i=1}^r \alpha_i e_i$ e teremos também $v = T(u) = T(u')$, pois os elementos da base E após e_r vão para zero, com $v = \sum_{i=1}^r \alpha_i T(e_i)$. Isso vale para qualquer $v \in \text{Im}(T)$, logo o conjunto $\{T(e_1), \dots, T(e_r)\}$ gera $\text{Im}(T)$. Esse conjunto também é linearmente independente, caso contrário, teríamos uma combinação linear com coeficientes $\beta_i \in \mathbb{K}$ não

todos nulos com $0 = \beta_1 T(e_1) + \cdots + \beta_r T(e_r) = T(\beta_1 e_1 + \cdots + \beta_r e_r)$, e $\sum_{i=1}^r \beta_i e_i \in N(T)$, portanto, sendo uma combinação linear não trivial de E' , o que não pode ser, pois, o conjunto E é linearmente independente, logo segue que $\{T(e_1), \dots, T(e_r)\}$ é uma base de $\text{Im}(T)$ e, fazendo as contas finais, temos $\dim(N(U)) + \dim(\text{Im}(T)) = n - r + r = n = \dim(U)$. \square

Assim como vetores se expressam como coordenadas em um vetor-coluna dada uma base, transformações lineares se expressam como matrizes dadas bases de U e V .

Proposição 10. *Sejam U e V espaços vetoriais de dimensão finita, com $\dim U = n > 0$ e $\dim V = m > 0$, e $T: U \rightarrow V$ uma transformação linear. Se B_U é uma base de U e B_V é uma base de V , existe uma matriz $[T]_{B_U B_V}$ de tamanho $m \times n$ tal que*

$$[T(v)]_{B_V} = [T]_{B_U B_V} [v]_{B_U}$$

para todo $v \in U$.

Demonstração. Sejam $B_U = \{e_1, \dots, e_n\}$ e $B_V = \{f_1, \dots, f_m\}$. Vamos expandir cada $T(e_j)$ na base B_V como

$$T(e_j) = \alpha_{1j} f_1 + \cdots + \alpha_{mj} f_m.$$

Seja $v \in U$ e escreva

$$v = \beta_1 e_1 + \cdots + \beta_n e_n.$$

Temos que

$$T(v) = \sum_{j=1}^n \beta_j T(e_j) = \sum_{j=1}^n \beta_j \sum_{i=1}^m \alpha_{ij} f_i = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_{ij} \beta_j \right) f_i,$$

o que, em termos de coordenadas, fica

$$[T(v)]_{B_V} = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix} [v]_{B_U},$$

ou seja, temos

$$[T]_{B_U B_V} = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix}$$

e essa matriz não depende de v , apenas dos valores de $T(e_j)$. \square

Observação 2. Se $T: V \rightarrow V$ é um operador linear e B é uma base de V , denotamos $[T]_{BB}$ como apenas $[T]_B$.

É possível também fazer mudança de bases para a matriz de uma transformação linear.

Proposição 11. *Sejam U e V espaços vetoriais de dimensão finita, com $\dim U = n > 0$ e $\dim V = m > 0$, e $T: U \rightarrow V$ uma transformação linear. Sejam também B_U e B'_U bases de U e B_V e B'_V bases de V . Sendo $P_{B_U \rightarrow B'_U}$ a matriz de mudança de base de B_U para B'_U e $P_{B_V \rightarrow B'_V}$ a matriz de mudança de base de B_V para B'_V , segue que*

$$[T]_{B'_U B'_V} = P_{B_V \rightarrow B'_V}^{-1} [T]_{B_U B_V} P_{B_U \rightarrow B'_U}.$$

Demonstração. Seja $u \in U$ e escrevemos $[u]_{B'_U} = P_{B_U \rightarrow B'_U}^{-1} [u]_{B_U}$, ou seja, $[u]_{B_U} = P_{B_U \rightarrow B'_U} [u]_{B'_U}$. Da mesma forma, temos $[T(u)]_{B_V} = P_{B_V \rightarrow B'_V} [T(u)]_{B'_V}$. A matriz de T nas bases B_U e B_V expressa que

$$[T(u)]_{B_V} = [T]_{B_U B_V} [u]_{B_U}$$

e, substituindo as expressões acima obtemos

$$P_{B_V \rightarrow B'_V} [T(u)]_{B'_V} = [T]_{B_U B_V} P_{B_U \rightarrow B'_U} [u]_{B'_U},$$

ou,

$$[T(u)]_{B'_V} = P_{B_V \rightarrow B'_V}^{-1} [T]_{B_U B_V} P_{B_U \rightarrow B'_U} [u]_{B'_U}.$$

Como isso vale para u arbitrário, segue que

$$[T]_{B'_U B'_V} = P_{B_V \rightarrow B'_V}^{-1} [T]_{B_U B_V} P_{B_U \rightarrow B'_U}.$$

□

Observação 3. Se $T: V \rightarrow V$ é um operador linear e B e B' são bases de V , com $P = P_{B \rightarrow B'}$ sendo a matriz de mudança de base de B para B' , temos que $[T]_{B'} = P^{-1}[T]_B P$.

Agora vamos começar a desenvolver a teoria de autovalores e autovetores de um operador linear.

Definição 17. Seja $T: V \rightarrow V$ um operador linear. Um escalar $\lambda \in \mathbb{K}$ é chamado um *autovalor* de T quando existe $p \in V$, com $p \neq 0$, chamado seu *autovetor associado*, quando

$$T(p) = \lambda p.$$

Definição 18. Seja $T: V \rightarrow V$ um operador linear e V . T é dito *diagonalizável* quando existe uma base de V formada por autovetores de T .

A justificativa para essa nomenclatura está na seguinte proposição.

Proposição 12. *Seja $T: V \rightarrow V$ um operador linear diagonalizável e V um espaço de dimensão finita. A representação matricial de T em uma base de seus autovetores é uma matriz diagonal.*

Demonstração. Seja $B = \{p_1, \dots, p_n\}$ uma base de autovetores de T . A matriz de T na base T é a matriz cujas colunas são os valores de

T na base B , com os coeficientes escritos também na base B . Sejam $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ os autovalores associados aos autovetores p_1, \dots, p_n , respectivamente. Temos

$$T(p_i) = \lambda_i p_i,$$

ou seja, $T(p_i)$, escrito na base B , é o vetor de coordenadas todas iguais a zero exceto na posição i , cujo valor é λ_i . Isso mostra que

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

que é uma matriz diagonal. □

Observação 4. Uma consequência dessa proposição é a seguinte discussão. Se uma matriz $A \in \mathbb{R}^{n \times n}$ é considerada a representação nas bases canônicas de uma transformação linear $\mathbb{R}^n \rightarrow \mathbb{R}^n$, essa transformação (e por consequência a matriz A) é diagonalizável quando existe uma matriz inversível $P \in \mathbb{R}^{n \times n}$ (a matriz de mudança de base) e uma matriz diagonal $\Lambda \in \mathbb{R}^{n \times n}$ tal que $\Lambda = P^{-1}AP$.

2.3 ESPAÇOS DE PRODUTO INTERNO

Nesta parte, introduziremos uma estrutura que permite medirem-se comprimentos e ângulos relacionados a vetores. Além disso, estudaremos transformações lineares relacionadas a essa estrutura.

Definição 19. Seja V um espaço vetorial sobre \mathbb{R} . Um produto in-

terno em V é uma forma bilinear $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ que satisfaz

$$\begin{aligned}\langle u + v, w \rangle &= \langle u, w \rangle + \langle v, w \rangle; \\ \langle \alpha u, v \rangle &= \alpha \langle u, v \rangle; \\ \langle u, v \rangle &= \langle v, u \rangle; \\ \langle u, u \rangle &\geq 0; \\ \langle u, u \rangle = 0 &\implies u = 0,\end{aligned}$$

para quaisquer $u, v, w \in V$ e $\alpha \in \mathbb{R}$.

Exemplo 8. Em \mathbb{R}^n , o produto interno usual é dado por

$$\langle u, v \rangle = u_1 v_1 + \cdots + u_n v_n = u^\top v,$$

considerando-se $u, v \in \mathbb{R}^n$ vetores-coluna. Este produto interno é conhecido como produto Euclidiano.

Definição 20. Seja V um espaço de produto interno. Definimos como $\|\cdot\|: V \rightarrow \mathbb{R}$ a função

$$\|u\| = \sqrt{\langle u, u \rangle},$$

chamada *norma*.

Exemplo 9. A norma induzida pelo produto interno usual em \mathbb{R}^n é dada por

$$\|u\| = \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{u^\top u},$$

considerando-se $u \in \mathbb{R}^n$ um vetor-coluna. Naturalmente, esta norma é conhecida como norma Euclidiana ou norma-2.

A noção de ortogonalidade entre vetores está relacionada a produtos internos nulos.

Definição 21. Dois vetores $u, v \in V$ são ditos *ortogonais* em um espaço de produto interno V quando $\langle u, v \rangle = 0$. Além disso, se os vetores possuírem norma 1, dizemos que u e v são *ortonormais*.

Definição 22. Uma base B de um espaço de produto interno V é dita *ortonormal* quando os vetores da base são 2-a-2 ortonormais.

É sempre possível encontrar uma base ortonormal para um espaço vetorial não-trivial, conforme mostra o seguinte teorema.

Teorema 5 (Ortonormalização de Gram-Schmidt). *Seja V um espaço de produto interno de dimensão finita. Se $V \neq \{0\}$, então V possui uma base ortonormal.*

Demonstração. V possui uma base finita $B = \{e_1, \dots, e_n\}$. Vamos tomar primeiro o vetor e_1 e construir o vetor

$$f_1 = e_1.$$

Após isso, vamos construir o vetor

$$f_2 = e_2 - \frac{\langle f_1, e_2 \rangle}{\langle f_1, f_1 \rangle} f_1.$$

Esse vetor é ortogonal a f_1 , pois

$$\langle f_1, f_2 \rangle = \langle f_1, e_2 \rangle - \frac{\langle f_1, e_2 \rangle}{\langle f_1, f_1 \rangle} \langle f_1, f_1 \rangle = 0.$$

Agora, construímos

$$f_3 = e_3 - \frac{\langle f_1, e_3 \rangle}{\langle f_1, f_1 \rangle} f_1 - \frac{\langle f_2, e_3 \rangle}{\langle f_2, f_2 \rangle} f_2$$

e temos tanto

$$\langle f_1, f_3 \rangle = \langle f_1, e_3 \rangle - \frac{\langle f_1, e_3 \rangle}{\langle f_1, f_1 \rangle} \langle f_1, f_1 \rangle - \frac{\langle f_2, e_3 \rangle}{\langle f_2, f_2 \rangle} \underbrace{\langle f_1, f_2 \rangle}_{=0} = 0,$$

quanto

$$\langle f_2, f_3 \rangle = \langle f_2, e_3 \rangle - \frac{\langle f_1, e_3 \rangle}{\langle f_1, f_1 \rangle} \underbrace{\langle f_2, f_1 \rangle}_{=0} - \frac{\langle f_2, e_3 \rangle}{\langle f_2, f_2 \rangle} \langle f_2, f_2 \rangle = 0,$$

e o conjunto $\{f_1, f_2, f_2\}$ é ortogonal. Note que dividir por $\langle f_2, f_2 \rangle$ é possível, pois, caso contrário, teríamos $f_2 = 0$, o que faria e_2 uma combinação linear de e_1 , o que é impossível, pois $\{e_1, e_2\}$ é um conjunto linearmente independente. Similarmente, continuamos o processo com

$$f_k = e_k - \sum_{i=1}^{k-1} \frac{\langle f_i, e_k \rangle}{\langle f_i, f_i \rangle} f_i,$$

sendo $\langle f_j, f_k \rangle$ para todo $j < k$, até a etapa $k = n$. Assim obtemos um conjunto $B' = \{f_1, \dots, f_n\}$, com $f_i \neq 0$, $i \in \mathbb{N}$, $1 \leq i \leq n$, ortogonal. Da expressão para f_k , podemos ver que com os elementos de B' podemos escrever cada um dos vetores $e_k \in B$, logo $\langle B' \rangle = V$. Além disso, sendo $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tais que

$$\alpha_1 f_1 + \dots + \alpha_n f_n = 0,$$

podemos tomar o produto interno para cada f_i , obtendo

$$0 = \langle f_i, \alpha_1 f_1 + \dots + \alpha_n f_n \rangle = \alpha_i \langle f_i, f_i \rangle \implies \alpha_i = 0,$$

para todo $i \in \mathbb{N}$, $1 \leq i \leq n$, mostrando que B' é um conjunto linearmente independente, portanto uma base de V . Finalmente, ao normalizar os vetores como

$$g_i = \frac{f_i}{\|f_i\|} = \frac{f_i}{\sqrt{\langle f_i, f_i \rangle}},$$

obtemos uma base $B'' = \{g_1, \dots, g_n\}$ ortonormal. \square

Dado um produto interno, podemos olhar para o conjunto de todos os vetores ortogonais a uma dada coleção de vetores (em particular, um dado subespaço).

Definição 23. Dado um espaço de produto interno V e um subespaço $S \subset V$, definimos o *complemento ortogonal* de S em V como o conjunto de todos os vetores de v que são ortogonais a todos os vetores de S . Esse conjunto é denotado por S^\perp .

Vamos demonstrar algumas proposições que caracterizam os complementos ortogonais em espaços de dimensão finita.

Proposição 13. *Seja V um espaço de produto interno e $S \subset V$ um subespaço. Então S^\perp também é um subespaço de V .*

Demonstração. Seja $v \in S$ arbitrário. Segue que $\langle v, 0 \rangle = 0$, logo $0 \in S^\perp$. Seja $x, y \in S^\perp$. Segue que $\langle v, x \rangle = 0$ e $\langle v, y \rangle = 0$, logo

$$\langle v, x + y \rangle = \langle v, x \rangle + \langle v, y \rangle = 0 + 0 = 0,$$

o que mostra que $x + y \in S^\perp$. Finalmente, com $\alpha \in \mathbb{R}$, temos

$$\langle v, \alpha x \rangle = \alpha \langle v, x \rangle = \alpha \cdot 0 = 0,$$

portanto $\alpha x \in S^\perp$ e concluímos que S^\perp é um subespaço de V . \square

Proposição 14. *Seja V um espaço de produto interno e $R \subset V$ e $S \subset V$ subespaços tais que $R \subset S$. Então $S^\perp \subset R^\perp$.*

Demonstração. Seja $x \in S^\perp$ e $s \in R$. Então, para todo $t \in S$, vale que $\langle x, t \rangle = 0$. Em particular, $s \in S$, logo $\langle x, s \rangle = 0$ e $x \in R^\perp$. \square

Proposição 15. *Seja V um espaço de produto interno de dimensão finita e $S \subset V$ um subespaço. Então $V = S \oplus S^\perp$.*

Demonstração. Seja $B_S = \{e_1, \dots, e_k\}$ uma base ortonormal de S . Podemos estender B_S para uma base ortonormal de V igual a $B = \{e_1, \dots, e_n\}$, com $B_S \subset B$. Isso pode ser feito via o Teorema 1 e ortonormalizando a base conforme o Teorema 5, usando primeiro os vetores de B_S e notando que eles se manterão inalterados durante o processo.

Seja $B' = B \setminus B_S$. Note que esse conjunto $B' = \{e_{k+1}, \dots, e_n\}$ é linearmente independente. Os elementos de B' são, por construção,

ortogonais aos elementos de B_S , ou seja, $B' \subset S^\perp$. Seja $v \in S^\perp$. Escreva v como

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Esse vetor deve ser ortogonal a todos os vetores de B_S , levando a

$$0 = \langle e_i, v \rangle = \langle e_i, \alpha_1 e_1 + \cdots + \alpha_n e_n \rangle = \alpha_i \langle e_i, e_i \rangle \implies \alpha_i = 0,$$

para todo $i \in \mathbb{N}$, $1 \leq i \leq k$. Assim,

$$v = \alpha_1 e_{k+1} + \cdots + \alpha_n e_n,$$

mostrando que S^\perp é gerado por B' . Disso segue que B' é uma base de S^\perp . Além disso, da construção de B , segue que $V = S + S^\perp$. Sendo $v \in S$ e $v \in S^\perp$, temos, em particular, $\langle v, v \rangle = 0$, levando a $v = 0$, o que mostra que $S \cap S^\perp = \{0\}$ e finalmente $V = S \oplus S^\perp$. \square

Proposição 16. *Seja V um espaço de produto interno de dimensão finita e $S \subset V$ um subespaço. Então $(S^\perp)^\perp = S$.*

Demonstração. Ao usar as bases B_S , B e B' da proposição anterior, vemos que os elementos ortogonais aos elementos de S^\perp são justamente os elementos de S . \square

Dada essa caracterização inicial de espaços de produto interno, podemos agora definir o que é a adjunta de uma transformação linear. É a transformação que “troca de lado” dentro do produto interno.

Definição 24. *Seja U e V espaços de produto interno real, respectivamente $\langle \cdot, \cdot \rangle_U$ e $\langle \cdot, \cdot \rangle_V$, de dimensão finita e $A: U \rightarrow V$ uma transformação linear. Definimos como $A^\top: V \rightarrow U$ a transformação linear tal que*

$$\langle A(u), v \rangle_V = \langle u, A^\top(v) \rangle_U,$$

para todos $u \in U$ e $v \in V$. A transformação A^\top é chamada *transformação adjunta* de A .

Exemplo 10. Seja $V = \mathbb{R}^2$ munido de seu produto interno usual e seja $T(x, y) = (x + y, y - x)$ um operador linear nesse espaço. Note que a base $B = \{(1, 0), (0, 1)\}$ é ortonormal nesse espaço. A expressão de T nessa base é dada por

$$[T]_B = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Seja $P: V \rightarrow V$ o operador linear cuja representação na base B é dada pela matriz transposta de $[T]_B$, na forma

$$[P]_B = [T]_B^\top = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Podemos descrever P também na forma $P(x, y) = (x - y, x + y)$. Seja $u = (x, y)$ e $v = (a, b)$. Ao calcular os produtos internos

$$\langle T(u), v \rangle = \langle (x + y, y - x), (a, b) \rangle = ax + ay - bx + by,$$

e

$$\langle u, P(v) \rangle = \langle (x, y), (a - b, a + b) \rangle = xa - xb + ya + yb,$$

vemos que eles são iguais, o que nos permite concluir que $P = T^\top$ nesse espaço de produto interno.

Observação 5. A transformação adjunta em espaços de dimensão finita sempre existe e é única. Basta generalizar o exemplo anterior e notar que, fixadas bases ortonormais, a matriz da transformação adjunta é a transposta da matriz da transformação original. Note também que é necessário que as bases sejam ortonormais para que esse argumento funcione.

Definição 25. Seja V um espaço de produto interno real de dimensão finita e $A: V \rightarrow V$ um operador linear. Diz-se que A é um *operador simétrico* quando

$$\langle A(u), v \rangle = \langle u, A(v) \rangle$$

para todos $u, v \in V$, ou seja, quando $A = A^\top$.

Observação 6. Um operador é simétrico quando, dada uma base ortonormal, sua matriz nessa base é uma matriz simétrica. Novamente, o requerimento da base ser ortonormal é essencial.

A importância de transformações simétricas é que elas satisfazem um resultado chamado do Teorema espectral, que diz que além delas serem diagonalizáveis, é também possível escolher uma base ortonormal de autovetores que fazem essa diagonalização. Para mostrar isso, provaremos antes uma sequência de lemas — que farão uso da próxima definição.

Definição 26. Seja V um espaço vetorial e $T: V \rightarrow V$ um operador linear. Um subespaço $U \subset V$ é dito *invariante* sobre T se $T(U) \subset U$.

Lema 2. *Seja V um espaço vetorial de dimensão 1 e $T: V \rightarrow V$. Então T é diagonalizável.*

Demonstração. Seja $B = \{p\}$ uma base de V . Segue, pela Proposição 3, que $p \neq 0$. Escrevendo $T(p)$ na base B , temos que $T(p) = \alpha p$, para algum $\alpha \in \mathbb{K}$. Logo p é um autovetor de T , com autovalor associado α , mostrando que B é uma base de V formada por autovetores de T . \square

Lema 3. *Seja V um espaço de produto interno real de dimensão finita, $T: V \rightarrow V$ um operador simétrico e $p \in V$ um autovetor de T . Então $T(\langle\{p\}\rangle^\perp) \subset \langle\{p\}\rangle^\perp$, isto é, o complemento ortogonal $\langle\{p\}\rangle^\perp$ é um subespaço invariante de T .*

Demonstração. Suponha $x \in \langle p \rangle^\perp$ e que $\lambda \in \mathbb{R}$ seja o autovalor associado a p . Então

$$\langle T(x), p \rangle = \langle x, T(p) \rangle = \langle x, \lambda p \rangle = \lambda \langle x, p \rangle = 0,$$

portanto $T(x) \in \langle p \rangle^\perp$. \square

Lema 4. *Seja V um espaço de produto interno real de dimensão finita, $T: V \rightarrow V$ um operador simétrico e $p \in V$ um autovetor de T . Então a restrição de T como operador no subespaço $\langle \{p\} \rangle^\perp$ está bem definida e é também um operador simétrico.*

Demonstração. Pelo Lema 3, podemos sempre construir uma restrição da transformação linear representada por A ao subespaço $\langle v \rangle^\perp$, da forma $A|_{\langle v \rangle^\perp}: \langle v \rangle^\perp \rightarrow \langle v \rangle^\perp$. Essa restrição é simétrica, pois, para todos $x, y \in \langle v \rangle^\perp$, temos

$$\langle A|_{\langle v \rangle^\perp} x, y \rangle = \langle x, A|_{\langle v \rangle^\perp} y \rangle.$$

\square

Lema 5. *Seja V um espaço de produto interno real não-trivial de dimensão finita $n \geq 2$ e $T: V \rightarrow V$ um operador simétrico. Então T possui pelo menos um par de autovalores e autovetores.*

Demonstração. Seja $f: V \rightarrow \mathbb{R}$ dada por $f(x) = \langle T(x), x \rangle$, denominada a forma quadrática de T . Seja $S = \{x \in V: \|x\| = 1\}$. O conjunto S é compacto, logo f atinge um máximo em S . Seja $\lambda \in \mathbb{R}$ esse máximo e $p \in S$ um ponto em que f o atinge, isto é, $f(p) = \lambda$. Afirmamos que λ é um autovalor de A associado ao autovetor p .

Seja $\dim V = n$. Agora, denote por $\langle p \rangle$ o subespaço gerado por p . Seja $u \in \langle p \rangle^\perp$ um vetor de seu complemento ortogonal com norma unitária, que existe porque $\dim(\langle p \rangle^\perp) = n - 1 > 0$. Quando $s \in \mathbb{R}$ é tal que $-1 < s < 1$, segue que $\sqrt{1 - s^2}p + su \in S$, pois

$$\left\| \sqrt{1 - s^2}p + su \right\|^2 = (1 - s^2)\|p\|^2 + s^2\|u\|^2 = 1,$$

ou seja, quando s percorre $(-1, 1)$, o vetor $\sqrt{1 - s^2}p + su$ percorre um caminho em S .

Vamos então definir uma função $g: (-1, 1) \rightarrow \mathbb{R}$ como

$$g(s) = f\left(\sqrt{1-s^2}p + su\right).$$

Note que $g(0) = f(p) = \lambda$, o ponto de máximo em S . Agora, note que, pelo fato de T ser simétrico, temos $\langle T(x), y \rangle = \langle x, T(y) \rangle$, para todos $x, y \in \mathbb{R}^n$. Vamos usar isso para escrever g de forma explícita como

$$\begin{aligned} g(s) &= f\left(\sqrt{1-s^2}p + su\right) \\ &= \left\langle \sqrt{1-s^2}T(p) + sT(u), \sqrt{1-s^2}p + su \right\rangle \\ &= (1-s^2)\langle T(p), p \rangle + s\sqrt{1-s^2}\langle T(p), u \rangle + s\sqrt{1-s^2}\langle T(u), p \rangle \\ &\quad + s^2\langle T(u), u \rangle \\ &= (1-s^2)\langle T(p), p \rangle + 2s\sqrt{1-s^2}\langle T(p), u \rangle + s^2\langle T(u), u \rangle. \end{aligned}$$

Essa função é diferenciável e

$$g'(s) = 2s(\langle T(u), u \rangle - \langle T(p), p \rangle) + 2\left(\sqrt{1-s^2} - \frac{s^2}{\sqrt{1-s^2}}\right)\langle T(p), u \rangle.$$

Lembramos que $s = 0$ é um ponto de máximo, logo deve ser que $g'(0) = 0$, uma relação que impõe $\langle T(p), u \rangle = 0$. Por construção, isso vale para qualquer $u \in \langle p \rangle^\perp$ com $\|u\| = 1$. Para outros vetores $u' \in \langle p \rangle^\perp$, $u' \neq 0$, com norma não necessariamente unitária, temos também

$$\langle T(p), u' \rangle = \|u'\| \left\langle T(p), \frac{u'}{\|u'\|} \right\rangle = 0,$$

ou seja, $T(p)$ é ortogonal a qualquer vetor de $\langle p \rangle^\perp$, nos levando à conclusão de que $T(p) \in \langle p \rangle$. Em outras palavras, existe $\mu \in \mathbb{R}$ tal que $T(p) = \mu p$. Além disso, $\lambda = f(p) = \langle T(p), p \rangle = \langle \mu p, p \rangle = \mu \|p\|^2 = \mu$, sendo λ um autovalor associado ao autovetor $p \neq 0$. \square

Agora demonstraremos o Teorema Espectral para operadores simétricos..

Teorema 6 (Espectral). *Seja V um espaço de produto interno real de dimensão finita e $T: V \rightarrow V$ um operador simétrico. Então T é diagonalizável com uma base ortonormal de autovetores. Como consequência, todos seus autovalores são reais.*

Demonstração. A prova é por indução. O caso $n = 1$ é trivial, como dito antes. Agora, vamos assumir que o resultado é válido para $\dim V = n$. Vamos provar que disso decorre que o resultado é válido para $n + 1$. Seja $\dim V = n + 1$. O operador $T: V \rightarrow V$ possui um autovalor $\lambda \in \mathbb{R}$ associado a um autovetor $v \in V$, $v \neq 0$, conforme a discussão anterior. Podemos então tomar a restrição $T|_{\langle v \rangle^\perp}: \langle v \rangle^\perp \rightarrow \langle v \rangle^\perp$, sendo $\dim(\langle v \rangle^\perp) = n + 1 - 1 = n$. Por indução existe uma base ortonormal $B' = \{v_1, \dots, v_n\}$ de $\langle v \rangle^\perp$ formada por autovetores de $T|_{\langle v \rangle^\perp}$, que também são autovetores de T . Por construção, v é ortogonal a qualquer um dos vetores de B' , logo $B = B' \cup \left\{ \frac{v}{\|v\|} \right\}$ é uma base ortonormal de V . \square

Observação 7. Se as bases B e B' de um espaço de produto interno V são ortonormais, então a matriz de mudança de base $O = P_{B \rightarrow B'}$ é uma matriz ortogonal, no sentido de que $O^\top O = I_n = O O^\top$. Isso significa que, dada uma matriz simétrica $A \in \mathbb{R}^{n \times n}$, segue que ela é ortonormalmente diagonalizável — como a representação de um operador linear na base canônica sob o produto interno usual. Em outras palavras, existe uma matriz ortonormal $O \in \mathbb{R}^{n \times n}$ (a matriz de mudança de base para a base ortonormal de autovetores) e uma matriz diagonal $\Lambda \in \mathbb{R}^{n \times n}$ tal que $\Lambda = O^\top A O$. Note que poderíamos ter escrito $\Lambda = O^{-1} A O$, já que $O^{-1} = O^\top$.

Uma classe especial de operadores simétricos são os operadores positivos, cuja definição e nomenclatura precisa está a seguir.

Definição 27. Seja V um espaço de produto interno real e $T: V \rightarrow V$ um operador simétrico. T é dito *positivo semi-definido* quando, para todo $v \in V$, vale $\langle T(v), v \rangle \geq 0$. Além disso, quando $\langle T(v), v \rangle = 0 \implies v = 0$, T é dito *positivo definido*.

O próximo resultado caracteriza completamente os operadores positivos.

Teorema 7. *Seja V um espaço de produto interno real de dimensão finita e $T: V \rightarrow V$ um operador simétrico. Então T é*

- *positivo definido se, e somente se, todos os autovalores de T são positivos;*
- *positivo semi-definido se, e somente se, todos os autovalores de T são maiores ou iguais a zero.*

Demonstração. • Seja $\{v_1, \dots, v_n\}$ uma base ortonormal de autovetores de T . Se T é positivo definido, então $\langle T(x), x \rangle > 0$ para qualquer $x \neq 0$. Em particular, isso é válido para os autovetores, pois eles são não-nulos. Seja λ_i o autovalor associado a v_i . Então

$$0 < \langle T(v_i), v_i \rangle = \langle \lambda v_i, v_i \rangle = \lambda \|v_i\|^2 \implies \lambda > 0.$$

- Seja $\{v_1, \dots, v_n\}$ uma base ortonormal de autovetores de T . Se T é positivo semi-definido, então $\langle T(x), x \rangle \geq 0$ para qualquer $x \in V$. Em particular, isso é válido para os autovetores. Seja λ_i o autovalor associado a v_i . Então

$$0 \leq \langle T(v_i), v_i \rangle = \langle \lambda v_i, v_i \rangle = \lambda \|v_i\|^2 \implies \lambda \geq 0.$$

□

Observação 8. Isso significa que uma matriz simétrica $A \in \mathbb{R}^{n \times n}$ é positiva (semi-)definida quando existe uma matriz ortogonal $O \in \mathbb{R}^{n \times n}$

e uma matriz diagonal $\Lambda \in \mathbb{R}^{n \times n}$, com sua diagonal formada por elementos positivos (não-negativos), tal que $\Lambda = O^T A O$.

2.4 DECOMPOSIÇÃO EM VALORES SINGULARES (SVD)

Para desenvolver a teoria da decomposição em valores singulares de uma transformação linear $A: U \rightarrow V$, precisamos estudar os quatro espaços fundamentais dessa transformação: $N(A)$, $\text{Im}(A)$, $N(A^T)$ e $\text{Im}(A^T)$. O primeiro resultado investiga a relação desses espaços por meio de complementos ortogonais.

Proposição 17. *Sejam U e V espaços de produto interno de dimensão finita e $A: U \rightarrow V$ uma transformação linear. Então vale que*

1. $N(A^T) = \text{Im}(A)^\perp$;
2. $\text{Im}(A^T) = N(A)^\perp$.

Demonstração. 1. Seja primeiro $x \in N(A^T)$. Nós vemos que, sendo $y \in \text{Im}(A)$, existe $z \in \mathbb{R}^n$ tal que $y = A(z)$ e

$$\langle x, y \rangle = \langle x, A(z) \rangle = \langle A^T(x), z \rangle = \langle 0, z \rangle = 0,$$

logo $x \in \text{Im}(A)^\perp$ e $N(A^T) \subset \text{Im}(A)^\perp$. Agora, seja $x \in \text{Im}(A)^\perp$. Para qualquer $y \in \text{Im}(A)$ vale que $\langle x, y \rangle = 0$. Note que $A(A^T(x)) \in \text{Im}(A)$, logo

$$\|A^T(x)\|^2 = \langle A^T(x), A^T(x) \rangle = \langle x, A(A^T(x)) \rangle = 0,$$

portanto $A^T(x) = 0$, ou seja, $x \in N(A^T)$, logo $\text{Im}(A)^\perp \subset N(A^T)$, completando a prova.

2. Seja $y \in \text{Im}(A^T)$. Existe $x \in \mathbb{R}^m$ tal que $y = A^T x$. Seja $z \in N(A)$ arbitrário. Então temos que

$$\langle y, z \rangle = \langle A^T(x), z \rangle = \langle x, A(z) \rangle = \langle x, 0 \rangle = 0$$

e $y \in N(A)^\perp$, o que mostra que $\text{Im}(A^\top) \subset N(A)^\perp$. Para a outra direção, $N(A)^\perp \subset \text{Im}(A^\top)$, equivalentemente mostrar que $\text{Im}(A^\top)^\perp \subset N(A)$, pois $(S^\perp)^\perp = S$. Seja $x \in \text{Im}(A^\top)^\perp$. Para todo, $y \in \text{Im}(A^\top)$, segue que $\langle x, y \rangle = 0$. Note que, em particular, temos $A^\top(A(x)) \in \text{Im}(A^\top)$. Daí,

$$\|A(x)\|^2 = \langle A(x), A(x) \rangle = \langle x, A^\top(A(x)) \rangle = 0,$$

o que indica que $A(x) = 0$, ou seja, $x \in N(A)$. Isso mostra que $\text{Im}(A^\top)^\perp \subset N(A)$, que implica $N(A)^\perp \subset \text{Im}(A^\top)$ e $N(A)^\perp = \text{Im}(A^\top)$. □

A seguir, vamos começar a mostrar resultados relacionados à composição de A com sua adjunta. Primeiro, vamos mostrar o seguinte.

Proposição 18. *Sejam U , V e W espaços de produto interno de dimensão finita e $A: U \rightarrow V$ e $B: V \rightarrow W$ transformações lineares. Então $(B \circ A)^\top = A^\top \circ B^\top$.*

Demonstração. Basta verificar que

$$\begin{aligned} \langle (B \circ A)u, v \rangle &= \langle B(A(u)), v \rangle = \langle A(u), B^\top(v) \rangle \\ &= \langle u, A^\top(B^\top(v)) \rangle = \langle u, (A^\top \circ B^\top)(v) \rangle \end{aligned}$$

para quaisquer $u \in U$ e $v \in W$. □

Observação 9. Similarmente, é possível mostrar que $(A^\top)^\top = A$. Basta ver que

$$\langle (A^\top)^\top(u), v \rangle = \langle u, A^\top(v) \rangle = \langle A(u), v \rangle$$

para quaisquer $u \in U$ e $v \in V$.

Observação 10. Dadas transformações lineares $A: U \rightarrow V$ e $B: V \rightarrow W$, é trivial provar que $B \circ A$ é uma transformação linear $U \rightarrow W$. Além disso, se B_U , B_V e B_W são bases dos respectivos espaços, então, em termos de representações matriciais, $[B \circ A]_{B_U B_W} = [B]_{B_V B_W} [A]_{B_U B_V}$.

Agora vamos mostrar que as composições de A com sua adjunta são operadores simétricos e positivos semi-definidas, o que mais para frente nos permitirá usar o teorema espectral nesses operadores.

Proposição 19. *Sejam U e V espaços de produto interno reais de dimensão finita e $A: U \rightarrow V$ uma transformação linear. Então as transformações $A^\top \circ A: U \rightarrow U$ e $A \circ A^\top: V \rightarrow V$ são simétricas e positivas semi-definidas.*

Demonstração. Temos

$$\begin{aligned}(A^\top \circ A)^\top &= A^\top \circ (A^\top)^\top = A^\top \circ A \\ \text{e } (A \circ A^\top)^\top &= (A^\top)^\top \circ A^\top = A \circ A^\top,\end{aligned}$$

logo as transformações são simétricas. Agora, seja $u \in U$ e $v \in V$. Temos

$$\begin{aligned}\langle (A^\top \circ A)(u), u \rangle &= \langle A^\top(A(u)), u \rangle = \langle A(u), A(u) \rangle \geq 0 \\ \text{e } \langle (A \circ A^\top)(v), v \rangle &= \langle A(A^\top(v)), v \rangle = \langle A^\top(v), A^\top(v) \rangle \geq 0,\end{aligned}$$

completando a prova. □

Em seguida, um par de resultados sobre os espaços fundamentais dessas composições.

Proposição 20. *Sejam U e V espaços de produto interno reais de dimensão finita e $A: U \rightarrow V$ uma transformação linear. Valem as seguintes igualdades de subespaços:*

1. $\text{Im}(A^\top \circ A) = \text{Im}(A^\top)$;
2. $N(A^\top \circ A) = N(A)$;
3. $\text{Im}(A \circ A^\top) = \text{Im}(A)$;
4. $N(A \circ A^\top) = N(A^\top)$.

Demonstração. Mostraremos apenas 1 e 2, pois 3 e 4 seguem deles ao trocar o papel de A^\top por A e A^\top por $(A^\top)^\top = A$.

1. É equivalente a 2 ao tomarmos complementos ortogonais dos espaços.
2. É óbvio que $N(A) \subset N(A^\top \circ A)$. Seja então $u \in N(A^\top \circ A)$, ou seja, $A^\top(A(u)) = 0$. Temos

$$\|A(u)\|^2 = \langle A(u), A(u) \rangle = \langle u, A^\top(A(u)) \rangle = \langle u, 0 \rangle = 0,$$

ou seja, $A(u) = 0$ e $u \in N(A)$.

□

Corolário 4. *Sejam U e V espaços de produto interno reais de dimensão finita e $A: U \rightarrow V$ uma transformação linear. Valem as seguintes igualdades de posto e nulidade:*

1. $\text{posto}(A^\top \circ A) = \text{posto}(A^\top) = \text{posto}(A) = \text{posto}(A \circ A^\top)$;
2. $\text{nul}(A^\top \circ A) = \text{nul}(A)$;
3. $\text{nul}(A \circ A^\top) = \text{nul}(A^\top)$.

Demonstração. Todos os resultados são consequência imediata da proposição acima, exceto a igualdade $\text{posto}(A^\top) = \text{posto}(A)$.

Temos que $\dim(V) = \text{nul}(A^\top) + \text{posto}(A^\top)$. Como $N(A^\top) = \text{Im}(A)^\perp$, segue que $\text{nul}(A^\top) = \dim \text{Im}(A)^\perp = \dim(V) - \text{posto}(A)$, pois $V = \text{Im}(A) \oplus \text{Im}(A)^\perp$, donde

$$\dim(V) = \dim(V) - \text{posto}(A) + \text{posto}(A^\top)$$

e $\text{posto}(A) = \text{posto}(A^\top)$. □

Agora resta mostrar a resultado principal do capítulo.

Teorema 8 (Decomposição em Valores Singulares). *Sejam U e V espaços de produto interno reais de dimensão finita e $A: U \rightarrow V$ uma transformação linear de posto r . Seja $\dim U = n$ e $\dim V = m$. Denote por A^\top a transformação adjunta de A . Então existem uma base ortonormal de U dada por*

$$B_U = \left\{ \underbrace{u_1, \dots, u_r}_{\text{base para } \text{Im}(A^\top)}, \underbrace{u_{r+1}, \dots, u_n}_{\text{base para } N(A)} \right\}$$

e uma base ortonormal de V dada por

$$B_V = \left\{ \underbrace{v_1, \dots, v_r}_{\text{base para } \text{Im}(A)}, \underbrace{v_{r+1}, \dots, v_m}_{\text{base para } N(A^\top)} \right\}$$

tais que, conforme denotado acima, $\{u_1, \dots, u_r\}$ é uma base ortonormal para $\text{Im}(A^\top)$, $\{u_{r+1}, \dots, u_n\}$ é uma base ortonormal de $N(A)$, $\{v_1, \dots, v_r\}$ é uma base ortonormal de $\text{Im}(A)$ e $\{v_{r+1}, \dots, v_m\}$ é uma base ortonormal de $N(A^\top)$. Além disso, para cada $i \in \mathbb{N}$, com $1 \leq i \leq r$,

$$A(u_i) = \sigma_i v_i,$$

$$A^\top(v_i) = \sigma_i u_i,$$

sendo $\sigma_i > 0$ escalares — chamados de valores singulares de A — tais que

$$\begin{aligned}(A^\top \circ A)(u_i) &= \sigma_i^2 u_i \\ e \quad (A \circ A^\top)(v_i) &= \sigma_i^2 v_i,\end{aligned}$$

podendo eles ser ordenados como

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0.$$

Demonstração. Temos que a transformação $A^\top \circ A$ é um operador simétrico em U . Logo, pelo teorema espectral, e sendo $\text{posto}(A^\top \circ A) = \text{posto}(A)$, existe uma base ortonormal de U formada por autovetores de $A^\top \circ A$ dada por

$$B_U = \{u_1, \dots, u_n\},$$

com autovalores associados iguais a $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Os autovetores associados ao autovalor 0 formam uma base de $N(A^\top \circ A) = N(A)$. Como $\text{nul}(A) = n - r$, vamos ordenar a base como

$$B_U = \{u_1, \dots, u_r, u_{r+1}, \dots, u_n\},$$

sendo $\{u_{r+1}, \dots, u_n\}$ uma base de $N(A)$. Além disso, o operador $A^\top \circ A$ é positivo semi-definido, logo seus autovalores são não-negativos, o que nos permite ordenar a base B_U de forma que

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0.$$

Dado que $\text{Im}(A^\top) = N(A)^\perp$ e que $\{u_1, \dots, u_r\}$ é uma base de $\text{Im}(A^\top \circ A) = \text{Im}(A^\top)$, segue que $\{u_1, \dots, u_r\}$ é uma base de $N(A)^\perp$, mostrando, em particular, que $A(u_i) \neq 0$ para qualquer $i \in \mathbb{N}$, com $1 \leq i \leq r$.

Agora, vamos definir os valores $\sigma_i = \sqrt{\lambda_i}$, para $i \in \mathbb{N}$, com $1 \leq i \leq r$, e construir os vetores

$$v_i = \frac{1}{\sigma_i} A(u_i)$$

para $i \in \mathbb{N}$, $1 \leq i \leq r$. Disso resulta que

$$A(u_i) = \sigma_i \frac{1}{\sigma_i} A(u_i) = \sigma_i v_i$$

e

$$A^\top(v_i) = \frac{1}{\sigma_i} A^\top(A(u_i)) = \frac{1}{\sigma_i} \sigma_i^2 u_i = \sigma_i u_i,$$

para todos $i \in \mathbb{N}$, com $1 \leq i \leq r$. Como os vetores u_1, \dots, u_r são ortonormais, segue que

$$\begin{aligned} \langle v_i, v_j \rangle &= \left\langle \frac{1}{\sigma_i} A(u_i), \frac{1}{\sigma_j} A(u_j) \right\rangle = \frac{1}{\sigma_i \sigma_j} \langle A(u_i), A(u_j) \rangle \\ &= \frac{1}{\sigma_i \sigma_j} \langle u_i, A^\top(A(u_j)) \rangle = \frac{1}{\sigma_i \sigma_j} \langle u_i, \sigma_j^2 u_j \rangle \\ &= \frac{\sigma_j}{\sigma_i} \langle u_i, u_j \rangle = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}, \end{aligned}$$

para qualquer $i, j \in \mathbb{N}$, com $1 \leq i, j \leq r$, logo os vetores v_1, \dots, v_r são também ortonormais. Do fato que $A(u_i) = \sigma_i v_i$, para todo $i \in \mathbb{N}$, com $1 \leq i \leq r$, e que $A(u_1), \dots, A(u_r)$ é uma base para $\text{Im}(A)$, segue que $\{v_1, \dots, v_r\}$ é uma base para $\text{Im}(A) = N(A^\top)^\perp$, que pode ser estendida para uma base ortonormal de V dada por

$$B_V = \{v_1, \dots, v_r, v_{r+1}, \dots, v_m\},$$

com $\{v_{r+1}, \dots, v_m\}$ sendo uma base ortonormal de $N(A^\top)$. Para finalizar, notamos que

$$(A \circ A^\top)(v_i) = A(A^\top(v_i)) = A(\sigma_i u_i) = \sigma_i A(u_i) = \sigma_i^2 v_i,$$

para $i \in \mathbb{N}$, com $1 \leq i \leq r$, e, sendo $N(A \circ A^\top) = N(A^\top)$, que $\{v_{r+1}, \dots, v_m\}$ é uma base para $N(A \circ A^\top)$, segue que B_V é uma base ortonormal de V formada por autovetores de $A \circ A^\top$, com autovalores σ_i^2 associados a v_i , para $i \in \mathbb{N}$, com $1 \leq i \leq r$, e autovalores 0 associados a $\{v_{r+1}, \dots, v_m\}$. \square

Observação 11. Quando $A \in \mathbb{R}^{m \times n}$ é uma matriz, representando uma transformação linear dos espaços $\mathbb{R}^n \rightarrow \mathbb{R}^m$ munidos dos produtos internos usuais, sua decomposição em valores singulares consiste na existência de uma matriz ortogonal $U \in \mathbb{R}^{m \times m}$, uma matriz ortogonal $V \in \mathbb{R}^{n \times n}$ e uma matriz $\Sigma \in \mathbb{R}^{m \times n}$ com elementos fora da diagonal $[\Sigma]_{ii}$ todos eles iguais a zero, sendo a diagonal formada pelos valores singulares $[\Sigma]_{ii} = \sigma_i \geq 0$, matrizes essas tais que $\Sigma = U^T A V$.

3 MOLECULAR DISTANCE GEOMETRY PROBLEM (MDGP) COM SVD

Seja $n \in \mathbb{N}$ o número de átomos de uma dada molécula e x_1, \dots, x_n os valores de coordenadas de tais átomos, em que

$$x_i = (x_{i,1}, x_{i,2}, x_{i,3}),$$

com $x_{i,k} \in \mathbb{R}$, ($k = 1, 2, 3$), indica a posição do átomo i para cada $i \in \{1, \dots, n\}$ é a k -ésima coordenada do átomo i .

Suponha que as coordenadas x_1, \dots, x_n são desconhecidas, ou seja, desconhecemos as posições em \mathbb{R}^3 de seus n átomos. Suponha também que conhecemos uma lista de distâncias entre cada par de átomos. Seria possível descobrir as posições de cada um dos átomos no espaço, a partir das distâncias entre eles?

Seja S um conjunto de pares (i, j) , com $i, j \in \{1, \dots, n\}$, para os quais conhecemos a distância $d_{i,j} \in \mathbb{R}$ entre eles. Portanto, as coordenadas das posições x_1, \dots, x_n dos n átomos definem o sistema não-lineares de equações:

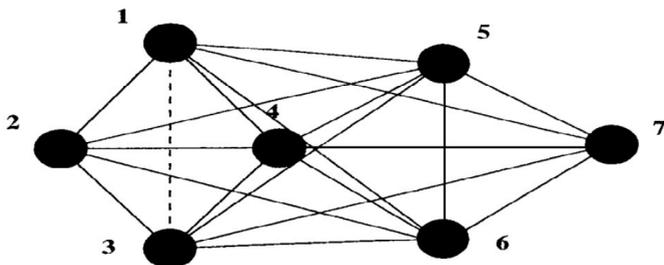
$$\|x_i - x_j\| = d_{i,j} \quad (i, j) \in S \quad (1)$$

Portanto o problema consiste em determinar de modo eficiente as soluções das equações acima. O problema é chamado de *Molecular Distance Geometry Problem* (MDGP), ou Problema Molecular de Geometria de Distâncias.

Vale notar que no problema nem sempre se pode encontrar todas as coordenadas, visto que precisamos um número suficiente de distâncias para definir um sistema viável. Na prática, os valores das distâncias podem conter erros [7], de modo que cada distância deverá ser definida em um intervalo. Entretanto, neste trabalho vamos considerar apenas conjuntos completos de distâncias exatas.

Uma maneira de resolver o MDGP com uma lista completa de distâncias exatas é dando-lhe um tratamento via Decomposição em

Figura 1 – Molécula com sete átomos com conjunto completo de distâncias



Fonte: [8].

Valores Singulares (*Singular Value Decomposition (SVD)*). Tal construção foi publicada no artigo [6]. Acima temos uma figura extraída de [8] exemplificando o problema

3.1 RESOLUÇÃO DO MDGP VIA DECOMPOSIÇÃO EM VALORES SINGULARES

Considerando as equações (1) temos:

$$\|x_i - x_j\| = d_{i,j} \quad i, j = 1, \dots, n \quad (2)$$

e, elevando ao quadrado e expandindo, temos

$$\|x_i\|^2 - 2x_i^T x_j + \|x_j\|^2 = d_{i,j}^2 \quad i, j = 1, \dots, n. \quad (3)$$

Considerando que movimentos rígidos como rotação, reflexão e translação não alteram a estrutura molecular da molécula (em aplicações mais comuns, uma proteína), podemos, sem perda de generalidade, definir um sistema de coordenadas de modo que o último átomo

da lista se torne a origem. Ou seja, consideremos $x_n = (0, 0, 0)$, então:

$$\|x_i\|^2 = \|x_i - (0, 0, 0)\|^2 = \|x_i - x_n\|^2 = d_{i,n}^2 \quad (4)$$

para $i = 1, \dots, n - 1$. Das equações (3) e (4), obtém-se

$$d_{i,n}^2 - d_{i,j}^2 + d_{j,n}^2 = 2x_i^\top x_j, \quad (5)$$

para $i, j = 1, \dots, n - 1$. Define-se, então, a matriz $D = [D_{i,j}]$, em que

$$D_{i,j} = \frac{d_{i,n}^2 - d_{i,j}^2 + d_{j,n}^2}{2}, \quad (6)$$

para $i, j = 1, \dots, n - 1$. Segue das equações (5) e (6), que

$$D_{i,j} = x_i^\top x_j, \quad (7)$$

para $i, j = 1, \dots, n - 1$. Se $X \in \mathbb{R}^{(n-1) \times 3}$ é dada por

$$X = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_{n-1}^\top \end{bmatrix},$$

então, pela Definição de D e X , temos que

$$D = \begin{bmatrix} x_1^\top x_1 & \cdots & x_1^\top x_{n-1} \\ \vdots & \ddots & \vdots \\ x_{n-1}^\top x_1 & \cdots & x_{n-1}^\top x_{n-1} \end{bmatrix} = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_{n-1}^\top \end{bmatrix} \cdot \begin{bmatrix} x_1 & \cdots & x_{n-1} \end{bmatrix},$$

ou seja,

$$D = XX^\top. \quad (8)$$

Como $X \in \mathbb{R}^{(n-1) \times 3}$, temos que o X tem posto menor ou igual a 3. Isso segue do fato de que $\text{posto}(X) = \text{posto}(X^\top)$, provado no

Corolário 4, logo o posto deve ser menor ou igual do que qualquer uma das dimensões de X . Além disso, se

$$x \in \text{Im}(D) \implies \exists y \in \mathbb{R}^{n-1} \text{ tal que } Dy = x,$$

segue que

$$XX^\top y = x \implies X(X^\top y) = x \implies x \in \text{Im}(X) \implies \text{Im}(D) \subseteq \text{Im}(X).$$

Desse modo, o posto de D deve ser menor ou igual a três, uma vez que $\text{posto}(D) \leq \text{posto}(X) \leq 3$. Além disso, D é simétrica e positiva semi-definida, pois $D = XX^\top = D^\top$. Portanto, é possível calcular sua decomposição em valores singulares (SVD) com $U = V$, na forma

$$D = U\Sigma U^\top, \tag{9}$$

onde $U \in \mathbb{R}^{(n-1) \times (n-1)}$ é uma matriz ortogonal, significando que $U^\top U = I_3 = UU^\top$, e $\Sigma \in \mathbb{R}^{(n-1) \times (n-1)}$ é uma matriz diagonal cujos elementos diagonais são $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n-1} \geq 0$, os valores singulares de D . Como $\text{posto}(\Sigma) = \text{posto}(D) \leq 3$, e supondo $n > 3$, segue que podemos escrever a matriz Σ como

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

que podemos reduzir, eliminando colunas nulas, para

$$\Sigma_r = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n-1) \times 3}.$$

Observação 12. Quando $n = 3$, temos que a matriz de valores singulares terá a forma

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix},$$

que não poderá ser reduzida. As contas a seguir poderiam ser baseadas nessa matriz, resultando em vetores de posição estimada em \mathbb{R}^2 , o que reflete o fato de que quaisquer três pontos são coplanares, então uma solução eficiente resultaria em pontos em um espaço bidimensional. Apesar disso, para efeito de consistência, vamos usar, para $n = 3$, a matriz “reduzida” — que na realidade é estendida — igual a

$$\Sigma_r = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix},$$

que, na prática, resultará em soluções da forma $x_i = (x_{i,1}, x_{i,2}, 0)$, o que pode ser interpretado como pertencente a \mathbb{R}^2 se considerarmos a inclusão $\mathbb{R}^2 \subset \mathbb{R}^3$. Considerações similares podem ser feitas para o caso limite de $n = 1$, que resultará em soluções no \mathbb{R}^1 .

Para obter a solução X do problema, usaremos a matriz de

raízes quadradas dos valores singulares, que denotaremos por

$$\Sigma_r^{1/2} = \begin{bmatrix} \sqrt{\sigma_1} & 0 & 0 \\ 0 & \sqrt{\sigma_2} & 0 \\ 0 & 0 & \sqrt{\sigma_3} \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n-1) \times 3}.$$

Essa matriz será usar para construir

$$X = U\Sigma_r^{1/2} \in \mathbb{R}^{(n-1) \times 3}, \tag{10}$$

uma vez que

$$\begin{aligned} XX^\top &= U\Sigma_r^{1/2}(U\Sigma_r^{1/2})^\top = U\Sigma_r^{1/2}(\Sigma_r^{1/2})^\top U^\top \\ &= U\Sigma U^\top = D, \end{aligned}$$

demonstrando a seguinte proposição.

Proposição 21. *Dada ma molécula com n átomos, sendo que o último e colocado na origem do sistema de coordenadas, sejam $D \in \mathbb{R}^{(n-1) \times (n-1)}$ a matriz definida por*

$$D_{i,j} = \frac{d_{i,n}^2 - d_{i,j}^2 + d_{j,n}^2}{2}$$

e

$$D = U\Sigma U^\top$$

sua decomposição em valores singulares, então, a matriz

$$X = U\Sigma_r^{1/2} \in \mathbb{R}^{(n-1) \times 3}$$

cujas linhas são as posições dos $n-1$ primeiros átomos da molécula, resolve a equação matricial

$$D = XX^\top.$$

Quando todas as distâncias entre os átomos da molécula são conhecidas, o MDGP envolvido pode ser resolvido em ordem de complexidade polinomial de grau três. [8]

4 EXPERIÊNCIA COMPUTACIONAL

Para ilustrar os desenvolvimentos do capítulo anterior, vamos mostrar dois exemplos de uso do processo de solução do MDGP. O primeiro exemplo é puramente algébrico e considera uma molécula com $n = 3$ átomos.

Exemplo 11. Vamos supor que temos $n = 3$ vetores de posição $x_1 = (1, 2, -1)$, $x_2 = (0, 3, 1)$ e $x_3 = (0, 0, 0)$. Temos que as distâncias ao quadrado entre esses vetores é

$$[d_{ij}^2] = \begin{bmatrix} 0 & 18 & 6 \\ 18 & 0 & 10 \\ 6 & 10 & 0 \end{bmatrix},$$

o que nos permite construir a matriz de distâncias

$$D = \begin{bmatrix} 6 & -1 \\ -1 & 10 \end{bmatrix}.$$

Essa matriz tem sua SVD dada por

$$D = \begin{bmatrix} -\sqrt{\frac{1}{2} - \frac{1}{\sqrt{5}}} & \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} \\ \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} & \frac{1}{\sqrt{10+4\sqrt{5}}} \end{bmatrix} \begin{bmatrix} 8 + \sqrt{5} & 0 \\ 0 & 8 - \sqrt{5} \end{bmatrix} \begin{bmatrix} -\sqrt{\frac{1}{2} - \frac{1}{\sqrt{5}}} & \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} \\ \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} & \frac{1}{\sqrt{10+4\sqrt{5}}} \end{bmatrix}^T,$$

o que, nos dá, pelo processo de solução do problema,

$$\begin{aligned} X &= \begin{bmatrix} -\sqrt{\frac{1}{2} - \frac{1}{\sqrt{5}}} & \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} \\ \sqrt{\frac{1}{2} + \frac{1}{\sqrt{5}}} & \frac{1}{\sqrt{10+4\sqrt{5}}} \end{bmatrix} \begin{bmatrix} \sqrt{8 + \sqrt{5}} & 0 & 0 \\ 0 & \sqrt{8 - \sqrt{5}} & 0 \end{bmatrix} \\ &= \begin{bmatrix} -\sqrt{3 - \frac{11}{2\sqrt{5}}} & \sqrt{3 + \frac{11}{2\sqrt{5}}} & 0 \\ \sqrt{5 + \frac{2}{2\sqrt{5}}} & \sqrt{5 - \frac{21}{2\sqrt{5}}} & 0 \end{bmatrix} = \begin{bmatrix} x_1^{\text{est}T} \\ x_2^{\text{est}T} \end{bmatrix}, \end{aligned}$$

o que, junto com $x_3^{\text{est}} = (0, 0, 0)$, nos retorna a matriz de distâncias quadradas estimada original:

$$[d_{ij}^2]^{\text{est}} = \begin{bmatrix} 0 & 18 & 6 \\ 18 & 0 & 10 \\ 6 & 10 & 0 \end{bmatrix},$$

indicando que as posições estimadas estão relacionadas às posições originais por meio de uma isometria. Note também que, pelo problema conter apenas três pontos, eles se encontram em um mesmo plano. Isso é indicado pelos pontos da solução estimada terem a terceira coordenada igual a zero.

Em seguida, aplicamos esses mesmos princípios no desenvolvimento de um programa em Julia [5], que implementa a solução do problema para um número arbitrário de átomos.

Exemplo 12. O próximo exemplo usa um código em Julia que gera posições aleatórias para os pontos x_i , os usa para obter uma matriz de distâncias e tenta re-estimar as posições a partir dela. O código usado para isso está listado abaixo.

```
1 using LinearAlgebra
2
3 N = 10;
4
5 X = randn(3,N);
6 X[:,N] = [0;0;0];
7
8 d_sq = zeros(N,N);
9 for i = 1:N
10     for j = 1:N
11         d_sq[i,j] = norm(X[:,i] - X[:,j])^2;
12     end
13 end
14
```

```
15 D = zeros(N-1,N-1);
16
17 for i = 1:N-1
18     for j = 1:N-1
19         D[i,j] = (d_sq[i,N] - d_sq[i,j] + d_sq[j,N])/2;
20     end
21 end
22
23 U, SIGMA0, V = svd(D);
24 SIGMAM = Diagonal(SIGMA0);
25 SIGMASqrt = sqrt(SIGMAM);
26 if N > 3
27     SIGMA = SIGMASqrt[:,1:3];
28 else
29     SIGMA = [SIGMASqrt, zeros(2,1)];
30 end
31 X_est0 = U*SIGMA;
32 X_est = [X_est0; zeros(1,3)];
33
34 d_est = zeros(N,N);
35 for i = 1:N
36     for j = 1:N
37         d_est[i,j] = norm(X_est[i,:] - X_est[j,:])^2;
38     end
39 end
40
41 err = norm(d_est - d_sq);
42 print(err);
```

Inicialmente foi feito um teste com um total de $n = 10$ posições. A norma da diferença obtida entre a matriz de distâncias quadráticas original e estimada — representando o erro de estimação — foi de 4.17×10^{-14} .

Em seguida, foram feitos testes com $n = 100$ e $n = 1000$, obtendo erros iguais a 5.21×10^{-13} e 6.00×10^{-12} , respectivamente. Todos os casos de simulação mostram um casamento perfeito entre os dados originais e os estimados — a menos de uma isomeria. Nota-se apenas

uma degradação na precisão da estimação conforme a dimensionalidade do problema aumenta, embora essa degradação seja pequena o suficiente (de fato, a precisão está, ainda, diversas ordens de grandeza menor do que o esperado para as medidas de posição em qualquer aplicação) para não ser relevante para o presente caso.

Exemplo 13. Os exemplos a seguir usam dados de distâncias de átomos obtidos de moléculas reais, a partir do banco de dados presente em <https://www.rcsb.org/>. Os dados foram obtidos por meio dos *scripts* desenvolvidos por Guilherme Philippi presentes em <https://github.com/caomem/HCProtCLI>.

Foi usada a seguinte proteína para testes:

- 1KYJ (<https://www.rcsb.org/structure/1KYJ>, $N = 144$).

A listagem do código para o processamento desses dados está a seguir.

```
1 using LinearAlgebra
2
3 include("pdb1kyj.m");
4 sz = size(E);
5 N = sz[1];
6 d_s2;
7
8 D = zeros(N-1,N-1);
9
10 for i = 1:N-1
11     for j = 1:N-1
12         D[i,j] = (d_sq[i,N] - d_sq[i,j] + d_sq[j,N])/2;
13     end
14 end
15
16 U, SIGMA0, V = svd(D);
17 SIGMAM = Diagonal(SIGMA0);
18 SIGMASqrt = sqrt(SIGMAM);
```

```
19 if N > 3
20     SIGMA = SIGMAsqrt[:,1:3];
21 else
22     SIGMA = [SIGMAsqrt, zeros(2,1)];
23 end
24 X_est0 = U*SIGMA;
25 X_est = [X_est0; zeros(1,3)];
26
27 d_est = zeros(N,N);
28 for i = 1:N
29     for j = 1:N
30         d_est[i,j] = norm(X_est[i,:] - X_est[j,:])^2;
31     end
32 end
33
34 err = norm(d_est - d_sq);
35 print(err);
```

A métrica de erro obtida para esse caso foi:

- 1KYJ: $\text{err} = 1.3654741375800942 \times 10^{-11}$.

Isso mostra como o programa está correto, produzindo erros pequenos com dados reais.

5 CONCLUSÃO

Em resumo, apresentamos, neste trabalho, um algoritmo para tratar do Problema Molecular de Geometria de Distâncias com um conjunto de distâncias completo e exato. Para isso, utilizamos como ferramenta a decomposição SVD da matriz de distâncias, que teve suas bases algébricas fundamentadas no segundo capítulo. O método foi testado com dados sintéticos e reais, obtidos de uma estrutura de proteína, mostrando que ele corretamente estima as posições do átomo de uma molécula, dada sua matriz de distâncias. Para estudos futuros devemos abordar os casos mais gerais do MDGP, em que serão tratados conjuntos de distâncias totalmente arbitrários.

REFERÊNCIAS

- [1] CB Anfinsen e HA Scheraga. “Experimental and theoretical aspects of protein folding”. Em: *Advances in protein chemistry* 29 (1975), pp. 205–300.
- [2] Christian B Anfinsen. “Principles that govern the folding of protein chains”. Em: *Science* 181.4096 (1973), pp. 223–230.
- [3] Sheldon Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer International Publishing, 2014. ISBN: 97833191110806.
- [4] Konrad Bachmann. “Progress and pitfalls in systematics: cladistics, DNA and morphology”. Em: *Acta Botanica Neerlandica* 44.4 (1995), pp. 403–419.
- [5] Jeff Bezanson, Stefan Karpinski, Viral Shah e Alan Edelman et al. *The Julia Programming Language*. 2008. Disp. em: <https://julialang.org/> (acesso em 18/03/2022).
- [6] Q. Dong e Z. Wu: “A linear-time algorithm for solving the molecular distance geometry problem with exact inter-atomic distances”. Em: *Journal of Global Optimization* 26 (2003), pp. 321–333.
- [7] Felipe Delfini Caetano Fidalgo. “Algoritmos para problemas de geometria molecular”. Diss. de mestr. Campinas: Universidade Estadual de Campinas, 2011, p. 6.

-
- [8] Felipe Delfini Caetano Fidalgo. “Algoritmos para problemas de geometria molecular”. Diss. de mestr. Campinas: Universidade Estadual de Campinas, 2011, p. 11.
- [9] Mário Barone Jr. *Álgebra Linear*. 3a ed. IME-USP, 2005.
- [10] Elon Lages Lima. *Álgebra Linear*. Vol. 1. IMPA, 2014.
- [11] Desmond ST Nicholl. *An introduction to genetic engineering*. Cambridge University Press, 2008.
- [12] Leslie A. Pray. *Discovery of DNA Structure and Function: Watson and Crick*. 2008. Disp. em: <https://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397/>.
- [13] Steven Roman. *Advanced Linear Algebra*. 3a ed. Vol. 3. Graduate Texts in Mathematics. Springer, 2008. ISBN: 978-0-387-72828-5. DOI: <https://doi.org/10.1007/978-0-387-72831-5>.
- [14] Gilberto Souto. *Decomposição em Valores Singulares*. Trabalho de Conclusão de Curso. Florianópolis, 2000.
- [15] P. Suppes. *Axiomatic Set Theory*. Dover Books on Mathematics. Dover Publications, 2012. ISBN: 9780486136875.