

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL
PROFMAT**

CHEIRUM MICHAEL RIBEIRO

**ESTUDO DE PROJEÇÕES COM UMA PROPOSTA DE ATIVIDADES PARA
APLICAÇÃO NO ENSINO MÉDIO**

CURITIBA

2019

CHEIRUM MICHAEL RIBEIRO

**ESTUDO DE PROJEÇÕES COM UMA PROPOSTA DE ATIVIDADES PARA
APLICAÇÃO NO ENSINO MÉDIO**

Dissertação apresentada ao Mestrado Profissional em
Matemática em Rede Nacional da Universidade Tec-
nológica Federal do Paraná em Curitiba - PROFMAT-
UTCT como requisito parcial para obtenção do grau
de Mestre.

Orientador: Rodolfo Gotardi Begiato

CURITIBA

2019

Dados Internacionais de Catalogação na Publicação

Ribeiro, Chéirum Michaael

Estudo de projeções com uma proposta de atividades para aplicação no ensino médio [recurso eletrônico] / Chéirum Michaael Ribeiro. -- 2019.

1 arquivo texto (120 f.): PDF; 22,9 MB.

Modo de acesso: World Wide Web

Título extraído da tela de título (visualizado em 19 dez. 2019)

Texto em português com resumo em inglês Dissertação (Mestrado em Matemática) - Universidade Tecnológica Federal do Paraná. Mestrado Profissional em Matemática em Rede Nacional, Curitiba, 2019

Bibliografia: f. 119-120.

1. Matemática - Dissertações. 2. Geometria projetiva. 3. Planos projetivos. 4. Geometria - Estudo e ensino (Ensino médio). 5. Jogos no ensino de matemática. 6. Matemática - Estudo e ensino (Ensino médio). 7. Geometria analítica plana - Problemas, exercícios, etc. 8. Inovação educacional. 9. Sistemas lineares. I. Begiato, Rodolfo Gotardi. II. Universidade Tecnológica Federal do Paraná. Programa de Mestrado Profissional em Matemática em Rede Nacional. III. Título.

CDD: Ed. 23 – 510

Biblioteca Central da UTFPR, Câmpus Curitiba

Bibliotecário: Adriano Lopes CRB-9/1429

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 73

A Dissertação de Mestrado intitulada “**Estudo de projeções com uma proposta de atividades para aplicação no ensino médio**”, defendida em sessão pública pelo candidato **Cheirum Michael Ribeiro**, no dia 12 de dezembro de 2019, foi julgada para a obtenção do título de Mestre, área de concentração Matemática, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Matemática em Rede Nacional - PROFMAT.

BANCA EXAMINADORA:

Prof. Dr. Rodolfo Gotardi Begiato – UTFPR

Prof. Dr. João Luis Gonçalves – UTFPR

Prof. Dr. Lucas Garcia Pedroso – UFPR

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 12 de dezembro de 2019.

Carimbo e Assinatura do(a) Coordenador(a) do Programa

Dedico este trabalho a Deus, por ser essencial em minha vida, autor de meu destino, meu guia, meu socorro, minha paz.

AGRADECIMENTOS

A Deus, todo poderoso, que desde o início desta caminhada me ungiu e deu sabedoria para vencer todos os desafios que surgiram nestes três anos.

A minha família, em especial na pessoa da minha mãe Maria Aparecida Ribeiro, meu pai Antonio Osni Ribeiro e irmãos Cheisum Micheel Ribeiro e Eduardo Antonio Ribeiro, pelo apoio incondicional.

A minha esposa Grasiela Jung Ribeiro e meus filhos Isaque Jung Ribeiro e Rebeca Jung Ribeiro pelo apoio. A compreensão deles foi essencial para que este trabalho pudesse ter sido realizado.

Ao professor Dr Rodolfo Gotardi Begiato pela orientação e dedicação. Desde o início dos trabalhos foi muito atencioso, paciente e zeloso.

A todos os professores do corpo docente do Profmat responsáveis pelas disciplinas que cursei, pois com muito zelo deram sua contribuição à minha formação.

*"Tudo posso naquele que me fortalece."
(Filipenses 4, 13)*

RESUMO

RIBEIRO, Cheirum Michaael. **Estudo de projeções com uma proposta de atividades para aplicação no Ensino Médio**. 122 f. Dissertação - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

O estudo de projeções é um assunto clássico e de extrema importância para diversas áreas da Matemática. Neste contexto, o presente trabalho tem por objetivo evidenciar essa importância trazendo inicialmente um estudo teórico aprofundado sobre projeções e algumas aplicações na análise numérica, através da exposição dos métodos iterativos GMRES e Gradientes Conjugados, baseados em projeção sobre subespaços de Krylov. Num segundo momento, procurou-se com o conhecimento adquirido propor algumas atividades que podem ser utilizadas na educação básica. Foram elaboradas atividades interdisciplinares envolvendo o conteúdo de Geometria Analítica com as áreas de conhecimento Física e Educação Física, onde propõe-se trabalhar as projeções no contexto de vetores no plano cartesiano e também nas regras de impedimento no futebol, sob o contexto recente da tecnologia VAR (*Video Assistant Referee*) que vem sendo adotada pela FIFA (*Fédération Internationale de Football Association*) na última década. Para realização das atividades, a fim de auxiliar na prática dos experimentos e de fornecer estímulo ao estudante foi produzido um material concreto, denominado *caixa octante*. Tal material representa um octante do \mathbb{R}^3 , e é um recurso que pode ser utilizado para todo conteúdo de Geometria Analítica.

Palavras-chave: Projeções. Ensino Básico. Geometria Analítica. Sistemas Lineares.

ABSTRACT

RIBEIRO, Cheirum Michaael. **Projections study with a proposal of activities for application in High School**. 106 f. Dissertation - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

The projections study is a classic subject of extreme importance for various Math's fields. In this context, the present thesis has the main goal to evidence its importance firstly bringing a depth theoretical study about projections and some applications on the numeric analysis through the exposure of interplay methods GMRES and Conjugates Gradients that are based on projections of Krylov subspace. On the second moment with the knowledge acquired in this work it was sought to propose some activities which can be apply on the basic education. It has been elaborated interdisciplinary activities engaging the content of Analytic Geometry with the Physics and Physical Education knowledge fields. It has proposed working the projections in the vectors context in the cartesian plan and also in the rules of soccer offside under the currently context of VAR (*Video Assistant Referee*) technology that has been assumed for FIFA (*International Federation of Association Football*) in the last decade. It has been built a concrete material called octant box to accomplish the activities as a result to assist the experiments practice as well as giving encouragement to the students. Such material represents an octant of \mathbb{R}^3 and it is a source that can be utilized for all Analytic Geometry content.

Keywords: Projections. Basic Education. Analytic Geometry. Linear Systems.

LISTA DE ILUSTRAÇÕES

Figura 1 – Rotação de vetores I	23
Figura 2 – Rotação de vetores II	24
Figura 3 – Rotação de vetores III	24
Figura 4 – $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$; Projeção ortogonal sobre plano xy	27
Figura 5 – $\mathcal{N}(T) = \{(0, 0, z) \in \mathbb{R}^3 \mid z \in \mathbb{R}\}$	27
Figura 6 – Núcleo e imagem de uma transformação linear $T : E \rightarrow F$	29
Figura 7 – Transformação injetiva	30
Figura 8 – Transformação sobrejetiva	30
Figura 9 – Projeção de vetores	36
Figura 10 – Diferença $v - u$	42
Figura 11 – Projeção ortogonal	46
Figura 12 – Projeção ortogonal do vetor $v = (3, 4, 6)$ sobre o vetor $u = (6, 8, 0)$	47
Figura 13 – $v = proj_W(v) + proj_{W^\perp}(v)$	52
Figura 14 – Projeção do resíduo inicial sobre o subespaço $\mathcal{AK}_l(A, r_0)$	58
Figura 15 – Construção de uma base ortonormal $\{v_1, v_2, \dots, v_l\}$ do subespaço de Krylov do resíduo inicial (r_0) sobre a matriz A por Arnoldi	59
Figura 16 – Gráfico de $f(x)$: o ponto mínimo dessa superfície é solução de $As = b$	77
Figura 17 – Curvas de nível: interseção da superfície com hiperplanos	77
Figura 18 – Fluxograma das atividades	84
Figura 19 – Representação geométrica do vetor definido pelos pontos A e B	86
Figura 20 – $\ \vec{AB}\ = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$	87
Figura 21 – Vetor com extremidade na origem: $\ v\ = d(0, A) = \sqrt{x_1^2 + y_1^2}$	87
Figura 22 – Representação geométrica da projeção do vetor $(3, 2)$ sobre eixo x	88
Figura 23 – Representação geométrica da projeção do vetor \vec{AB} sobre eixo x	88
Figura 24 – Representação geométrica da projeção do vetor \vec{AB} sobre eixo y	88
Figura 25 – Representação geométrica da projeção de vetor paralelo ao eixo y em relação ao eixo x	89
Figura 26 – Representação geométrica da projeção de vetor determinado pelos pontos $A(-1, 1)$ e $B(4, 3)$ sobre reta $y = 1$	89
Figura 27 – Representação geométrica da projeção de vetor definido pelos pontos $A(3, 2)$ e $B(1, 3)$ sobre a reta $r : y = \frac{1}{3}x + 1$	90
Figura 28 – Caixa octante	93
Figura 29 – Exemplo de atividade a ser realizada na caixa octante	93
Figura 30 – Caixa octante montada I	94
Figura 31 – Caixa octante montada II	94
Figura 32 – Representação geométrica do vetor definido pelos pontos A e B	95

Figura 33 – $\ \vec{AB}\ = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$	95
Figura 34 – Vetor com extremidade na origem: $\ v\ = d(0, A) = \sqrt{x_1^2 + y_1^2}$	96
Figura 35 – Representação dos pontos A, B, C e D	96
Figura 36 – Representação das coordenadas do paralelepípedo ABCDEFGH	97
Figura 37 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante I	97
Figura 38 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante II	98
Figura 39 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante com adaptação para alunos com deficiência visual	98
Figura 40 – $\ \vec{AB}\ = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$	99
Figura 41 – Exemplo de atividade realizada na caixa octante com semelhança de triângulos	100
Figura 42 – Sombra de um objeto projetada ortogonalmente por uma lâmpada puntiforme fixa	100
Figura 43 – Projeção não ortogonal com $\theta < 90^\circ$	101
Figura 44 – Projeção não ortogonal com $\theta > 90^\circ$	101
Figura 45 – Obtendo comprimento da sombra de uma lapiseira posicionada no ponto $A(20, 3, 0)$ com inclinação de 45° com relação à base da caixa	102
Figura 46 – Sombra de um objeto qualquer posicionado em ponto arbitrário e com incli- nação de θ com relação à base da caixa	103
Figura 47 – Exemplo de impedimento I	105
Figura 48 – Exemplo de impedimento II	105
Figura 49 – Exemplo de impedimento III	106
Figura 50 – Exemplo de impedimento IV	106
Figura 51 – Posicionamento de câmera para verificação de impedimento I	107
Figura 52 – Posicionamento de câmera para verificação de impedimento II	107
Figura 53 – Lance capturado por emissora de TV	108
Figura 54 – Reação de torcedores	109
Figura 55 – Reação de dirigentes de um dos clubes envolvidos	109
Figura 56 – Declaração do chefe de arbitragem	109
Figura 57 – Imagem capturada e analisada pelo VAR com projeções	110
Figura 58 – Funcionamento do VAR	110
Figura 59 – Experimento simulado do VAR	111
Figura 60 – Definição de vetores correspondentes a cada jogador envolvido na jogada . .	112
Figura 61 – Definição de triângulo retângulo correspondente a cada jogador com relação à base da caixa (campo)	112
Figura 62 – Verificação da não existência de impedimento	113
Figura 63 – Exemplo de definição de vetor correspondente a um jogador	114
Figura 64 – Verificando se atacante está em posição de impedimento	114
Figura 65 – Verificando projeção ortogonal do atacante	115
Figura 66 – Verificando projeção ortogonal do defensor	115

Figura 67 – Verificação da existência de impedimento	116
----------------------------------------------------------------	-----

SUMÁRIO

	INTRODUÇÃO	14
1	TÓPICOS INTRODUTÓRIOS	18
1.1	Transformações Lineares	18
1.1.1	Propriedades	20
1.2	Produto de Transformações Lineares	25
1.3	Núcleo e Imagem	26
1.4	Soma Direta e Projeção	34
1.5	Espaço Vetorial com Produto Interno	38
1.5.1	Produto interno	38
1.5.2	Ortogonalidade e Projeção ortogonal	43
2	PROJEÇÕES E ANÁLISE NUMÉRICA	54
2.1	Motivação	54
2.2	Subespaços de Krylov	55
2.3	Método GMRES	57
2.3.1	Iteração de Arnoldi	58
2.3.2	Rotações de Givens	65
2.4	Método dos Gradientes conjugados	75
3	PROPOSTAS DE ATIVIDADES PARA USO NO ENSINO MÉDIO	84
3.1	Modelagem matemática e interdisciplinaridade	85
3.2	Atividade 1: Projetando vetores em retas no plano	86
3.2.1	Objetivos	86
3.2.2	Desenvolvimento da proposta	86
3.2.2.1	Linguagem vetorial	86
3.2.2.2	Projeção de vetores sobre eixos coordenados	87
3.2.2.3	Projeção de vetores sobre retas no plano	89
3.2.3	Avaliação	91
3.3	Atividade 2: Prática com Luz e Sombra	92
3.3.1	Objetivos	92
3.3.2	Caixa octante	92
3.3.3	Desenvolvimento da proposta	94
3.3.3.1	Linguagem vetorial	94
3.3.3.2	Escolha de um objeto para experimento	99
3.3.3.3	Experimento com objeto	100

3.3.4	Avaliação	104
3.4	Atividade 3: Verificando impedimentos no futebol	104
3.4.1	Objetivos	104
3.4.2	Desenvolvimento da proposta	104
3.4.2.1	Regra de Impedimento	104
3.4.2.2	Funcionamento da tecnologia VAR	106
3.4.2.3	Experimento com VAR	111
3.4.3	Avaliação	116
4	CONCLUSÃO	117
	REFERÊNCIAS	119

INTRODUÇÃO

Neste trabalho vamos tratar o tema projeção sob duas perspectivas. Na primeira perspectiva, visando o aprofundamento do conhecimento matemático na área de Álgebra Linear, foi feito um intenso estudo teórico sobre projeções e algumas de suas aplicações na Análise numérica. Na segunda perspectiva, procurou-se com o conhecimento adquirido propor algumas atividades que poderiam ser utilizadas na educação básica. Sob essa segunda perspectiva foram elaboradas atividades interdisciplinares envolvendo a Geometria Analítica e as áreas de conhecimento Física e Educação Física, onde propõe-se trabalhar as projeções com utilização de vetores no plano cartesiano e também nas regras de impedimento no futebol, sob o contexto recente da tecnologia VAR (Video Assistant Referee) que vem sendo adotada pela FIFA (Fédération Internationale de Football Association) na última década.

Dentro dessas duas perspectivas produzimos o texto que segue, onde empenhamo-nos para fornecer um trabalho que seja útil para professores de matemática, daí o cuidado de fornecer uma redação didática e um grande número de exemplos.

O estudo de projeções é um assunto clássico no estudo de Álgebra Linear com muita utilidade na área de Análise Numérica, onde projetores podem ser utilizados como ferramenta para o desenvolvimento de métodos que podem ser utilizados para a resolução de diversos problemas da área, tais como fatoração de matrizes, otimização linear, otimização não-linear, resolução de sistemas lineares e não lineares, etc. Posteriormente, tais métodos desenvolvidos podem ser empregados para aplicação nas mais diversas áreas de conhecimento, como, Física, Engenharias, Economia, etc.

Considerando um espaço vetorial qualquer E , uma projeção é um operador linear $P : E \rightarrow E$ que satisfaz a condição $P^2 = P$. O tipo de projeção mais estudado são as chamadas projeções ortogonais, em que o operador linear deve satisfazer, adicionalmente, $P = P^T$.

A noção de projeção está intimamente ligada à noção de Soma Direta. Se W_1, W_2 são dois subespaços do espaço vetorial E , dizemos que E é soma direta de W_1 e W_2 se todo elemento $v \in E$ pode ser escrito, de maneira única, como $v = w_1 + w_2$, onde $w_1 \in W_1$ e $w_2 \in W_2$. Num primeiro momento, considerando a projeção $P : E \rightarrow E$, é possível mostrar que a condição $P^2 = P$ garante que E é a soma direta do núcleo da projeção com a sua imagem. Além disso, se E pode ser escrito como soma direta dos subespaços W_1 e W_2 , define-se o operador linear $P : E \rightarrow E$, projeção sobre W_1 , paralelamente a W_2 , do seguinte modo: todo vetor $w \in V$ se escreve, de modo único, como soma $w = u + v$ de um vetor $u \in W_1$ com um vetor $v \in W_2$. Põe-se então $P(w) = u$. Neste caso, pode-se perceber que a imagem da projeção é o espaço W_1 e, além disso, o núcleo é o espaço W_2 .

No caso da projeção ortogonal, pode-se ainda verificar que o núcleo do operador será

complemento ortogonal da imagem. E que, sendo W subespaço de V , tem-se que $V = W \oplus W^\perp$ e que, neste caso, a projeção sobre W , paralelamente a W^\perp é uma projeção ortogonal.

Na área de Análise Numérica, as projeções têm papel importante no desenvolvimento de métodos para resolução dos mais diversos tipos de problema. Para trabalhar com decomposição de matrizes, pode ser aplicado o conceito de projeções para determinar a fatoração QR. Quando se fala da resolução de sistemas lineares, podemos citar, por exemplo, métodos baseados em projeção sobre subespaços de Krylov (KELLEY, 1995). A resolução de sistemas sobredeterminados (mais equações do que incógnitas) através dos métodos Quadrados Mínimos, que constantemente aparecem na área de Otimização, são resolvidos através de projeções em subespaços. Entre os métodos desenvolvidos para resolução de problemas de minimização (linear ou não, restrito ou não) podemos encontrar uma série de métodos que utilizam o conceito de projeção em algumas de suas etapas ou até em sua essência.

Neste trabalhos, vamos estudar dois dos principais métodos iterativos para resolução de sistemas lineares: Método dos Gradientes Conjugados (*Conjugate Gradient Method*) e o GMRES (*Generalized Minimum Residual Method*) que são baseados em projeções sobre subespaços de Krylov.

Após o estudo teórico das projeções, foi possível elaborar algumas propostas para trabalho em sala de aula no ensino básico, nas quais, partindo de algumas atividades teóricas, desenvolve-se uma série de atividades práticas para as quais foi desenvolvido o material concreto caixa octante.

As atividades propostas levam em consideração duas possibilidades advindas das atividades práticas e que constituem importantes fundamentos na aprendizagem qualitativa de Matemática: a modelagem matemática e a interdisciplinaridade.

Sobre a modelagem matemática, as Orientações Curriculares para o Ensino Médio discorrem que como estratégia de ensino, a modelagem apresenta fortes ligações com a ideia de resolução de problema. Neste contexto, frente a situações práticas o aluno precisa mobilizar e articular muitas competências, o que acaba contribuindo para sua aprendizagem, uma vez que precisa selecionar variáveis, problematizar, formular hipóteses, recorrer ao conhecimento matemático, confrontar as conclusões teóricas e eventualmente ainda, quando surge a necessidade, modificar o modelo para que esse melhor corresponda à situação real, aqui se revelando o aspecto dinâmico da construção do conhecimento (BRASIL, 2006a).

Já em relação à interdisciplinaridade, temos que segundo os Parâmetros Curriculares Nacionais (PCN) para o Ensino Médio

Na perspectiva escolar, a interdisciplinaridade não tem a pretensão de criar novas disciplinas ou saberes, mas de utilizar os conhecimentos de várias disciplinas para resolver um problema concreto ou compreender um determinado fenômeno sob diferentes pontos de vista. Em suma, a interdisciplinaridade tem uma função instrumental. Trata-se de recorrer a um saber diretamente útil e utilizável para responder às questões e aos problemas sociais contemporâ-

neos. (...) A integração dos diferentes conhecimentos pode criar as condições necessárias para uma aprendizagem motivadora, na medida em que ofereça maior liberdade aos professores e alunos para a seleção de conteúdos mais diretamente relacionados aos assuntos ou problemas que dizem respeito à vida da comunidade. (BRASIL, 2000)

As propostas de atividades para o uso no ensino médio têm como objetivo central promover um reforço e um estímulo na aprendizagem de conteúdos referente à Geometria Analítica, de maneira interdisciplinar com as áreas de Física e Educação Física, trabalhando com projeções em \mathbb{R}^2 e \mathbb{R}^3 e priorizando conhecimentos geométricos relacionados ao estudo do triângulo retângulo, como as relações trigonométricas e Teorema de Pitágoras, com uma introdução a linguagem vetorial, sobre o plano cartesiano.

Dentre as propostas, a primeira é de cunho teórico, onde se propõe um estudo sobre projeções de vetores do \mathbb{R}^2 em retas quaisquer do plano. Já as outras duas propostas visam trabalhar projeções por meio de dois experimentos práticos. Para tanto, foi desenvolvido um material concreto, o qual chamamos de *caixa octante*. Tal material representa um octante do \mathbb{R}^3 e tem por objetivo não só auxiliar na prática dos experimentos, como também fornecer estímulo ao estudante, a ponto de o mesmo poder visualizar alguns conceitos abstratos.

Uma das atividades a ser realizada com a caixa octante está relacionada com um experimento de luz e sombra. Pretende-se trabalhar com a projeção ortogonal através do problema de calcular o comprimento da sombra de um objeto projetada ortogonalmente por uma lâmpada puntiforme presa ao teto. Nesta atividade, pretende-se sobretudo que o aluno entenda a importância da projeção ortogonal, considerando que ela é a que fornece a menor distância entre a sombra e o objeto projetado.

A segunda atividade prática proposta com a caixa octante está relacionada a uma aplicação da definição de projeção no esporte, mais precisamente a marcações de impedimentos em partidas de futebol.

Conforme as regras da IFAB (The International Football Association Board), que é o órgão responsável pela formulação e manutenção das regras que compõem o futebol, um jogador estará em posição de impedimento quando qualquer parte de sua cabeça, corpo ou pés (exceto mãos e braços, inclusive dos goleiros) estiver no campo adversário e mais próximo da linha de meta adversária do que a bola e o penúltimo adversário. Ademais, um jogador em posição de impedimento somente será punido se no momento em que a bola for jogada ou tocada por um companheiro de equipe (considerando o momento do primeiro ponto de contato com a bola) participar ativamente do jogo, seja interferindo no jogo, interferindo num adversário ou mesmo ganhando vantagem de sua posição de impedimento (IFAB, 2019).

Tal atividade tem como motivação levar os estudantes a um pleno entendimento da regra do impedimento por meio do estudo de projeções minimizando a proposição de críticas e reclamações por vezes decorrentes do desentendimento dessa regra. Na prática, o professor

conduzirá um experimento com utilização de bonecos, onde os mesmos utilizarão conhecimentos relacionados ao estudo de projeção para verificar a existência de impedimentos no experimento.

A ideia é simular a tecnologia VAR (*Video Assistant Referee*) recentemente aprovada pela FIFA (do francês: *Fédération Internationale de Football Association*) em relação à marcação de impedimento. Um árbitro assistente de vídeo (VAR) é um árbitro de partida que possui acesso independente às imagens gravadas da partida, o qual poderá auxiliar o árbitro na eventualidade de um “erro claro e óbvio”, como por exemplo a marcação de impedimento.

Quanto à organização da dissertação, no primeiro capítulo deste trabalho, descrevemos os fundamentos teóricos estudados sob a primeira perspectiva (aprofundamento do conhecimento). Neste sentido, faremos um relato do conceito de projeção sob a ótica da Álgebra Linear. Para tanto, é esperado que o leitor já tenha alguns conhecimentos prévios relacionados a definições e propriedades de matrizes, vetores e espaços vetoriais.

No Capítulo 2, ainda sob a perspectiva de aprofundamento de conhecimento, apresentamos uma aplicação do estudo de projeções no campo da Análise Numérica através dos métodos iterativos para resolução de sistemas lineares: GMRES e Gradientes Conjugados.

O terceiro capítulo é dedicado à perspectiva de aplicações no Ensino Básico. Nele, descrevemos as nossas propostas de atividades, o material concreto desenvolvido e algumas possibilidades de utilização.

Por fim, terminamos este texto com algumas conclusões que pudemos alcançar durante a produção deste trabalho.

1 TÓPICOS INTRODUTÓRIOS

Neste capítulo estudamos conceitos importantes para a compreensão da definição de projeção e aplicações relacionadas, dentre eles o de Transformação Linear e Ortogonalidade. Para tanto, na composição deste capítulo utilizamos alternadamente um conjunto de autores como base teórica, tais quais, (LIMA, 2014), (TEIXEIRA, 2015), (SANTOS, 2006), (HEFEZ; FERNADEZ, 2016), (COELHO; LOURENÇO, 2013) e (NASCIMENTO, 2013).

1.1 TRANSFORMAÇÕES LINEARES

As chamadas transformações lineares, formam uma classe muito especial de funções que têm muitas aplicações na Física, nas Engenharias e em vários ramos da Matemática. Tais funções nas quais se está interessado na Álgebra Linear são as funções cujos domínios e contradomínios são espaços vetoriais e que, além disso, preservam as operações de adição de vetores e de multiplicação de um vetor por um escalar (HEFEZ; FERNADEZ, 2016).

Definição 1.1. *Sejam E e F espaços vetoriais sobre \mathbb{R} . Uma transformação linear é uma correspondência que associa a cada vetor $u \in E$ um vetor $T(u) \in F$ de modo que valham, para quaisquer $u, v \in E$ e qualquer $\alpha \in \mathbb{R}$, as relações:*

$$T(u + v) = T(u) + T(v) \quad (1.1)$$

$$T(\alpha u) = \alpha T(u) \quad (1.2)$$

Exemplo 1.2. A função $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, dada por $T(x, y) = (2x, 0, x + y)$, é uma transformação linear.

De fato, se $u = (x_1, y_1) \in \mathbb{R}^2$, $v = (x_2, y_2) \in \mathbb{R}^2$ e $\alpha \in \mathbb{R}$, temos que

$$\begin{aligned} T(u + v) = T((x_1, y_1) + (x_2, y_2)) &= T(x_1 + x_2, y_1 + y_2) \\ &= (2(x_1 + x_2), 0, (x_1 + x_2) + (y_1 + y_2)) \\ &= (2x_1 + 2x_2, 0, (x_1 + x_2) + (y_1 + y_2)) \\ &= (2x_1, 0, x_1 + y_1) + (2x_2, 0, x_2 + y_2) \\ &= T(u) + T(v) \end{aligned}$$

e

$$\begin{aligned} T(\alpha u) = T(\alpha(x_1, y_1)) = T(\alpha x_1, \alpha y_1) &= (2\alpha x_1, 0, \alpha x_1 + \alpha y_1) \\ &= \alpha(2x_1, 0, x_1 + y_1) \\ &= \alpha T(u). \end{aligned}$$

Exemplo 1.3. A aplicação

$$T : M_2 \rightarrow M_2$$

$$T \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \begin{pmatrix} a+b & b+c \\ c+d & d+a \end{pmatrix}$$

também é uma aplicação linear, pois, se $u = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$, $v = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ e $\alpha \in \mathbb{R}$, então

$$\begin{aligned} T \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right) &= T \left(\begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} a_1 + a_2 + b_1 + b_2 & b_1 + b_2 + c_1 + c_2 \\ c_1 + c_2 + d_1 + d_2 & d_1 + d_2 + a_1 + a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1 + b_1 & b_1 + c_1 \\ c_1 + d_1 & d_1 + a_1 \end{pmatrix} + \begin{pmatrix} a_2 + b_2 & b_2 + c_2 \\ c_2 + d_2 & d_2 + a_2 \end{pmatrix} \\ &= T \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \right) + T \left(\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right) \end{aligned}$$

e

$$\begin{aligned} T \left(\alpha \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \right) &= T \left(\begin{pmatrix} 49\alpha a_1 & \alpha b_1 \\ \alpha c_1 & \alpha d_1 \end{pmatrix} \right) = \begin{pmatrix} \alpha a_1 + \alpha b_1 & \alpha b_1 + \alpha c_1 \\ \alpha c_1 + \alpha d_1 & \alpha d_1 + \alpha a_1 \end{pmatrix} \\ &= \alpha \begin{pmatrix} a_1 + b_1 & b_1 + c_1 \\ c_1 + d_1 & d_1 + a_1 \end{pmatrix} \\ &= \alpha T \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \right). \end{aligned}$$

Exemplo 1.4. Seja $P(\mathbb{R})$ o espaço vetorial dos polinômios sobre \mathbb{R} e considere a função

$$\begin{aligned} D : P(\mathbb{R}) &\longrightarrow P(\mathbb{R}) \\ p(x) &\longmapsto p'(x) \end{aligned}$$

onde D é a função derivação restrita aos polinômios $p(x) \in P(\mathbb{R})$, $D(p(x)) = p'(x)$. Temos que D é uma transformação linear pois, se $p(x), q(x) \in P(\mathbb{R})$ e $\alpha \in \mathbb{R}$:

$$D((p+q)(x)) = (p+q)'(x) = p'(x) + q'(x) = D(p(x)) + D(q(x))$$

e

$$D((\alpha p)(x)) = (\alpha p)'(x) = \alpha p'(x) = \alpha D(p(x)).$$

Exemplo 1.5. Seja:

$$\begin{aligned} T : M_n &\longrightarrow \mathbb{R} \\ A &\longmapsto T(A) = \det(A). \end{aligned}$$

Esta aplicação não é uma transformação linear, pois em geral

$$\det(A_1 + A_2) \neq \det(A_1) + \det(A_2).$$

De fato, considerando as matrizes A_1 , A_2 e $A_1 + A_2$ (por exemplo), tais quais:

$$\begin{aligned} A_1 &= \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \\ A_2 &= \begin{pmatrix} -3 & 0 \\ 5 & 2 \end{pmatrix}, \\ A_1 + A_2 &= \begin{pmatrix} -2 & 2 \\ 3 & 3 \end{pmatrix}. \end{aligned}$$

Temos que $\det(A_1) = 5$, $\det(A_2) = -6$ e $\det(A_1 + A_2) = -12$. No entanto, $\det(A_1) + \det(A_2) = 5 + (-6) = -1 \neq -12 = \det(A_1 + A_2)$.

Exemplo 1.6. Sejam E e F espaços vetoriais. A função O , que leva todo vetor de E no vetor nulo de F , ou seja, $O(v) = \vec{0}$, para todo $v \in E$, é uma transformação linear e é chamada a transformação linear nula. De fato, se v_1 e v_2 são vetores quaisquer de E , temos que $O(v_1) + O(v_2) = \vec{0} + \vec{0} = \vec{0} = O(v_1 + v_2)$. Ademais, se $\alpha \in \mathbb{R}$, então $\alpha O(v_1) = \alpha \vec{0} = \vec{0} = O(\alpha v_1)$.

Exemplo 1.7. A transformação identidade, I , de E em E que leva todo vetor de E nele mesmo, ou seja, $I(v) = v$, para todo $v \in E$ é uma transformação linear. Pois, se considerarmos $v_1, v_2 \in E$ e $\alpha \in \mathbb{R}$, $I(v_1) + I(v_2) = v_1 + v_2 = I(v_1 + v_2)$ e $\alpha I(v_1) = \alpha v_1 = I(\alpha v_1)$.

1.1.1 PROPRIEDADES

Proposição 1.8. Toda transformação linear $T : E \rightarrow F$ leva o vetor nulo de E no vetor nulo de F .

Demonstração. Se x é um vetor qualquer de E , então

$$T(\vec{0}) = T(0 \cdot x) = 0 \cdot T(x) = \vec{0}.$$

□

Definição 1.9. Sejam $S : E \rightarrow F$ e $T : E \rightarrow F$ transformações lineares. Para todo $v \in E$ e $\alpha \in \mathbb{R}$,

(i) Definimos a soma de S e T , denotada por $S + T$, como a função $S + T : E \rightarrow F$ dada por

$$(S + T)(v) = S(v) + T(v); \tag{1.3}$$

(ii) Definimos o produto de α por T , denotado por αT , como a função $\alpha T : E \rightarrow F$ dada por

$$(\alpha T)(v) = \alpha T(v). \tag{1.4}$$

Teorema 1.10. *A soma de duas transformações lineares $S, T : E \rightarrow F$ e o produto de uma transformação linear $T : E \rightarrow F$ por um número $\alpha \in \mathbb{R}$ são transformações lineares.*

Demonstração. Sejam as transformações lineares $S, T : E \rightarrow F$, e os escalares $\alpha, \beta \in \mathbb{R}$, com as operações de soma de duas transformações lineares e produto de uma transformação linear por um escalar, definidas respectivamente em (1.3) e (1.4) na Definição 1.9.

(a) Mostrando que $S + T : E \rightarrow F$ é uma transformação linear:

$$\begin{aligned} (S + T)(v_1) + (S + T)(v_2) &= S(v_1) + T(v_1) + S(v_2) + T(v_2) \\ &= S(v_1 + v_2) + T(v_1 + v_2) \\ &= (S + T)(v_1 + v_2) \end{aligned}$$

e,

$$\begin{aligned} \beta(S + T)(v_1) &= \beta S(v_1) + \beta T(v_1) \\ &= S(\beta v_1) + T(\beta v_1) \\ &= (S + T)(\beta v_1). \end{aligned}$$

(b) Mostrando que $\alpha T : E \rightarrow F$ é uma transformação linear:

$$\begin{aligned} (\alpha T)(v_1) + (\alpha T)(v_2) &= \alpha T(v_1) + \alpha T(v_2) \\ &= \alpha(T(v_1) + T(v_2)) \\ &= \alpha(T(v_1 + v_2)) \\ &= (\alpha T)(v_1 + v_2) \end{aligned}$$

e,

$$\begin{aligned} \beta(\alpha T)(v_1) = \beta\alpha T(v_1) &= (\beta\alpha)T(v_1) \\ &= (\alpha\beta)T(v_1) \\ &= \alpha(\beta T)(v_1) \\ &= \alpha T(\beta v_1) \\ &= (\alpha T)(\beta v_1). \end{aligned}$$

□

Denotemos $\mathcal{L}(E; F)$ o conjunto das transformações lineares de E em F . Quando $E = F$, usaremos a notação $\mathcal{L}(E)$ em vez de $\mathcal{L}(E; E)$. As transformações lineares $T : E \rightarrow E$ do espaço vetorial E em si mesmo são chamados *operadores lineares* em E . Por sua vez, as transformações lineares $\varphi : E \rightarrow \mathbb{R}$, com valores numéricos, são chamadas *funcionais lineares*. Escreve-se E^* em vez de $\mathcal{L}(E; \mathbb{R})$ e o conjunto E^* dos funcionais lineares $\varphi : E \rightarrow \mathbb{R}$ chama-se o espaço vetorial *dual* de E .

Corolário 1.11. $\mathcal{L}(E; F)$ é um Espaço Vetorial.

Demonstração. Pelo Teorema 1.10 temos que a soma de duas transformações lineares $S, T : E \rightarrow F$ e o produto de uma transformação linear $T : E \rightarrow F$ por um número $\alpha \in \mathbb{R}$ também são transformações lineares, o que nos mostra que $\mathcal{L}(E; F)$ é um subespaço vetorial e consequentemente $\mathcal{L}(E; F)$ é um Espaço Vetorial. \square

Uma função $T : E \rightarrow F$ é uma transformação linear se, e somente se, $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$, para todos os vetores $v_1, v_2 \in E$ e todos os escalares α e β . Pois, se T é linear, então $T(\alpha v_1 + \beta v_2) = T(\alpha v_1) + T(\beta v_2) = \alpha T(v_1) + \beta T(v_2)$. Por outro lado, se T é uma função tal que $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$, para todos os vetores $v_1, v_2 \in E$ e todos os escalares α e β , então fazendo $\alpha = 1, \beta = 1$ e depois $\beta = 0$ segue-se que T é uma transformação linear. Mais geralmente, $T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 T(v_1) + \dots + \alpha_n T(v_n), \forall \alpha_i \in \mathbb{R}$ e $v_i \in E$.

Teorema 1.12. *Sejam E, F espaços vetoriais e B uma base de E . A cada vetor $u \in B$, façamos corresponder (de maneira arbitrária) um vetor $u' \in F$. Então existe uma única transformação linear $T : E \rightarrow F$ tal que $T(u) = u'$ para cada $u \in B$.*

Demonstração. Para cada vetor $v \in E$ temos, de modo único, uma combinação linear $v = \alpha_1 u_1 + \dots + \alpha_m u_m$ de elementos u_1, \dots, u_m da base B . Definimos $T : E \rightarrow F$ pondo

$$T(v) = \alpha_1 u'_1 + \dots + \alpha_m u'_m.$$

Dados $v, w \in E$ temos

$$v = \alpha_1 u_1 + \dots + \alpha_m u_m$$

e

$$w = \beta_1 u_1 + \dots + \beta_m u_m$$

(mesmo que a base B seja infinita, podemos exprimir v e w como combinações lineares dos mesmos elementos de B , completando com coeficientes zero os múltiplos dos u_i que aparecem apenas numa das duas expressões). Então

$$v + w = (\alpha_1 u_1 + \dots + \alpha_m u_m) + (\beta_1 u_1 + \dots + \beta_m u_m) = \sum_{i=1}^m (\alpha_i + \beta_i) u_i.$$

Logo

$$T(v + w) = \sum_{i=1}^m (\alpha_i + \beta_i) u'_i = \sum_{i=1}^m \alpha_i u'_i + \sum_{i=1}^m \beta_i u'_i = T(v) + T(w).$$

De maneira análoga se vê que $T(\alpha v) = \alpha T(v)$, portanto $T : E \rightarrow F$, assim definida, é uma transformação linear, tal que $T(u) = u'$, para todo $u \in B$. Quanto à unicidade, seja $S : E \rightarrow F$

outra transformação linear tal que $S(u) = u'$ para todo $u \in B$. Então, para cada $v = \sum_{i=1}^m \alpha_i u_i \in E$ tem-se

$$S(v) = S\left(\sum_{i=1}^m \alpha_i u_i\right) = \sum_{i=1}^m \alpha_i \cdot S(u_i) = \sum_{i=1}^m \alpha_i \cdot u'_i = T(v).$$

Portanto $S = T$. Isto completa a demonstração. \square

Em outras palavras, o Teorema 1.12 garante que uma transformação linear é totalmente determinada se conhecermos seus valores nos vetores de uma base de seu domínio.

Exemplo 1.13. Vamos determinar a transformação linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que $T(1, 1) = (0, 2, 1)$ e $T(0, 2) = (1, 0, 1)$.

Como $\alpha = \{(1, 1), (0, 2)\}$ é linearmente independente e possui dimensão 2, temos que α é uma base de \mathbb{R}^2 . Além disso, se $(x, y) \in \mathbb{R}^2$, então

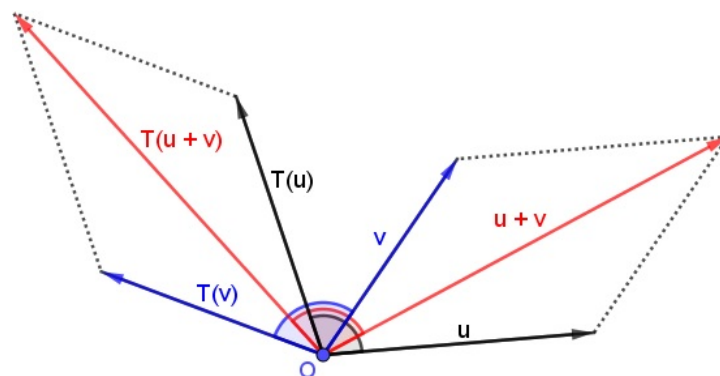
$$(x, y) = a_1(1, 1) + a_2(0, 2) = (a_1, a_1 + 2a_2)$$

se, e somente se, $a_1 = x$ e $a_2 = \frac{y-x}{2}$. Portanto, pelo Teorema 1.12,

$$\begin{aligned} T(x, y) &= a_1 T(1, 1) + a_2 T(0, 2) \\ &= x T(1, 1) + \left(\frac{y-x}{2}\right) T(0, 2) \\ &= x(0, 2, 1) + \left(\frac{y-x}{2}\right) (1, 0, 1) \\ &= \left(\frac{y-x}{2}, 2x, \frac{x+y}{2}\right). \end{aligned}$$

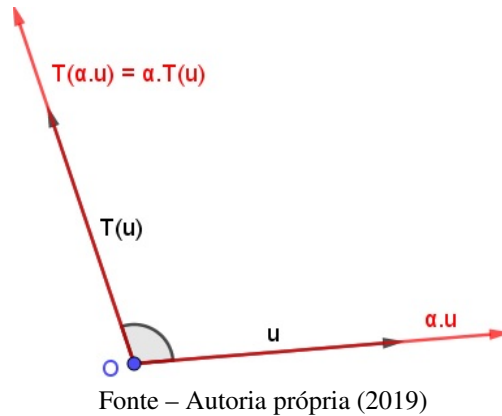
Exemplo 1.14. (Rotação de ângulo θ em torno da origem em \mathbb{R}^2). Trata-se de uma transformação linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, que leva cada vetor u no vetor $T(u)$ que dele resulta pela rotação de ângulo θ em torno da origem. As Figuras 1 e 2 mostram que $T(u+v) = T(u) + T(v)$ e $T(\alpha u) = \alpha T(u)$ para $u \in \mathbb{R}^2$ e $\alpha \in \mathbb{R}$, o que nos indica que T é uma transformação linear.

Figura 1 – Rotação de vetores I



Fonte – Autoria própria (2019)

Figura 2 – Rotação de vetores II



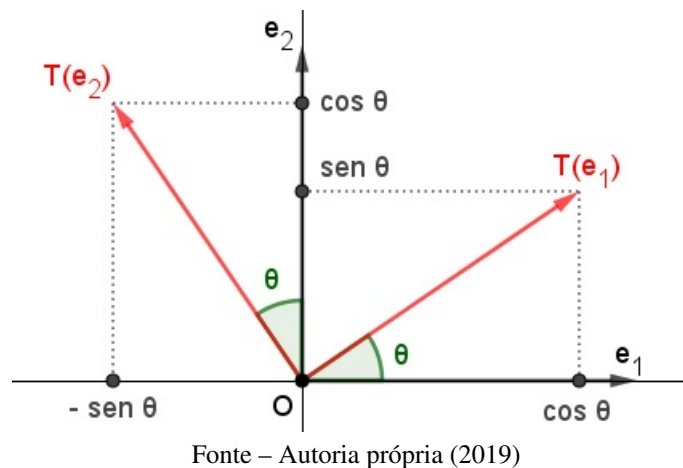
Considerando a base canônica de \mathbb{R}^2 formada pelos vetores unitários $e_1 = (1, 0)$ e $e_2 = (0, 1)$ e as definições de seno e cosseno, o vetor unitário $T(e_1)$, que forma com e_1 um ângulo θ , tem coordenadas $\cos \theta$ e $\sin \theta$, ou seja, $T(e_1) = (\cos \theta, \sin \theta)$. Além disso, como e_2 forma com e_1 um ângulo reto, temos que $T(e_2)$ também forma com $T(e_1)$ um ângulo reto. Logo $T(e_2) = (-\sin \theta, \cos \theta)$ (Figura 3). Utilizando o mesmo raciocínio do Exemplo 1.13, se $(x, y) \in \mathbb{R}^2$, então

$$(x, y) = a_1 e_1 + a_2 e_2 = a_1(1, 0) + a_2(0, 1) = (a_1, a_2).$$

Logo,

$$\begin{aligned} T(x, y) &= a_1 T(e_1) + a_2 T(e_2) \\ &= x(\cos \theta, \sin \theta) + y(-\sin \theta, \cos \theta) \\ &= (x \cos \theta, x \sin \theta) + (-y \sin \theta, y \cos \theta) \\ &= (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta). \end{aligned}$$

Figura 3 – Rotação de vetores III



1.2 PRODUTO DE TRANSFORMAÇÕES LINEARES

Definição 1.15. Dadas as transformações lineares $T : E \rightarrow F$, $S : F \rightarrow G$ (note que o domínio de S coincide com o contradomínio de T), o produto de S por T é a função $ST : E \rightarrow G$ pondo, para cada $v \in E$,

$$ST(v) = S(T(v)). \quad (1.5)$$

Observe que ST nada mais é do que a composta $S \circ T$ das funções S e T .

$$\begin{array}{ccc} E & \xrightarrow{T} & F & \xrightarrow{S} & G \\ & \searrow & & \nearrow & \\ & & & & \\ & \xrightarrow{ST} & & & \end{array} \quad (1.6)$$

Proposição 1.16. Se $T : E \rightarrow F$ e $S : F \rightarrow G$ são transformações lineares, então a composição $ST : E \rightarrow G$ é uma transformação linear.

Demonstração. Sejam $v_1, v_2 \in E$ e α, β escalares

$$\begin{aligned} (ST)(\alpha v_1 + \beta v_2) &= S(T(\alpha v_1 + \beta v_2)) \\ &= S(\alpha T(v_1) + \beta T(v_2)) \\ &= \alpha S(T(v_1)) + \beta S(T(v_2)) \\ &= \alpha (ST)(v_1) + \beta (ST)(v_2). \end{aligned}$$

□

Proposição 1.17. Sejam S , T e U transformações lineares com domínios e contradomínios apropriados. Então

(a) **Associatividade:** $S(TU) = (ST)U$;

(b) **Distributividade à direita:** $S(T + U) = ST + SU$;

(c) **Distributividade à esquerda:** $(S + T)U = SU + TU$;

(d) **Homogeneidade:** $\alpha(TU) = (\alpha T)U = T(\alpha U)$, para qualquer $\alpha \in \mathbb{R}$;

(e) $IT = TI = T$, onde I é a transformação identidade.

Demonstração. Seja v pertencente ao domínio de U e $\alpha \in \mathbb{R}$.

(a) $S(TU)(v) = S((TU)(v)) = S(T(U(v))) = ST((U(v)))$;

(b) Pela linearidade de S temos que $S(T + U)(v) = S((T + U)(v)) = S(T(v) + U(v)) = S(T(v)) + S(U(v)) = ST(v) + SU(v) = (ST + SU)(v)$;

- (c) Da definição de $S + T$ temos: $(S + T)U(v) = S(U(v)) + T(U(v)) = SU(v) + TU(v)$;
- (d) (1ª parte) $\alpha(TU)(v) = \alpha T(U(v)) = (\alpha T)(U(v)) = (\alpha T)U(v)$;
- (d) (2ª parte) $\alpha(TU)(v) = \alpha T(U(v)) = T(\alpha U(v)) = T(\alpha U)(v)$;
- (e) Seja I a transformação identidade, temos que $IT(v) = I(T(v)) = T(v) = (T(v))I = T(v)I = TI(v)$.

□

Observe que o produto ST só está definido quando T toma valores no domínio de S . Esta restrição desaparece, naturalmente, quando se trata de operadores lineares no mesmo espaço E : então o produto ST está definido quaisquer que sejam $T_1, T_2 \in \mathcal{L}(E)$. Neste contexto, há diferenças notáveis entre o produto de transformações lineares e o produto de números reais, como as ausências da comutatividade, da lei do corte e da inversa multiplicativa para uma transformação $\neq \vec{0}$, além da presença de transformações nilpotentes, para as quais tem-se $T^n = \vec{0}$ com $T_1 \neq \vec{0}$.

1.3 NÚCLEO E IMAGEM

Definição 1.18. *Sejam E e F espaços vetoriais. Seja $T : E \rightarrow F$ uma transformação linear.*

(a) O **núcleo** de T é o conjunto definido como:

$$\mathcal{N}(T) = \{v \in E \mid T(v) = \vec{0}\}. \quad (1.7)$$

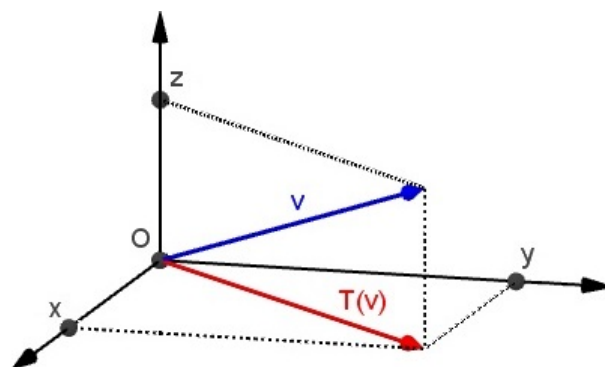
(b) A **imagem** de T é o conjunto definido como:

$$\mathcal{I}m(T) = \{w \in F \mid w = T(v), \text{ para algum } v \in E\}. \quad (1.8)$$

Exemplo 1.19. *Sejam $T : E \rightarrow F$ a transformação linear nula e $I : E \rightarrow E$ a transformação identidade.*

- (a) $\mathcal{N}(T) = E$ e $\mathcal{I}m(T) = \vec{0}$, pois pela definição de transformação linear nula, para qualquer $v \in E$, temos que $T(v) = \vec{0}$, o que implica que o único elemento de $\mathcal{I}m(T)$ é o vetor nulo e todo elemento de E pertence ao $\mathcal{N}(T)$;
- (b) $\mathcal{N}(I) = \vec{0}$ e $\mathcal{I}m(I) = E$, pois pela definição de transformação identidade, $T(v) = \vec{0}$ somente quando $v = \vec{0}$ onde $v \in E$, da mesma forma que $T(v) = v$, onde $v \in E$.

Figura 4 – $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$; Projeção ortogonal sobre plano xy



Fonte – Autoria própria (2019)

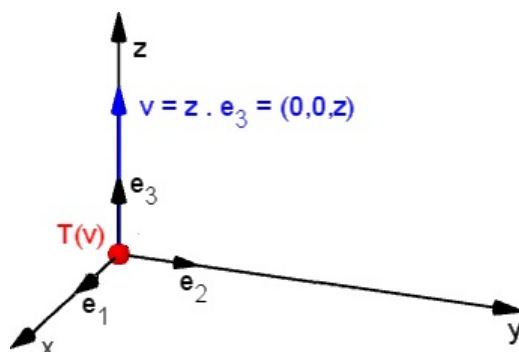
Exemplo 1.20. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a projeção ortogonal sobre o plano xy (Figura 4).

Considerando a base canônica de \mathbb{R}^3 , $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ perceba que a transformação $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida como projeção ortogonal sobre o plano xy , sabendo que $T(\vec{e}_1) = \vec{e}_1$, $T(\vec{e}_2) = \vec{e}_2$ e $T(\vec{e}_3) = \vec{0}$, algebricamente pode ser descrita como

$$\begin{aligned} T(x, y, z) &= T(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3) \\ &= xT(\vec{e}_1) + yT(\vec{e}_2) + zT(\vec{e}_3) \\ &= x\vec{e}_1 + y\vec{e}_2 + \vec{0} \\ &= (x, y, 0). \end{aligned}$$

Se $T(x, y, z) = (0, 0, 0) \Rightarrow (x, y, z) = (0, 0, 0) \Rightarrow x = 0$ e $y = 0$. Como nada é dito sobre a variável z , temos que z é qualquer, logo $\mathcal{N}(T) = \{(0, 0, z) \in \mathbb{R}^3 \mid z \in \mathbb{R}\}$, ou seja o núcleo de T são todos os vetores que estão sobre o eixo z (Figura 5).

Figura 5 – $\mathcal{N}(T) = \{(0, 0, z) \in \mathbb{R}^3 \mid z \in \mathbb{R}\}$



Fonte – Autoria própria (2019)

A proposição a seguir mostra como podemos determinar geradores para a imagem de uma transformação linear.

Proposição 1.21. *Seja $T : E \rightarrow F$ uma transformação linear. Se $\{v_1, \dots, v_n\}$ é uma base de E , então a imagem de T é gerada por $T(v_1), \dots, T(v_n)$.*

Demonstração. Seja $w \in \mathcal{I}m(T)$. Então, existe $v \in E$ tal que $T(v) = w$. Como $\{v_1, \dots, v_n\}$ é uma base de E , existem escalares $\alpha_1, \dots, \alpha_n$ tais que $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Assim,

$$\begin{aligned} w = T(v) &= T(\alpha_1 v_1 + \dots + \alpha_n v_n) = T(\alpha_1 v_1) + \dots + T(\alpha_n v_n) \\ &= \alpha_1 T(v_1) + \dots + \alpha_n T(v_n). \end{aligned}$$

O que mostra que $T(v_1), \dots, T(v_n)$ geram $\mathcal{I}m(T)$. □

Exemplo 1.22. *Seja a transformação linear $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $T(x, y, z) = (2x - y - z, x - y - z, x + y - z)$. Se $w \in \mathcal{I}m(T)$ então $w = T(x, y, z)$, ou seja,*

$$\begin{aligned} w &= (2x - y - z, x - y - z, x + y - z) \\ &= x(2, 1, 1) + y(-1, -1, 1) + z(-1, -1, -1). \end{aligned}$$

Logo todo vetor que pertence a imagem de T é gerado pelos vetores $w_1 = (2, 1, 1)$, $w_2 = (-1, -1, 1)$ e $w_3 = (-1, -1, -1)$. Podemos então escrever que

$$\mathcal{I}m(T) = [(2, 1, 1), (-1, -1, 1), (-1, -1, -1)].$$

Exemplo 1.23. *Considere a transformação linear*

$$\begin{aligned} T &: \mathbb{R}^3 \rightarrow M_2 \\ T(a, b, c) &= \begin{pmatrix} a + b & 0 \\ 0 & c - b \end{pmatrix}. \end{aligned}$$

Um elemento $(a, b, c) \in \mathbb{R}^3$ pertence ao núcleo de T se $T(a, b, c) = 0$. Então,

$$\begin{pmatrix} a + b & 0 \\ 0 & c - b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

se, e somente se, $a = -b$ e $c = b$. Portanto,

$$\mathcal{N}(T) = \{(-b, b, b) / b \in \mathbb{R}\}.$$

Por outro lado, a imagem de T é formada pelas matrizes de M_2 da forma

$$\begin{aligned} \begin{pmatrix} a + b & 0 \\ 0 & c - b \end{pmatrix} &= \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix} \\ &= a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

com $a, b, c \in \mathbb{R}$. Portanto,

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

é um conjunto gerador de $\mathcal{I}m(T)$.

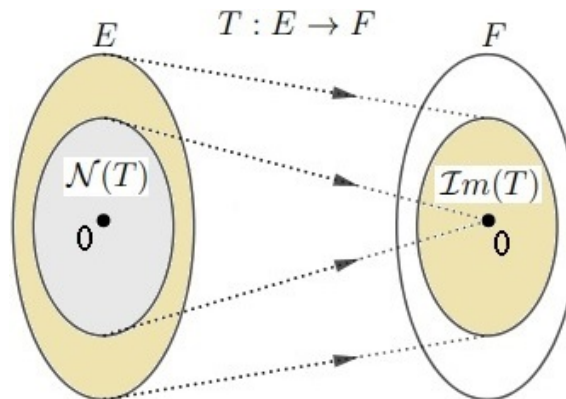
Teorema 1.24. *Seja $T : E \rightarrow F$ uma transformação linear. Então:*

- (i) $\mathcal{I}m(T)$ é um subespaço vetorial de F ;
- (ii) $\mathcal{N}(T)$ é um subespaço vetorial de E .

Demonstração. De fato, se $u_1, u_2 \in \mathcal{I}m(T)$ e $\alpha \in \mathbb{R}$, temos $u_1 = T(v_1)$ e $u_2 = T(v_2)$, onde $v_1, v_2 \in E$. Logo, $u_1 + u_2 = T(v_1) + T(v_2) = T(v_1 + v_2) \in \mathcal{I}m(T)$ e $\alpha u_1 = \alpha T(v_1) = T(\alpha v_1) \in \mathcal{I}m(T)$, o que mostra que $\mathcal{I}m(T)$ é um subespaço vetorial de F , pois $v_1 + v_2, \alpha v_1 \in E$ que é um espaço vetorial (provando (i)).

Por outro lado, se $v_1, v_2 \in \mathcal{N}(T)$ e $\alpha \in \mathbb{R}$, logo $T(v_1) = T(v_2) = \vec{0}$. Assim, $T(v_1 + v_2) = T(v_1) + T(v_2) = \vec{0} + \vec{0} = \vec{0}$ e $T(\alpha v_1) = \alpha T(v_1) = \alpha \vec{0} = \vec{0}$, o que significa dizer que $v_1 + v_2, \alpha v_1 \in \mathcal{N}(T)$ e conseqüentemente $\mathcal{N}(T)$ é um subespaço vetorial de E (provando (ii)). \square

Figura 6 – Núcleo e imagem de uma transformação linear $T : E \rightarrow F$



Fonte – Autoria própria (2019)

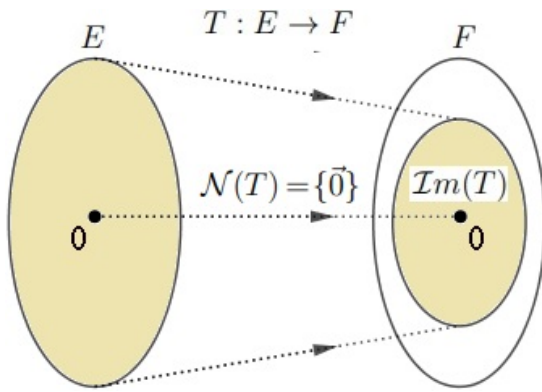
Definição 1.25. *Seja $T : E \rightarrow F$, quando $\mathcal{I}m(T) = F$, diz-se que a transformação T é **sobrejetiva**. Isto significa que, para qualquer $w \in F$ dado, pode-se achar $v \in E$ tal que $T(v) = w$.*

Definição 1.26. *Uma transformação linear $T : E \rightarrow F$ é dita **injetiva** quando tem-se $u \neq v$ em $E \Rightarrow T(u) \neq T(v)$ em F . Equivalentemente: $T(u) = T(v) \Rightarrow u = v$.*

Teorema 1.27. *A fim de que uma transformação linear $T : E \rightarrow F$ seja injetiva é necessário e suficiente que seu núcleo $\mathcal{N}(T)$ contenha apenas o vetor nulo.*

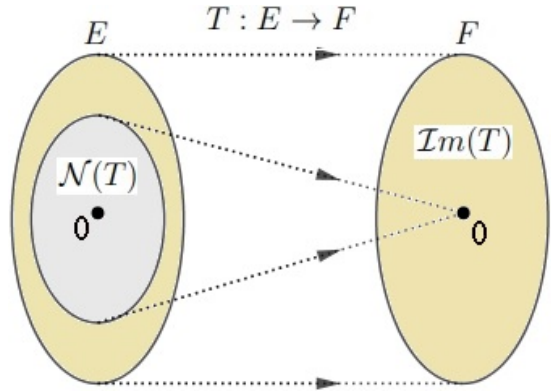
Demonstração. Seja T injetiva. Então $v \in \mathcal{N}(T) \Rightarrow T(v) = \vec{0} = T(\vec{0}) \Rightarrow v = \vec{0}$, logo $\mathcal{N}(T) = \{\vec{0}\}$. Reciprocamente, seja $\mathcal{N}(T) = \{\vec{0}\}$. Então $T(u) = T(v) \Rightarrow T(u - v) = T(u) - T(v) = \vec{0} \Rightarrow u - v \in \mathcal{N}(T) \Rightarrow u - v = \vec{0} \Rightarrow u = v$, logo T é injetiva. \square

Figura 7 – Transformação injetiva



Fonte – Autoria própria (2019)

Figura 8 – Transformação sobrejetiva



Fonte – Autoria própria (2019)

Exemplo 1.28. A transformação linear $T : \mathbb{R}^3 \rightarrow M_2$, onde $T(a, b, c) = \begin{pmatrix} a + b & 0 \\ 0 & c - b \end{pmatrix}$ do Exemplo 1.23 não é injetiva, pois

$$\mathcal{N}(T) = \{(-b, b, b) / b \in \mathbb{R}\} \neq \{(0, 0, 0)\}.$$

Já a transformação linear $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(x, y) = (x - y, x + y)$, é injetiva, pois

$$(x - y, x + y) = (0, 0) \Leftrightarrow (x, y) = (0, 0)$$

e, portanto, $\mathcal{N}(S) = \{(0, 0)\}$.

Teorema 1.29. Uma transformação linear é injetiva se, e somente se, leva vetores linearmente independentes em vetores linearmente independentes.

Demonstração. Seja $T : E \rightarrow F$ uma transformação linear injetiva. Se os vetores $v_1, \dots, v_n \in E$ são linearmente independentes (LI), vamos provar que suas imagens $T(v_1), \dots, T(v_n)$ são vetores linearmente independentes em F . Com efeito,

$$\begin{aligned} \alpha_1 T(v_1) + \dots + \alpha_n T(v_n) = \vec{0} &\Rightarrow T(\alpha_1 v_1) + \dots + T(\alpha_n v_n) = \vec{0} \\ &\Rightarrow T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \vec{0} \\ &\Rightarrow \alpha_1 v_1 + \dots + \alpha_n v_n = \vec{0}, \end{aligned}$$

pois T é injetiva. Como v_1, \dots, v_n são LI, segue-se que $\alpha_1 = \dots = \alpha_n = \vec{0}$, portanto $T(v_1), \dots, T(v_n)$ são LI.

Por outro lado, se a transformação linear $T : E \rightarrow F$ leva vetores LI em vetores LI então, de qualquer conjunto LI de vetores $v \neq \vec{0}$ em E , temos um conjunto de vetores correspondente $T(v) \neq \vec{0}$ em F também LI, o que significa dizer que $\mathcal{N}(T) = \{\vec{0}\}$ e portanto, pelo Teorema 1.27, T é injetiva. \square

Corolário 1.30. *Seja $T : E \rightarrow F$ uma transformação linear. Seja $\{v_1, \dots, v_n\}$ uma base de E , T é sobrejetiva se, e somente se, $T(v_1), \dots, T(v_n)$ geram F .*

Demonstração. Como $\{v_1, \dots, v_n\}$ é base de E , pela Proposição 1.21 temos que $T(v_1), \dots, T(v_n)$ geram $\mathcal{I}m(T)$. Mas T é sobrejetiva, logo $T(v_1), \dots, T(v_n)$ geram F . Por outro lado, se $T(v_1), \dots, T(v_n)$ geram F , temos que qualquer $w \in F$ é escrito como combinação linear de $T(v_1), \dots, T(v_n)$. Como, por hipótese, $\{v_1, \dots, v_n\}$ é base de E , temos que T é sobrejetiva. \square

Segue-se do Teorema 1.29 que se E tem dimensão finita n e $T : E \rightarrow F$ é uma transformação linear injetiva então $\dim F \geq n$, pois, considerando uma base $\{(v_1, \dots, v_n)\}$ de E , logo pelo Teorema 1.29 e pela Proposição 1.21 temos que $\{(T(v_1), \dots, T(v_n))\}$ é linearmente independente, mostrando conseqüentemente a impossibilidade de ocorrer em uma transformação linear injetiva $T : E \rightarrow F$, que a $\dim F < n$. Por outro lado, do Corolário 1.30 temos que se E tem dimensão finita n e $T : E \rightarrow F$ é uma transformação linear sobrejetiva então $\dim F \leq n$.

Neste contexto, por exemplo, não existe uma transformação linear injetiva de \mathbb{R}^3 em \mathbb{R}^2 , assim como também não existe uma transformação linear sobrejetiva de \mathbb{R}^2 em \mathbb{R}^3 .

Exemplo 1.31. Considerando a transformação linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, onde $T(x, y) = (-y, x, x + y)$, temos que:

(i) é injetiva, pois

$$(-y, x, x + y) = (0, 0, 0) \Leftrightarrow (x, y) = (0, 0) \Rightarrow \mathcal{N}(T) = \{(0, 0)\}.$$

(ii) não é sobrejetiva, pois

$$(-y, x, x + y) = (0, x, x) + (-y, 0, y) = x(0, 1, 1) + y(-1, 0, 1)$$

com $x, y \in \mathbb{R}$. O que significa dizer que $\alpha = \{(0, 1, 1), (-1, 0, 1)\}$ é um conjunto gerador de $\mathcal{I}m(T)$. Ademais, como os vetores $(0, 1, 1)$ e $(-1, 0, 1)$ são *LI*, temos que α é base e conseqüentemente $\dim(\mathcal{I}m(T)) = 2 < 3 = \dim(\mathbb{R}^3)$.

Exemplo 1.32. Considerando a transformação linear $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $T(x, y, z) = (x + y, y + z)$, temos que:

(i) não é injetiva, pois

$$(x + y, y + z) = (0, 0) \Leftrightarrow (x, y, z) = (-y, y, -y) \Rightarrow \mathcal{N}(T) \neq \{(0, 0, 0)\}.$$

(ii) é sobrejetiva, pois

$$(x + y, y + z) = (x, 0) + (y, y) + (0, z) = x(1, 0) + y(1, 1) + z(0, 1)$$

com $x, y, z \in \mathbb{R}$. O que significa dizer que $\alpha = \{(1, 0), (1, 1), (0, 1)\}$ é um conjunto gerador de $\mathcal{I}m(T)$. Porém, temos que o vetor $(1, 1) = (1, 0) + (0, 1)$. Logo, os vetores $(1, 0)$ e $(0, 1)$ são *LI* formando uma base para \mathbb{R}^2 (base canônica). Logo, $\dim(\mathcal{I}m(T)) = \dim(\mathbb{R}^2)$.

Teorema 1.33. (*Dimensão do núcleo e da imagem*). *Sejam E, F espaços vetoriais de dimensão finita. Para toda transformação linear $T : E \rightarrow F$ tem-se*

$$\dim(\mathcal{N}(T)) + \dim(\mathcal{I}m(T)) = \dim(E). \quad (1.9)$$

Demonstração. Sejam E, F espaços vetoriais de dimensão finita e $T : E \rightarrow F$ onde $\mathcal{N}(T) \neq \{\vec{0}\}$ e $\mathcal{I}m(T) \neq \{\vec{0}\}$.

Dividiremos esta demonstração em duas etapas:

- (a) Vamos mostrar inicialmente que se $\{T(u_1), \dots, T(u_p)\}$ é uma base de $\mathcal{I}m(T)$ e $\{v_1, \dots, v_q\}$ é uma base do $\mathcal{N}(T)$ então $\{u_1, \dots, u_p, v_1, \dots, v_q\}$ é uma base de E .

Com efeito, em primeiro lugar, se tivermos

$$\alpha_1 u_1 + \dots + \alpha_p u_p + \beta_1 v_1 + \dots + \beta_q v_q = \vec{0}$$

então, aplicando o operador T a ambos os membros desta igualdade e lembrando que v_1, \dots, v_q pertencem ao núcleo de T , obtemos

$$\begin{aligned} T(\alpha_1 u_1 + \dots + \alpha_p u_p + \beta_1 v_1 + \dots + \beta_q v_q) &= T(\vec{0}) \\ \alpha_1 T(u_1) + \dots + \alpha_p T(u_p) + \beta_1 T(v_1) + \dots + \beta_q T(v_q) &= \vec{0} \\ \alpha_1 T(u_1) + \dots + \alpha_p T(u_p) &= \vec{0}. \end{aligned}$$

Como os vetores $T(u_1), \dots, T(u_p)$ são *LI*, resulta daí que $\alpha_1 = \dots = \alpha_p = \vec{0}$. Portanto a igualdade $\alpha_1 u_1 + \dots + \alpha_p u_p + \beta_1 v_1 + \dots + \beta_q v_q = \vec{0}$ se reduz a

$$\beta_1 v_1 + \dots + \beta_q v_q = \vec{0}.$$

Como v_1, \dots, v_q são *LI*, concluímos que $\beta_1 = \dots = \beta_q = \vec{0}$. Isto mostra que os vetores $u_1, \dots, u_p, v_1, \dots, v_q$ são *LI*.

Em seguida, consideremos um vetor arbitrário $w \in E$. Como $T(w) \in \mathcal{I}m(T)$ e ademais $\{T(u_1), \dots, T(u_p)\}$ é uma base da imagem de T , temos que

$$\begin{aligned} T(w) &= \alpha_1 T(u_1) + \dots + \alpha_p T(u_p) \\ T(w) - (\alpha_1 T(u_1) + \dots + \alpha_p T(u_p)) &= \vec{0} \\ T(w - (\alpha_1 u_1 + \dots + \alpha_p u_p)) &= \vec{0}. \end{aligned}$$

Assim, o vetor $w - (\alpha_1 u_1 + \dots + \alpha_p u_p)$ pertence ao núcleo de T , logo pode ser expresso como combinação linear dos elementos de sua base $\{v_1, \dots, v_q\}$. Temos então

$$w - (\alpha_1 u_1 + \dots + \alpha_p u_p) = \beta_1 v_1 + \dots + \beta_q v_q,$$

ou seja, $w = \alpha_1 u_1 + \dots + \alpha_p u_p + \beta_1 v_1 + \dots + \beta_q v_q$. Isto mostra que os vetores $u_1, \dots, u_p, v_1, \dots, v_q$ geram E e portanto constituem uma base.

(b) Considerando $\{T(u_1), \dots, T(u_p)\}$ uma base de $\mathcal{I}m(T)$, $\{v_1, \dots, v_q\}$ uma base do $\mathcal{N}(T)$ e $\{u_1, \dots, u_p, v_1, \dots, v_q\}$ uma base de E (provado em (a)), temos que $\dim(\mathcal{I}m(T)) + \dim(\mathcal{N}(T)) = p + q = \dim(E)$, o que prova o resultado. □

Exemplo 1.34. Seja $O : E \rightarrow F$ uma transformação linear nula, onde $O(v) = \vec{0}$, para todo vetor $v \in E$. O teorema é verificado, pois

$$\begin{aligned} \mathcal{I}m(O) = \{\vec{0}\} &\Rightarrow \dim(\mathcal{I}m(O)) = 0 \quad \text{e} \\ \mathcal{N}(O) = E &\Rightarrow \dim(\mathcal{N}(O)) = \dim(E). \end{aligned}$$

Exemplo 1.35. Seja $T : E \rightarrow F$ uma transformação linear, onde $\mathcal{N}(T) = \{\vec{0}\}$. Perceba que se $\mathcal{N}(T) = \{\vec{0}\}$ então T é injetiva (Teorema 1.27). Além disso, pela Proposição 1.21 e pelo Teorema 1.29 temos que se $\{v_1, \dots, v_n\}$ é uma base de E e T é injetiva, logo, $\{T(v_1), \dots, T(v_n)\}$ é uma base de $\mathcal{I}m(T)$, Sendo assim, o resultado do teorema anterior também é verificado, pois, conclui-se que $\dim(\mathcal{I}m(T)) = \dim(E)$.

Exemplo 1.36. (Retomando o Exemplo 1.23) Considerando a transformação linear

$$\begin{aligned} T &: \mathbb{R}^3 \rightarrow M_2 \\ T(a, b, c) &= \begin{pmatrix} a + b & 0 \\ 0 & c - b \end{pmatrix}. \end{aligned}$$

Temos que se $\mathcal{N}(T) = \{(-b, b, b) / b \in \mathbb{R}\}$ então o conjunto $\{(-1, 1, 1)\}$ é uma base do $\mathcal{N}(T)$, e conseqüentemente $\dim(\mathcal{N}(T)) = 1$.

Por outro lado, temos que o conjunto gerador da imagem de T é dado por

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Como $\dim(\mathcal{N}(T)) = 1$ e $\dim(\mathbb{R}^3) = 3$, pelo Teorema da dimensão do núcleo e da imagem, temos que $\dim(\mathcal{I}m(T)) = 2$. Neste contexto, podemos concluir que β não é base para $\mathcal{I}m(T)$.

Definição 1.37. Se uma transformação linear $T : E \rightarrow F$ é injetiva e sobrejetiva, então é chamada **isomorfismo**. Neste caso, dizemos que os espaços vetoriais E e F são **isomorfos**.

Teorema 1.38. Sejam E e F espaços vetoriais de dimensão finita n . Uma transformação linear $T : E \rightarrow F$ é injetiva se, e somente se, é sobrejetiva e portanto é um isomorfismo.

Demonstração. Com efeito, do Teorema da dimensão do núcleo e da imagem (1.33), temos que $n = \dim(\mathcal{N}(T)) + \dim(\mathcal{I}m(T))$. Logo,

$$\mathcal{N}(T) = \{\vec{0}\} \quad (T \text{ injetiva}) \Leftrightarrow \dim(\mathcal{I}m(T)) = n \Leftrightarrow \mathcal{I}m(T) = F \quad (T \text{ sobrejetiva}).$$

□

Teorema 1.39. *Sejam E e F espaços vetoriais de dimensão finita. E é isomorfo a F se, e somente se, $\dim(E) = \dim(F)$.*

Demonstração. Diante do exposto no Teorema 1.38, basta mostrar que se $\dim(E) = \dim(F)$, então E e F são isomorfos.

Sejam $\alpha = \{u_1, \dots, u_n\}$ e $\beta = \{v_1, \dots, v_n\}$ bases de E e F , respectivamente. Pelo Teorema 1.12, existe uma única transformação linear $T : E \rightarrow F$, tal que $T(u_i) = v_i$, para $i = 1, \dots, n$. Pela Proposição 1.21, a imagem de T é gerada pelos vetores $T(u_1), \dots, T(u_n)$ e assim

$$\mathcal{I}m(T) = [T(u_1), \dots, T(u_n)] = [v_1, \dots, v_n] = F \quad (T \text{ é sobrejetiva}).$$

Pelo Teorema 1.33 (dimensão do núcleo e da imagem), T também é injetiva e portanto um isomorfismo. □

1.4 SOMA DIRETA E PROJEÇÃO

Definição 1.40. *Se W_1, W_2 são dois subespaços do espaço vetorial E . Dizemos que E é a soma de W_1 e W_2 se todo elemento $v \in E$ pode ser escrito como $v = w_1 + w_2$, onde $w_1 \in W_1$ e $w_2 \in W_2$, no caso denotamos $E = W_1 + W_2$. Quando $W_1 \cap W_2 = \{\vec{0}\}$, diremos que a soma $W_1 + W_2$ é **direta** e, neste caso, escrevemos $E = W_1 \oplus W_2$.*

Exemplo 1.41. Sejam $W_1 = \left\{ \begin{pmatrix} a & b \\ c & a \end{pmatrix}; a, b, c \in \mathbb{R} \right\}$ e $W_2 = \left\{ \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}; a \in \mathbb{R} \right\}$, dois subespaços vetoriais de M_2 .

$$1. \quad M_2 = W_1 + W_2. \text{ De fato, seja } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_2.$$

Sendo $A = B_1 + B_2$, com $B_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & a_1 \end{pmatrix} \in W_1$ e $B_2 = \begin{pmatrix} a_2 & 0 \\ 0 & -a_2 \end{pmatrix} \in W_2$, temos que

$$\begin{aligned} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \begin{pmatrix} a_1 & b_1 \\ c_1 & a_1 \end{pmatrix} + \begin{pmatrix} a_2 & 0 \\ 0 & -a_2 \end{pmatrix} \\ \Rightarrow \begin{cases} a_1 + a_2 &= a_{11} \\ b_1 &= a_{12} \\ c_1 &= a_{21} \\ a_1 - a_2 &= a_{22} \end{cases} &\Rightarrow \begin{cases} a_1 &= \frac{a_{11} + a_{22}}{2} \\ b_1 &= a_{12} \\ c_1 &= a_{21} \\ a_2 &= \frac{a_{11} - a_{22}}{2} \end{cases} \end{aligned}$$

e portanto,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{a_{11}+a_{22}}{2} & a_{12} \\ a_{21} & \frac{a_{11}+a_{22}}{2} \end{pmatrix}}_{\in W_1} + \underbrace{\begin{pmatrix} \frac{a_{11}-a_{22}}{2} & 0 \\ 0 & \frac{a_{11}-a_{22}}{2} \end{pmatrix}}_{\in W_2}.$$

2. $W_1 \cap W_2 = \{\vec{0}\}$. De fato, se $A \in W_1 \cap W_2$ então $A \in W_1$ e, portanto,

(i) $A \in W_1$ e, portanto, $a_{11} = a_{22}$;

(ii) $A \in W_2$ e, portanto, $a_{12} = a_{21} = 0$ e $a_{11} = -a_{22}$.

Deste modo, por (i) e (ii) tem-se que $a_{11} = a_{22} = 0$.

Exemplo 1.42. Sejam W_1 e W_2 subespaços vetoriais de \mathbb{R}^2 , gerados respectivamente pelos vetores $(1, 0)$ e $(1, 1)$. Observe que $\mathbb{R}^2 = W_1 \oplus W_2$. É claro que se $(a, b) \in \mathbb{R}^2$, então podemos escrever $(a, b) = (a-b)(1, 0) + b(1, 1)$. Por outro lado, $W_1 \cap W_2 = \{\vec{0}\}$, pois se $(a, b) \in W_1 \cap W_2$, então $(a, b) = c(1, 0) = d(1, 1)$, o que implica que $c = d = 0$.

Teorema 1.43. Sejam W_1 e W_2 subespaços de um espaço vetorial E . Temos que $E = W_1 \oplus W_2$ se, e somente se, todo vetor $v \in E$ se escreve, de modo único, como soma $v = w_1 + w_2$, onde $w_1 \in W_1$ e $w_2 \in W_2$.

Demonstração. Suponhamos que $E = W_1 \oplus W_2$. Logo, $W_1 \cap W_2 = \{\vec{0}\}$. Sejam $u_1, w_1 \in W_1$ e $u_2, w_2 \in W_2$ tais que

$$u_1 + u_2 = w_1 + w_2.$$

Somando-se $-w_1 - u_2$ em ambos os membros da igualdade acima, obtemos

$$\underbrace{u_1 - w_1}_{\in W_1} = \underbrace{w_2 - u_2}_{\in W_2} \Rightarrow u_1 - w_1 = w_2 - u_2 \in W_1 \cap W_2 = \{\vec{0}\},$$

o que implica que $u_1 = w_1$ e $u_2 = w_2$.

Por outro lado, suponhamos que todo elemento de $v \in E$ se escreve, de modo único, como soma $v = w_1 + w_2$, onde $w_1 \in W_1$ e $w_2 \in W_2$. Seja $v \in W_1 \cap W_2$. Vamos mostrar que $v = \vec{0}$. Mas,

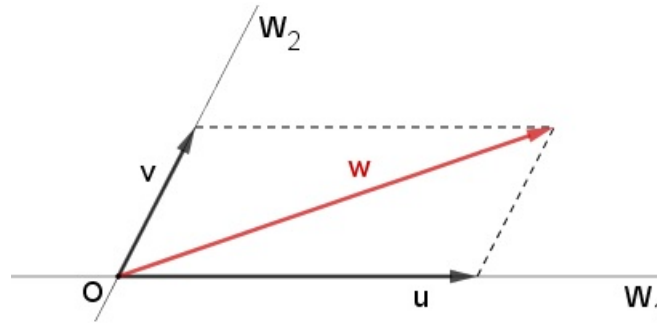
$$v = \underbrace{v}_{\in W_1} + \underbrace{\vec{0}}_{\in W_2} = \underbrace{v}_{\in W_2} + \underbrace{\vec{0}}_{\in W_1}.$$

Isto é, se $v \neq \vec{0}$, teríamos duas formas de escrever v como uma soma de um elemento de W_1 e um de W_2 . Logo, $v = \vec{0}$ e $W_1 \cap W_2 = \{\vec{0}\}$. Portanto, como claramente $E = W_1 + W_2$, temos que $E = W_1 \oplus W_2$.

□

A noção de soma direta está intimamente ligada à noção de projeção. Se $E = W_1 \oplus W_2$ é a decomposição do espaço vetorial E como soma direta dos subespaços W_1 e W_2 , define-se o operador linear $P : E \rightarrow E$, *projeção de E sobre W_1 , paralelamente a W_2* , do seguinte modo: todo vetor $w \in E$ se escreve, de modo único, como soma $w = u + v$ de um vetor $u \in W_1$ com um vetor $v \in W_2$. Põe-se, então, $P(w) = u$. Veja (Figura 9).

Figura 9 – Projeção de vetores



Fonte – Autoria própria (2019)

Observação 1: Perceba que o operador linear $P : E \rightarrow E$ definido acima é, de fato, uma transformação linear, pois sendo W_1, W_2 subespaços vetoriais de E , com $E = W_1 \oplus W_2$, e considerando $w_1, w_2 \in E$ tais quais $w_1 = u_1 + v_1$ e $w_2 = u_2 + v_2$, com $u_1, u_2 \in W_1$ e $v_1, v_2 \in W_2$. Então, temos que $P(w_1) + P(w_2) = u_1 + u_2 = P(w_1 + w_2)$ e $P(\alpha w_1) = \alpha u_1 = \alpha P(w_1)$.

Observação 2: O operador linear $P : E \rightarrow E$ assim definido tem imagem W_1 e núcleo W_2 . Além disso, temos que sobre P vale a igualdade $P^2 = P$, pois, se $w \in E$, onde $w = u + v$, sendo $u \in W_1$ e $v \in W_2$, logo

$$P^2(w) = (P \circ P)(w) = P(P(w)) = P(u) = u = P(w). \quad (1.10)$$

O teorema a seguir (1.44) mostra que, reciprocamente, todo operador linear onde $P^2 = P$ é uma projeção.

Teorema 1.44. *Seja $P : E \rightarrow E$ um operador linear. Se $P^2 = P$ então E é a soma direta do núcleo com a imagem de P . Ou seja, P é a projeção sobre $\mathcal{I}m(P)$ paralelamente a $\mathcal{N}(P)$.*

Demonstração. Todo $v \in E$ escreve-se como soma $v = v - (P(v) - P(v)) = (v - P(v)) + P(v)$, onde $P(v)$, evidentemente, pertence a $\mathcal{I}m(P)$ e, como $P(v - P(v)) = P(v) - P(P(v)) = P(v) - P(v) = \vec{0}$, vemos que $v - P(v) \in \mathcal{N}(P)$. Portanto $E = \mathcal{N}(P) + \mathcal{I}m(P)$. Se $w \in \mathcal{N}(P) \cap \mathcal{I}m(P)$, por um lado tem-se $P(w) = \vec{0}$ e, por outro, $P(w) = w$ (tendo em vista que estamos supondo que $w \in \mathcal{N}(P)$ e que $w \in \mathcal{I}m(P)$); logo $w = \vec{0}$. Assim $\mathcal{N}(P) \cap \mathcal{I}m(P) = \{\vec{0}\}$ e tem-se a soma direta $E = \mathcal{N}(P) \oplus \mathcal{I}m(P)$. A última afirmação do enunciado é uma consequência direta da definição de P . \square

Exemplo 1.45. Considere os subconjuntos

$$W_1 = \{(x, y) \in \mathbb{R}^2; x + y = 0\} \quad e$$

$$W_2 = \{(x, y) \in \mathbb{R}^2; x - y = 0\}.$$

Temos que $W_1 + W_2 = \mathbb{R}^2$, pois se $(x, y) \in \mathbb{R}^2$, então

$$(x, y) = \left(\frac{x-y}{2}, \frac{y-x}{2}\right) + \left(\frac{x+y}{2}, \frac{x+y}{2}\right),$$

o que mostra que todo elemento de \mathbb{R}^2 se escreve como a soma de um elemento de W_1 : $\left(\frac{x-y}{2}, \frac{y-x}{2}\right)$; e um elemento de W_2 : $\left(\frac{x+y}{2}, \frac{x+y}{2}\right)$. Além disso, temos que $W_1 \cap W_2 = \{\vec{0}\}$ e, portanto podemos definir o operador linear P , projeção de \mathbb{R}^2 sobre W_1 paralelamente a W_2 ,

$$P(x, y) = \left(\frac{x-y}{2}, \frac{y-x}{2}\right).$$

Cabe aqui mencionar que não podemos definir projeção se a soma não for direta, como veremos abaixo.

Exemplo 1.46. Sejam W_1 e W_2 subespaços vetoriais de \mathbb{R}^3 sobre a base canônica, dados por

$$W_1 = \{(x_1, 0, z_1); x_1, z_1 \in \mathbb{R}\} \quad (\text{plano } y = 0) \quad e$$

$$W_2 = \{(x_2, y_2, 0); x_2, y_2 \in \mathbb{R}\} \quad (\text{plano } z = 0).$$

Note que, $W_1 + W_2$, pois

$$(a, b, c) \in \mathbb{R}^3 \Rightarrow (a, b, c) = \underbrace{(a, 0, c)}_{\in W_1} + \underbrace{(0, b, 0)}_{\in W_2}.$$

Mas, a soma de W_1 com W_2 não é direta, pois

$$W_1 \cap W_2 = \{(x, 0, 0); x \in \mathbb{R}\} \quad (\text{eixo } x).$$

Neste contexto, não podemos definir projeção, pois se considerarmos, por exemplo, o vetor $(3, 0, 0)$, temos que

$$\underbrace{(1, 0, 0)}_{\in W_1} + \underbrace{(2, 0, 0)}_{\in W_2} = (3, 0, 0) = \underbrace{(1, 0, 0)}_{\in W_2} + \underbrace{(2, 0, 0)}_{\in W_1},$$

o que contraria o fato de que a decomposição deve ser única e o fato de a decomposição não ser única faz com que uma possível função projeção não esteja bem definida, no exemplo em questão, poderíamos projetar $(3, 0, 0)$ tanto em $(1, 0, 0)$ quanto em $(2, 0, 0)$ (ou, ainda, em qualquer vetor $(\alpha, 0, 0)$, com $\alpha \in \mathbb{R}$).

1.5 ESPAÇO VETORIAL COM PRODUTO INTERNO

1.5.1 PRODUTO INTERNO

Definição 1.47. *Seja E um espaço vetorial. Um **produto interno** em E é uma função que a cada par de vetores u e v em E associa um número real, denotado por $\langle u, v \rangle$, que satisfaz, para quaisquer vetores u, v, w de E e qualquer número real α , as seguintes propriedades:*

$$\text{P1 } \langle v, v \rangle \geq 0;$$

$$\text{P2 } \langle v, v \rangle = 0 \text{ se, e somente se, } v = \vec{0};$$

$$\text{P3 } \langle u, v \rangle = \langle v, u \rangle;$$

$$\text{P4 } \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle;$$

$$\text{P5 } \langle \alpha u, v \rangle = \alpha \langle u, v \rangle.$$

Se está definido um produto interno em E , dizemos que E é um **espaço vetorial com produto interno**. Vejamos a definição de um produto interno em \mathbb{R}^n , chamado de *produto interno usual de \mathbb{R}^n* ou *produto escalar de \mathbb{R}^n* , generalizando a noção de produto escalar de \mathbb{R}^2 e de \mathbb{R}^3 .

Exemplo 1.48. Sejam $u = (x_1, \dots, x_n)$ e $v = (y_1, \dots, y_n)$ vetores em \mathbb{R}^n .

Definamos

$$\langle u, v \rangle = x_1 y_1 + \dots + x_n y_n.$$

Note que

$$\langle u, u \rangle = x_1 x_1 + \dots + x_n x_n = x_1^2 + \dots + x_n^2 \geq 0,$$

e que

$$\langle u, u \rangle = x_1^2 + \dots + x_n^2 = 0 \iff x_1 = \dots = x_n = 0 \iff u = \vec{0},$$

mostrando que as propriedades **P1** e **P2** da Definição 1.47 são satisfeitas. A propriedade **P3** também é satisfeita já que

$$\langle u, v \rangle = x_1 y_1 + \dots + x_n y_n = y_1 x_1 + \dots + y_n x_n = \langle v, u \rangle.$$

Ademais, se $w = (z_1, \dots, z_n) \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, perceba que as propriedades **P4** e **P5** também são atendidas, pois

$$\begin{aligned} \langle u + v, w \rangle &= (x_1 + y_1)z_1 + \dots + (x_n + y_n)z_n \\ &= x_1 z_1 + y_1 z_1 + \dots + x_n z_n + y_n z_n \\ &= (x_1 z_1 + \dots + x_n z_n) + (y_1 z_1 + \dots + y_n z_n) \\ &= \langle u, w \rangle + \langle v, w \rangle \end{aligned}$$

e

$$\begin{aligned}
\langle \alpha u, v \rangle &= (\alpha x_1)y_1 + \dots + (\alpha x_n)y_n \\
&= \alpha(x_1y_1) + \dots + \alpha(x_ny_n) \\
&= \alpha(x_1y_1 + \dots + x_ny_n) \\
&= \alpha\langle u, v \rangle.
\end{aligned}$$

Proposição 1.49. *Seja E um espaço com produto interno. Se $u, v, w \in E$ e $\alpha \in \mathbb{R}$, então*

- (i) $\langle \vec{0}, u \rangle = \langle u, \vec{0} \rangle = 0$;
- (ii) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$;
- (iii) $\langle u, \alpha v \rangle = \alpha\langle u, v \rangle$;
- (iv) $\langle u, v - w \rangle = \langle u, v \rangle - \langle u, w \rangle$.

Demonstração. De fato, se E um espaço com produto interno, com $u, v, w \in E$ e $\alpha \in \mathbb{R}$, diante do exposto na Definição 1.47, temos que

- (i) Das propriedades **P3**, **P4** e **P5** da definição de produto interno

$$\begin{aligned}
\langle u, \vec{0} \rangle = \langle \vec{0}, u \rangle &= \langle u + (-u), u \rangle = \langle u, u \rangle + \langle -u, u \rangle \\
&= \langle u, u \rangle + (-1)\langle u, u \rangle \\
&= \langle u, u \rangle - \langle u, u \rangle \\
&= 0;
\end{aligned}$$

- (ii) Das propriedades **P3** e **P4**

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

- (iii) Das propriedades **P3** e **P5**

$$\langle u, \alpha v \rangle = \langle \alpha v, u \rangle = \alpha\langle v, u \rangle = \alpha\langle u, v \rangle;$$

- (iv) Dos itens (ii) e (iii) (colocando $\alpha = -1$)

$$\begin{aligned}
\langle u, v - w \rangle &= \langle u, v + (-w) \rangle = \langle u, v \rangle + \langle u, -w \rangle \\
&= \langle u, v \rangle + (-1)\langle u, w \rangle \\
&= \langle u, v \rangle - \langle u, w \rangle.
\end{aligned}$$

□

Considerando o exposto no Exemplo 1.48, note que podemos definir o produto interno usual de \mathbb{R}^n como multiplicação de matrizes (SANTOS, 2006). Sejam $u = (x_1, \dots, x_n)$ e $v = (y_1, \dots, y_n)$ vetores em \mathbb{R}^n . Escrevendo os vetores u e v como matrizes colunas, temos que o produto interno de $U = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $V = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ pode ser escrito em termos do produto entre as matrizes como

$$\langle U, V \rangle = U^T V. \quad (1.11)$$

Definição 1.50. *Seja E um espaço vetorial com produto interno. Para todo vetor $v \in E$, definimos a **norma** de E denotada por $\|v\|$ como sendo*

$$\|v\| = \sqrt{\langle v, v \rangle}. \quad (1.12)$$

Se $\|v\| = 1$, dizemos que v é um vetor unitário. Dado um vetor v , com $\|v\| \neq 1$, o vetor u unitário com mesma direção e sentido de v é dado por

$$u = \frac{1}{\|v\|} v \quad (1.13)$$

pois,

$$\|u\|^2 = \left\| \frac{1}{\|v\|} v \right\|^2 = \left\langle \frac{1}{\|v\|} v, \frac{1}{\|v\|} v \right\rangle = \frac{1}{\|v\|^2} \langle v, v \rangle = \frac{\|v\|^2}{\|v\|^2} = 1 \implies \|u\| = 1. \quad (1.14)$$

Neste caso, dizemos que o vetor u está **normalizado**, o que significa que seu comprimento é igual a 1 unidade.

Exemplo 1.51. Vamos determinar o vetor normalizado do vetor $u = (1, -2, 3)$.

Como $\|u\| = \sqrt{\langle u, u \rangle} = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}$, o vetor normalizado de u é o vetor:

$$u_1 = \frac{u}{\|u\|} = \frac{1}{\sqrt{14}}(1, -2, 3) = \left(\frac{1}{\sqrt{14}}, \frac{-2}{\sqrt{14}}, \frac{3}{\sqrt{14}} \right).$$

Proposição 1.52. *Seja E um espaço com produto interno. Se $u, v \in E$ e $\alpha \in \mathbb{R}$, então*

- (i) $\|v\| \geq 0$ e $\|v\| = 0$ se, e somente se, $v = \vec{0}$;
- (ii) $\|\alpha v\| = |\alpha| \|v\|$;
- (iii) $|\langle u, v \rangle| \leq \|u\| \|v\|$ (*Desigualdade de Cauchy-Schwarz*);
- (iv) $\|u + v\| \leq \|u\| + \|v\|$ (*Desigualdade triangular*).

Demonstração. Se E um espaço com produto interno, então dados $u, v \in E$ e $\alpha \in \mathbb{R}$, temos que

- (i) Decorre das propriedades **P1** e **P2** da definição de produto interno (1.47);
- (ii) $\|\alpha v\| = \sqrt{\langle \alpha v, \alpha v \rangle} = \sqrt{\alpha^2 \langle v, v \rangle} = \sqrt{\alpha^2} \sqrt{\langle v, v \rangle} = |\alpha| \|v\|$;
- (iii) Se os vetores u e v são linearmente dependentes, então existe $\alpha \in \mathbb{R}$ tal que $u = \alpha v$, e assim temos:

$$\begin{aligned} |\langle u, v \rangle| &= |\langle \alpha v, v \rangle| = |\alpha \langle v, v \rangle| = |\alpha| \langle v, v \rangle = |\alpha| \sqrt{\langle v, v \rangle} \sqrt{\langle v, v \rangle} \\ &= \sqrt{\alpha^2 \langle v, v \rangle} \sqrt{\langle v, v \rangle} \\ &= \sqrt{\langle \alpha v, \alpha v \rangle} \sqrt{\langle v, v \rangle} \\ &= \sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle} \\ &= \|u\| \|v\|. \end{aligned}$$

Por outro lado, se u e v forem linearmente independentes (LI), então para todo $\alpha \in \mathbb{R}$ temos que $\langle u + \alpha v, u + \alpha v \rangle > 0$ (propriedade **P1** da definição de produto interno (1.47)).

Logo,

$$\begin{aligned} \langle u + \alpha v, u + \alpha v \rangle &= \langle u, u \rangle + \alpha \langle u, v \rangle + \alpha \langle v, u \rangle + \alpha^2 \langle v, v \rangle \\ &= \langle v, v \rangle \alpha^2 + 2 \langle u, v \rangle \alpha + \langle u, u \rangle > 0 \end{aligned}$$

que é uma inequação do 2º grau na variável α . Neste contexto, temos que a equação $\langle v, v \rangle \alpha^2 + 2 \langle u, v \rangle \alpha + \langle u, u \rangle = 0$ não possui raízes reais, o que significa dizer que seu discriminante é menor que zero. Logo,

$$\begin{aligned} 4 \langle u, v \rangle^2 - 4 \langle v, v \rangle \langle u, u \rangle < 0 &\Rightarrow 4 \langle u, v \rangle^2 < 4 \langle v, v \rangle \langle u, u \rangle \\ &\Rightarrow \langle u, v \rangle^2 < \langle v, v \rangle \langle u, u \rangle \\ &\Rightarrow \langle u, v \rangle^2 < \|u\|^2 \|v\|^2 \\ &\Rightarrow \sqrt{\langle u, v \rangle^2} < \sqrt{\|u\|^2 \|v\|^2} \\ &\Rightarrow |\langle u, v \rangle| < \|u\| \|v\|, \end{aligned}$$

provando o resultado.

- (iv) Temos,

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2 |\langle u, v \rangle| + \|v\|^2 \quad \text{pois, } x \leq |x| \text{ para todo } x \in \mathbb{R} \\ &\leq \|u\|^2 + 2 \|u\| \|v\| + \|v\|^2 \quad \text{pela desigualdade de Cauchy-Schwarz} \\ &= (\|u\| \|v\|)^2. \end{aligned}$$

Extraindo as raízes quadradas em ambos os lados da desigualdade acima obtemos a desigualdade desejada.

□

Definição 1.53. Definimos o ângulo entre dois vetores não nulos u e v como sendo o número real θ entre 0 e π tal que

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}. \quad (1.15)$$

Perceba que pela desigualdade Cauchy-Schwarz o ângulo θ está bem definido, pois se dividirmos a expressão $|\langle u, v \rangle| \leq \|u\| \|v\|$ por $\|u\| \|v\|$ obtemos

$$\frac{|\langle u, v \rangle|}{\|u\| \|v\|} \leq 1 \Leftrightarrow -1 \leq \frac{\langle u, v \rangle}{\|u\| \|v\|} \leq 1. \quad (1.16)$$

Como $\cos \theta$ assume, uma única vez, cada valor no intervalo $[-1, 1]$ quando θ varia no intervalo $[0, \pi]$, segue de (1.15) que existe um único $\theta \in [0, \pi]$ tal que $\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}$.

Exemplo 1.54. Vamos calcular o o ângulo θ entre os vetores $u = (2, -2)$ e $v = (6, -1)$. Como

$$\begin{aligned} \|u\| &= \sqrt{2^2 + (-2)^2} = \sqrt{8} = 2\sqrt{2} \\ \|v\| &= \sqrt{6^2 + (-1)^2} = \sqrt{37} \\ \langle u, v \rangle &= 2 \cdot 6 + (-2) \cdot (-1) = 14, \end{aligned}$$

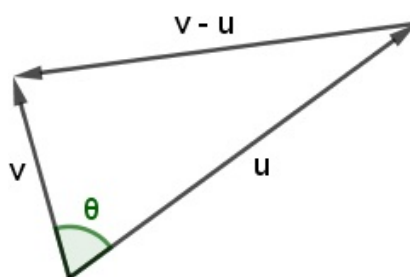
segue que

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|} = \frac{14}{2\sqrt{2}\sqrt{37}} = \frac{7}{\sqrt{74}} \approx 0,814$$

e, portanto, $\theta \approx 36^\circ$.

Exemplo 1.55. Sejam $u = (x_1, \dots, x_n)$ e $v = (y_1, \dots, y_n)$ vetores não nulos em \mathbb{R}^n . Vamos verificar geometricamente o ângulo θ entre u e v .

Figura 10 – Diferença $v - u$



Fonte – Autoria própria (2019)

Supondo que u e v não são múltiplos e aplicando Lei dos Cossenos ao triângulo determinado pelos vetores u , v e $v - u$ (Figura 10), obtemos:

$$\|v - u\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos \theta,$$

Dáí:

$$\begin{aligned} & 2\|u\|\|v\| \cos \theta \\ = & \|u\|^2 + \|v\|^2 - \|v - u\|^2 \\ = & (x_1^2 + \dots + x_n^2) + (y_1^2 + \dots + y_n^2) - ((y_1 - x_1)^2 + \dots + (y_n - x_n)^2) \\ = & x_1^2 + \dots + x_n^2 + y_1^2 + \dots + y_n^2 - (y_1^2 - 2x_1y_1 + x_1^2 + \dots + y_n^2 - 2x_ny_n + x_n^2) \\ = & x_1^2 + \dots + x_n^2 + y_1^2 + \dots + y_n^2 - y_1^2 + 2x_1y_1 - x_1^2 - \dots - y_n^2 + 2x_ny_n - x_n^2 \\ = & 2x_1y_1 + \dots + 2x_ny_n \\ = & 2(x_1y_1 + \dots + x_ny_n). \end{aligned}$$

Portanto,

$$\|u\|\|v\| \cos \theta = x_1y_1 + \dots + x_ny_n = \langle u, v \rangle \Leftrightarrow \cos \theta = \frac{\langle u, v \rangle}{\|u\|\|v\|}.$$

No caso em que u e v são múltiplos, temos que, existe um número real $\lambda \neq 0$ tal que $u = \lambda v$. Logo, $x_i = \lambda y_i$, para $i = 1, \dots, n$. Como $\cos \angle(\lambda v, v) = 1$ se $\lambda > 0$ e $\cos \angle(\lambda v, v) = -1$ se $\lambda < 0$, segue que

$$\langle u, v \rangle = \|\lambda v\|\|v\| = \|\lambda v\|\|v\| \cos \theta = \|u\|\|v\| \cos \theta$$

comprovando o resultado.

1.5.2 ORTOGONALIDADE E PROJEÇÃO ORTOGONAL

Definição 1.56. *Seja E um espaço vetorial com produto interno e sejam $u, v \in E$. Dizemos que u e v são **ortogonais** se $\langle u, v \rangle = 0$. Um subconjunto W de E é chamado de **ortogonal** se os seus elementos são ortogonais dois a dois e dizemos que W é um conjunto **ortonormal** se for um conjunto ortogonal e se $\|u\| = 1, \forall u \in W$.*

Perceba que pela Equação (1.15) u e v são ortogonais se, e somente se, $\cos \theta = 0$, onde θ é o ângulo entre u e v , e isto é verdade se, e somente se, $\theta = \frac{\pi}{2}$.

Usaremos a notação $u \perp v$ para indicar que os vetores u e v são ortogonais. Perceba que a ortogonalidade depende do produto interno, isto é, dois vetores podem ser ortogonais em relação a um produto interno mas não em relação a outro.

Exemplo 1.57. *Seja E um espaço vetorial com produto interno. O vetor nulo $\vec{0}$ é ortogonal a todos os elementos de E , pois $\langle 0, v \rangle = 0$, para todo $v \in E$ (item (i) da Proposição 1.49).*

Exemplo 1.58. Considere o \mathbb{R}^3 com o produto interno usual e o conjunto

$$X = \{w_1 = (1, 1, 1), w_2 = (-1, 1, 0), w_3 = (-1, -1, 2)\}.$$

(a) W é ortogonal, pois

$$\begin{aligned}\langle w_1, w_2 \rangle &= \langle (1, 1, 1), (-1, 1, 0) \rangle = -1 + 1 + 0 = 0, \\ \langle w_1, w_3 \rangle &= \langle (1, 1, 1), (-1, -1, 2) \rangle = -1 - 1 + 2 = 0, \\ \langle w_2, w_3 \rangle &= \langle (-1, 1, 0), (-1, -1, 2) \rangle = 1 - 1 + 0 = 0.\end{aligned}$$

(b) Note que

$$\begin{aligned}\|w_1\| &= \|(1, 1, 1)\| = \sqrt{\langle (1, 1, 1), (1, 1, 1) \rangle} = \sqrt{3}, \\ \|w_2\| &= \|(-1, 1, 0)\| = \sqrt{\langle (-1, 1, 0), (-1, 1, 0) \rangle} = \sqrt{2}, \\ \|w_3\| &= \|(-1, -1, 2)\| = \sqrt{\langle (-1, -1, 2), (-1, -1, 2) \rangle} = \sqrt{6}.\end{aligned}$$

Assim, obtendo vetores u_1, u_2 e u_3 unitários de mesma direção e sentido que os vetores $(1, 1, 1)$, $(-1, 1, 0)$ e $(-1, -1, 2)$, respectivamente, temos

$$\begin{aligned}u_1 &= \frac{1}{\|w_1\|}w_1 = \frac{1}{\sqrt{3}}(1, 1, 1) = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \\ u_2 &= \frac{1}{\|w_2\|}w_2 = \frac{1}{\sqrt{2}}(-1, 1, 0) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \\ u_3 &= \frac{1}{\|w_3\|}w_3 = \frac{1}{\sqrt{6}}(-1, -1, 2) = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right).\end{aligned}$$

Logo, $\{u_1, u_2, u_3\}$ é um conjunto ortonormal.

Teorema 1.59. Num espaço vetorial E com produto interno, todo conjunto ortogonal X de vetores não-nulos é LI.

Demonstração. Sejam $v_1, \dots, v_n \in X$. Pela ortogonalidade de X , temos que $\langle v_i, v_j \rangle = 0$ se $i \neq j$, ademais, da definição de norma $\langle v_i, v_i \rangle = \|v_i\|^2$. Se $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ é uma combinação linear nula desses vetores então, para cada $i = 1, 2, \dots, n$, tomando o produto interno de ambos os membros desta igualdade por v_i temos que

$$\begin{aligned}\langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_i \rangle &= \langle 0, v_i \rangle \Rightarrow \alpha_1 \langle v_1, v_i \rangle + \dots + \alpha_n \langle v_n, v_i \rangle = 0 \\ &\Rightarrow \alpha_i \langle v_i, v_i \rangle = \alpha_i \|v_i\|^2 = 0.\end{aligned}$$

Como os vetores pertencentes ao conjunto X são todos não-nulos, temos

$$\alpha_i \|v_i\|^2 = 0 \implies \alpha_i = 0,$$

o que significa dizer que os coeficientes da combinação linear $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ são todos iguais a zero e os vetores do conjunto X são, portanto, linearmente independentes. \square

A recíproca do resultado acima é falsa, pois, por exemplo, o conjunto $\{(1, 1), (1, 0)\}$ de vetores em \mathbb{R}^2 com o produto interno usual é linearmente independente, como se segue

$$a(1, 1) + b(1, 0) = \vec{0} \implies a = b = 0$$

mas não é um conjunto ortogonal, pois

$$\langle (1, 1), (1, 0) \rangle = 1 \neq 0.$$

Se $\alpha = \{v_1, \dots, v_n\}$ é um conjunto ortogonal de vetores não nulos de E , e além disso $\dim(E) = n$, segue do teorema anterior que α é uma base de E . Uma base de vetores ortogonais é chamada **base ortogonal** e uma base de vetores ortonormais é chamada **base ortonormal**.

Teorema 1.60. *Se $\alpha = \{v_1, \dots, v_n\}$ é uma base ortonormal de E , então, para todo $v \in E$, podemos escrever*

$$v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n. \quad (1.17)$$

Demonstração. Seja $v = a_1 v_1 + \dots + a_n v_n$ a combinação linear de v com relação aos elementos da base α . Fixe i , com $1 \leq i \leq n$. Temos

$$\begin{aligned} \langle v, v_i \rangle &= \langle a_1 v_1 + \dots + a_n v_n, v_i \rangle \\ &= a_1 \langle v_1, v_i \rangle + \dots + a_n \langle v_n, v_i \rangle = a_i \end{aligned}$$

pois $\langle v_j, v_i \rangle = 0$ se $j \neq i$ e $\langle v_i, v_i \rangle = \|v_i\|^2 = 1$. Como i foi tomado de modo arbitrário, a demonstração está completa. \square

Se $\beta = \{v_1, \dots, v_n\}$ é uma base ortogonal de E , normalizando cada um dos vetores de β , obtemos a base ortonormal α de E , onde

$$\alpha = \left\{ \frac{v_1}{\|v_1\|}, \dots, \frac{v_n}{\|v_n\|} \right\}.$$

Pelo Teorema 1.60, para cada vetor v em E , temos que

$$\begin{aligned} v &= \left\langle v, \frac{v_1}{\|v_1\|} \right\rangle \frac{v_1}{\|v_1\|} + \dots + \left\langle v, \frac{v_n}{\|v_n\|} \right\rangle \frac{v_n}{\|v_n\|} \\ &= \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \dots + \frac{\langle v, v_n \rangle}{\|v_n\|^2} v_n. \end{aligned}$$

O número real

$$a_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2} \quad (1.18)$$

é chamado de coeficiente de *Fourier* de v em relação ao vetor v_i . Este escalar admite uma interpretação geométrica relacionada com a noção de projeção.

Proposição 1.61. *Seja w um vetor não nulo de E . Se $v \in E$, então*

$$k = \frac{\langle v, w \rangle}{\langle w, w \rangle} = \frac{\langle v, w \rangle}{\|w\|^2} \quad (1.19)$$

é o único número real tal que $v' = v - kw$ é ortogonal a w .

Demonstração. Para que v' seja ortogonal a w devemos ter

$$\begin{aligned} \langle v - kw, w \rangle = 0 &\Rightarrow \langle v, w \rangle - \langle kw, w \rangle = 0 \\ &\Rightarrow \langle v, w \rangle = k\langle w, w \rangle \\ &\Rightarrow k = \frac{\langle v, w \rangle}{\langle w, w \rangle}. \end{aligned}$$

Reciprocamente, suponhamos que $k = \frac{\langle v, w \rangle}{\langle w, w \rangle}$. Então,

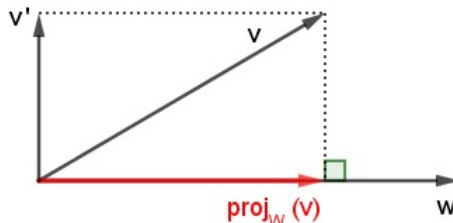
$$\langle v - kw, w \rangle = \langle v, w \rangle - k\langle w, w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, w \rangle = 0,$$

o que mostra que $v - kw$ é ortogonal a w . □

Definição 1.62. *Considerando o escalar k da Proposição 1.61 (coeficiente de Fourier de v em relação ao vetor w). A **projeção ortogonal de v sobre w** (Figura 11) é denotada por $\text{proj}_w(v)$ e é definida por*

$$\text{proj}_w(v) = kw = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = \frac{\langle v, w \rangle}{\|w\|^2} w. \quad (1.20)$$

Figura 11 – Projeção ortogonal

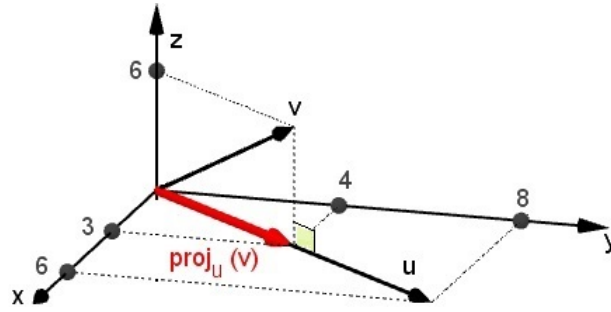


Fonte – Autoria própria (2019)

Exemplo 1.63. *Seja o espaço vetorial \mathbb{R}^3 com o produto interno usual. Vamos determinar a projeção ortogonal do vetor $v = (3, 4, 6)$ sobre o vetor $u = (6, 8, 0)$ (Figura 12).*

$$\text{proj}_u(v) = \frac{\langle v, u \rangle}{\|u\|^2} u = \frac{\langle (3, 4, 6), (6, 8, 0) \rangle}{\langle (6, 8, 0), (6, 8, 0) \rangle} (6, 8, 0) = \frac{1}{2} (6, 8, 0) = (3, 4, 0).$$

Figura 12 – Projecção ortogonal do vetor $v = (3, 4, 6)$ sobre o vetor $u = (6, 8, 0)$



Fonte – Autoria própria (2019)

Exemplo 1.64. Seja $E = C^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Vamos determinar a projeção de t^3 sobre t .

$$\left(\text{proj}_t(t^3)\right)(t) = \frac{\langle t^3, t \rangle}{\|t\|^2}t = \frac{\langle t^3, t \rangle}{\langle t, t \rangle}t = \frac{\int_{-1}^1 t^4 dt}{\int_{-1}^1 t^2 dt}t = \frac{2/5}{2/3}t = \frac{3}{5}t.$$

Proposição 1.65. (Generalização da Proposição 1.61) Suponhamos que $\{w_1, \dots, w_r\}$ seja um conjunto ortogonal de vetores não nulos de E . Se $v \in E$, então

$$k_i = \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} = \frac{\langle v, w_i \rangle}{\|w_i\|^2}, \quad 1 \leq i \leq r, \quad (1.21)$$

são os únicos números reais tais que o vetor

$$v' = v - k_1w_1 - k_2w_2 - \dots - k_rw_r \quad (1.22)$$

é ortogonal aos vetores w_1, \dots, w_r .

Demonstração. Considerando que $\{w_1, \dots, w_r\}$ é ortogonal, temos que $\langle w_i, w_j \rangle = 0$ se $i \neq j$. Neste contexto, para que v' seja ortogonal ao vetor w_i , $1 \leq i \leq r$, devemos ter

$$\begin{aligned} \langle v - k_1w_1 - k_2w_2 - \dots - k_rw_r, w_i \rangle &= 0 \\ \Rightarrow \langle v, w_i \rangle - k_1\langle w_1, w_i \rangle - \dots - k_r\langle w_r, w_i \rangle &= 0 \\ \Rightarrow \langle v, w_i \rangle - k_i\langle w_i, w_i \rangle &= 0 \\ \Rightarrow \langle v, w_i \rangle &= k_i\langle w_i, w_i \rangle, \end{aligned}$$

mostrando que $k_i = \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} = \frac{\langle v, w_i \rangle}{\|w_i\|^2}$. Por outro lado, se $k_i = \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle}$, então

$$\begin{aligned} \langle v - k_1w_1 - \dots - k_rw_r, w_i \rangle &= \langle v, w_i \rangle - k_1\langle w_1, w_i \rangle - \dots - k_r\langle w_r, w_i \rangle \\ &= \langle v, w_i \rangle - k_i\langle w_i, w_i \rangle \\ &= \langle v, w_i \rangle - \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle}\langle w_i, w_i \rangle = 0, \end{aligned}$$

o que mostra que v' é ortogonal aos vetores w_1, \dots, w_r , completando a demonstração. \square

Veremos a seguir que todo espaço com produto interno, não nulo, de dimensão finita tem uma base ortonormal. A construção dada na prova do resultado abaixo é chamada de **processo de ortogonalização de Gram-Schmidt**, pois leva os nomes de Jorgen Pedersen Gram (Dinamarca, 1850 - 1916) e de Erhard Schmidt (Alemanha, 1876 - 1959).

Teorema 1.66. *O espaço E possui uma base ortogonal.*

Demonstração. Seja $\{v_1, \dots, v_n\}$ uma base de E . Tomemos

$$\begin{aligned} w_1 &= v_1, \\ w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1, \\ w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2, \\ &\vdots \\ w_n &= v_n - \frac{\langle v_n, w_1 \rangle}{\|w_1\|^2} w_1 - \dots - \frac{\langle v_n, w_{n-1} \rangle}{\|w_{n-1}\|^2} w_{n-1}. \end{aligned}$$

Pela Proposição 1.65, o conjunto $\{w_1, \dots, w_n\}$ é um conjunto ortogonal. Além disso, como o conjunto $\{v_1, \dots, v_n\}$ é linearmente independente, cada vetor w_i é não nulo. Assim, o conjunto $\{w_1, \dots, w_n\}$ é um conjunto ortogonal de vetores não nulos de E . Como, por definição, $n = \dim(E)$, segue pelo Teorema 1.59 que $\{w_1, \dots, w_n\}$ é uma base ortogonal de E . \square

Exemplo 1.67. Considere o espaço \mathbb{R}^3 com o produto interno usual. Apliquemos o processo de Gram-Schmidt ao conjunto $T = \{(1, 0, 0), (1, 1, 1), (0, 0, 1)\}$ para obtermos uma base ortogonal $\{w_1, w_2, w_3\}$ de \mathbb{R}^3 .

Primeiramente perceba que T é uma base de \mathbb{R}^3 , já que $\dim(\mathbb{R}^3) = 3$ e ademais T é linearmente independentes, pois

$$a(1, 0, 0) + b(1, 1, 1) + c(0, 0, 1) = (a + b, b, b + c) = 0 \implies a = b = c = 0.$$

Façamos

$$\begin{aligned} w_1 &= (1, 0, 0), \\ w_2 &= (1, 1, 1) - \frac{\langle (1, 1, 1), (1, 0, 0) \rangle}{\|(1, 0, 0)\|^2} (1, 0, 0) = (0, 1, 1), \\ w_3 &= (0, 0, 1) - \frac{\langle (0, 0, 1), (1, 0, 0) \rangle}{\|(1, 0, 0)\|^2} (1, 0, 0) \\ &\quad - \frac{\langle (0, 0, 1), (0, 1, 1) \rangle}{\|(0, 1, 1)\|^2} (0, 1, 1) = \left(0, -\frac{1}{2}, \frac{1}{2}\right). \end{aligned}$$

Assim, $\left\{ (1, 0, 0), (0, 1, 1), \left(0, -\frac{1}{2}, \frac{1}{2}\right) \right\}$ é uma base ortogonal de \mathbb{R}^3 .

Corolário 1.68. *Qualquer espaço vetorial E possui uma base ortonormal.*

Demonstração. Como E possui uma base ortogonal (Teorema 1.66), basta normalizar cada vetor da respectiva base para obter uma base ortonormal. \square

Exemplo 1.69. Do Exemplo 1.67, segue que $\left\{(1, 0, 0), (0, 1, 1), \left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\}$ é uma base ortogonal de \mathbb{R}^3 . Como

$$\begin{aligned}\|(1, 0, 0)\| &= \sqrt{\langle(1, 0, 0), (1, 0, 0)\rangle} = 1, \\ \|(0, 1, 1)\| &= \sqrt{\langle(0, 1, 1), (0, 1, 1)\rangle} = \sqrt{2}, \\ \left\|\left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\| &= \sqrt{\left\langle\left(0, -\frac{1}{2}, \frac{1}{2}\right), \left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\rangle} = \frac{1}{\sqrt{2}},\end{aligned}$$

Normalizando os vetores, temos que

$$\begin{aligned}\frac{1}{\|(1, 0, 0)\|}(1, 0, 0) &= (1, 0, 0), \\ \frac{1}{\|(0, 1, 1)\|}(0, 1, 1) &= \frac{1}{\sqrt{2}}(0, 1, 1) = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \\ \frac{1}{\left\|\left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\|}\left(0, -\frac{1}{2}, \frac{1}{2}\right) &= \sqrt{2}\left(0, -\frac{1}{2}, \frac{1}{2}\right) = \left(0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).\end{aligned}$$

Logo, $\left\{(1, 0, 0), \left(0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)\right\}$ é um base ortonormal de \mathbb{R}^3 .

Definição 1.70. Seja E um espaço vetorial com produto interno, e seja $W \subseteq E$ um subconjunto de E . Chamamos de **complemento ortogonal de W** ao conjunto

$$W^\perp = \{v \in E : \langle v, u \rangle = 0, \forall u \in W\}. \quad (1.23)$$

Exemplo 1.71. Para \mathbb{R}^2 com o produto interno usual e $W = \{(1, 2)\}$, temos

$$\begin{aligned}W^\perp &= \{(x, y) \in \mathbb{R}^2; \langle(x, y), (1, 2)\rangle = 0\} \\ &= \{(x, y) \in \mathbb{R}^2; x + 2y = 0\} \\ &= \{(x, y) \in \mathbb{R}^2; x = -2y\} \\ &= \{(-2y, y) \in \mathbb{R}^2; y \in \mathbb{R}\}.\end{aligned}$$

Proposição 1.72. Seja W um subconjunto de um espaço com produto interno E . Então W^\perp é um subespaço de E .

Demonstração. $\forall u \in W, \alpha \in \mathbb{R}$ e $v_1, v_2 \in W^\perp$, temos que

- $\vec{0} \in W^\perp$ pois $\langle \vec{0}, u \rangle = 0$;
- $\langle v_1, u \rangle = \langle v_2, u \rangle = 0$. Logo, $\langle v_1 + v_2, u \rangle = \langle v_1, u \rangle + \langle v_2, u \rangle = 0$, e então $v_1 + v_2 \in W^\perp$;
- $\langle \alpha v_1, u \rangle = \alpha \langle v_1, u \rangle = 0$.

O que garante que W^\perp é um subespaço vetorial de E (mesmo que W não tenha estrutura de espaço vetorial). \square

Proposição 1.73. *Seja E um espaço vetorial sobre \mathbb{R} com produto interno. Sejam $W \subseteq E$ um subconjunto e $\beta = \{w_1, \dots, w_k\}$ um conjunto gerador para W . Então $v \in W^\perp$ se, e somente se, $\langle v, w_i \rangle = 0$, para cada $i = 1, \dots, k$.*

Demonstração. Sejam $w \in W$ e $\beta = \{w_1, \dots, w_k\}$ um conjunto gerador para W . Então, existem $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tais que $w = \alpha_1 w_1 + \dots + \alpha_k w_k$. Logo,

$$\langle v, w \rangle = \langle v, \alpha_1 w_1 + \dots + \alpha_k w_k \rangle = \alpha_1 \langle v, w_1 \rangle + \dots + \alpha_k \langle v, w_k \rangle.$$

Se assumirmos que $\langle v, w_i \rangle = 0$ para cada $i = 1, \dots, k$, segue que $\langle v, w \rangle = 0$, o que significa dizer que, $v \in W^\perp$. Por outro lado, se $v \in W^\perp$, então $\langle v, w \rangle = 0$, para cada $v \in W$. Em particular, $\langle v, w_i \rangle = 0$ para cada $i = 1, \dots, k$. \square

Exemplo 1.74. Seja M_2 o espaço vetorial das matrizes de ordem 2, com o produto interno

$$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22},$$

onde $A = (a_{ij})_{i,j}$ e $B = (b_{ij})_{i,j}$, $i, j = 1, 2$. Seja

$$W = \left\{ \begin{pmatrix} x & y \\ x+y & w \end{pmatrix}; x, y, w \in \mathbb{R} \right\}.$$

(i) Determinar uma base ortogonal de W .

Note que,

$$\begin{aligned} \begin{pmatrix} x & y \\ x+y & w \end{pmatrix} &= \begin{pmatrix} x & 0 \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & w \end{pmatrix} \\ &= x \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + w \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

o que significa dizer que $\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ é um conjunto gerador de W . Além disso, vê-se facilmente que β é LI, pois

$$\begin{pmatrix} a & b \\ a+b & c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} a = 0 \\ b = 0 \\ c = 0 \end{cases}.$$

Assim, β é uma base de W . No entanto, β não é ortogonal, pois

$$\left\langle \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle = 1 \neq 0.$$

Aplicando o processo de ortogonalização de Gram-Schmidt à base β , temos

$$\begin{aligned}
 w_1 &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\
 w_2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\|^2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\
 w_3 &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\|^2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
 &\quad - \frac{\left\langle \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|^2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.
 \end{aligned}$$

Portanto uma base ortogonal para W é dada pelo conjunto

$$\left\{ \begin{pmatrix} 1 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

(ii) Determinar uma base ortogonal de W^\perp .

Seja $v = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in W^\perp$, com $a, b, c, d \in \mathbb{R}$. Então

$$\begin{cases} \left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle = 0 \\ \left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle = 0 \\ \left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} 1 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \right\rangle = 0 \end{cases} \Rightarrow \begin{cases} d = 0 \\ c = -b \\ a = b \end{cases}.$$

Logo $W^\perp = \left\{ \begin{pmatrix} b & b \\ -b & 0 \end{pmatrix}; b \in \mathbb{R} \right\}$ e, portanto, uma base para W^\perp é dada pelo conjunto

$$\left\{ \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \right\}.$$

Teorema 1.75. Se W é um subespaço de E , então

$$E = W \oplus W^\perp. \tag{1.24}$$

Demonstração. Vamos provar inicialmente que $W \cap W^\perp = \{\vec{0}\}$. Para tanto, suponhamos por absurdo que $W \cap W^\perp \neq \{\vec{0}\}$. Isto significa dizer que existe um vetor não nulo $v \in W \cap W^\perp$ tal que $v \in W$ e $v \in W^\perp$. Como W^\perp é complemento ortogonal de W temos que $\|v\|^2 = \langle v, v \rangle = 0$, o que é um absurdo pois $v \neq \vec{0}$. Portanto, $W \cap W^\perp = \{\vec{0}\}$.

Vejamus que $E = W + W^\perp$. Pelo processo de ortogonalização de Gram-Schmidt, existe uma base ortonormal $\{v_1, \dots, v_n\}$ de W . Tomemos $v \in E$. Defina

$$\begin{aligned} w_1 &= \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n = \text{proj}_W(v), \\ w_2 &= v - w_1. \end{aligned}$$

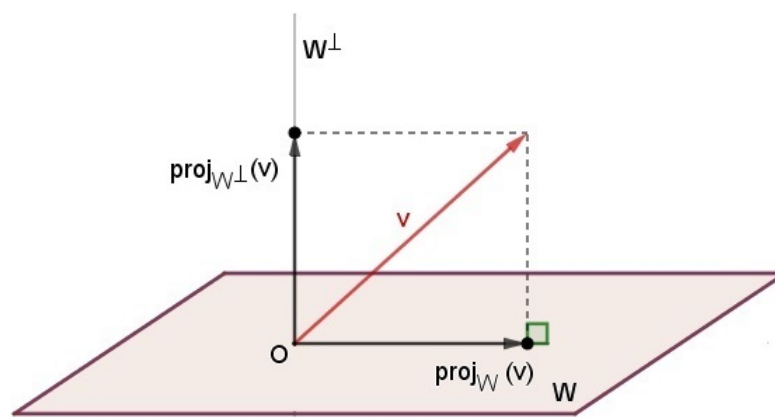
Note que $w_1 + w_2 = w_1 + (v - w_1) = v$. Além disso, $w_1 \in W$, pois w_1 é uma combinação linear dos vetores da base de W . Portanto, resta mostrar que $w_2 \in W^\perp$, ou seja, w_2 é ortogonal a W . Para isto, pela Proposição 1.73, basta verificar que w_2 é ortogonal a cada vetor v_i da base de W , com $i = 1, \dots, n$, o que é verdade pelo exposto na Proposição 1.65, pois a base $\{v_1, \dots, v_n\}$ de W é ortonormal. \square

Do Teorema 1.75, acima demonstrado, temos que cada vetor v de E pode ser escrito de modo único como

$$v = w_1 + w_2 \tag{1.25}$$

onde $w_1 \in W$ e $w_2 \in W^\perp$. O vetor w_1 é chamado *projeção ortogonal de v em W* e é denotado por $\text{proj}_W(v)$. O vetor w_2 é chamado *componente de v ortogonal a W* e é denotado por $\text{proj}_{W^\perp}(v)$. Assim, $v = \text{proj}_W(v) + \text{proj}_{W^\perp}(v)$ (Figura 13).

Figura 13 – $v = \text{proj}_W(v) + \text{proj}_{W^\perp}(v)$



Fonte – Autoria própria (2019)

Exemplo 1.76. (Retomando o Exemplo 1.74) Considerando o espaço vetorial M_2 com produto interno definido,

$$W = \left\{ \begin{pmatrix} x & y \\ x + y & w \end{pmatrix}; x, y, w \in \mathbb{R} \right\}$$

e

$$W^\perp = \left\{ \begin{pmatrix} b & b \\ -b & 0 \end{pmatrix}; b \in \mathbb{R} \right\}.$$

Pelo Teorema 1.75 temos que $M_2 = W \oplus W^\perp$.

Além disso, se supormos por absurdo que $W \cap W^\perp \neq \{\vec{0}\}$, então existe um vetor $v \in W \cap W^\perp$ não nulo tal que $v \in W$ e $v \in W^\perp$. Assim, como W^\perp é complemento ortogonal de W temos que $\langle v, v \rangle = \|v\|^2 \neq 0$, o que é um absurdo. Portanto, $W \cap W^\perp = \{\vec{0}\}$.

2 PROJEÇÕES E ANÁLISE NUMÉRICA

O estudo de projeções é um assunto clássico no estudo de álgebra linear com muita utilidade na área de Análise Numérica, onde desempenha importante papel no desenvolvimento de métodos para resolução dos mais diversos tipos de problemas. Em se tratando da resolução de sistemas lineares, podemos citar dois dos principais métodos iterativos baseados em projeção sobre subespaços de Krylov: Métodos dos Gradientes Conjugados (*Conjugate Gradient Method*) e o GMRES (*Generalized Minimum Residual Method*). Ambos serão trabalhados nas próximas seções.

Vale mencionar que para a composição deste capítulo utilizamos alternadamente um conjunto de autores como base teórica, tais quais, (KELLEY, 1995), (BEGIATO, 2007), (SAAD, 2003), (DEMME, 1997), (GREENBAUM, 1997), (TREFETHEN; BAU, 1997), (HESTENES; STIEFEL, 1952) e (LAGO, 2010).

2.1 MOTIVAÇÃO

Considere o sistema linear

$$As = b, \quad (2.1)$$

em que b é um vetor de \mathbb{R}^m e A é uma matriz $m \times n$. Escolhendo arbitrariamente $s_0 \in \mathbb{R}^m$ e aplicando ao Sistema (2.1), temos duas possibilidades:

- (i) s_0 é solução do sistema, o que implica que $As_0 = b$;
- (ii) s_0 não é solução do sistema, implicando $As_0 \neq b$.

Independente das situações distintas citadas acima, podemos reescrever o Sistema (2.1) da forma

$$As + r = b \quad (2.2)$$

com r sendo denominado *resíduo*, o que acaba implicando que a solução do referido Sistema (2.2) somente ocorre quando $r = 0$.

Exemplo 2.1. Considere o sistema linear

$$\begin{cases} x + y + z & = & 6 \\ 2x - y + z & = & 3 \\ -x + 3y - 2z & = & -1 \end{cases} \quad (2.3)$$

cuja forma matricial é dada por

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ -1 & 3 & -2 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_s = \underbrace{\begin{pmatrix} 6 \\ 3 \\ -1 \end{pmatrix}}_b.$$

Tomando $s_0 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ arbitrário como possível candidato a solução, temos que o resíduo correspondente é

$$r_0 = b - As_0 = \begin{pmatrix} 6 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

implicando que $s_0 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ não é solução do sistema (2.3). Por outro lado, tomando $s_1 =$

$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, e o aplicando ao sistema (2.3) temos que o resíduo correspondente é

$$r_1 = b - As_1 = \begin{pmatrix} 6 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

o que significa dizer que $s_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ é solução do sistema 2.3.

Neste contexto é que se desenvolvem os métodos iterativos Gradiente Conjugados e GMRES, pois o objetivo (resguardado suas respectivas características) é minimizar r , de modo que se possa explicitar ou estimar (com a precisão exigida) a solução do respectivo sistema linear partindo de uma estimativa inicial qualquer, no caso s_0 .

A minimização de $r = As - b$ é obtida a partir de uma sequência de aproximações sucessivas x_1, x_2, \dots, x_n partindo-se de uma aproximação inicial s_0 . No caso dos métodos iterativos Gradientes Conjugados e GMRES, tais aproximações ocorrem nos subespaços de Krylov.

2.2 SUBESPAÇOS DE KRYLOV

Definição 2.2. Dada uma matriz $A : n \times n$, não singular, e um vetor $v : n \times 1$, a sequência: v, Av, A^2v, A^3v, \dots é chamada sequência de Krylov de A com relação a v . O subespaço

$\mathcal{K}_l(A, v) = \text{span}\{v, Av, A^2v, A^3v, \dots, A^{l-1}v\}$ é denominado *subespaço de Krylov de ordem l de A com relação a v* .

Exemplo 2.3. Seja a matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ e o vetor $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Temos que

$$\begin{aligned} \mathcal{K}_3(A, v) &= \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \\ &= \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 18 \end{pmatrix} \right\} \end{aligned}$$

é o subespaço de ordem 3 de A com relação ao vetor v .

Convém conhecermos algumas propriedades do subespaço de Krylov, a fim de justificarmos futuros resultados.

Proposição 2.4. *O subespaço de Krylov é o subespaço de todos os vetores em \mathbb{R}^n que podem ser escritos como $x = p(A)v$, onde p é um polinômio de grau não superior a $l - 1$.*

Demonstração. Seja $x \in \mathcal{K}_l(A, v) = \text{span}\{v, Av, A^2v, A^3v, \dots, A^{l-1}v\}$, logo x pode ser escrito como combinação linear dos elementos do subespaço de Krylov. Desta forma, temos que

$$\begin{aligned} x &= \alpha_1 v + \alpha_2 Av + \alpha_3 A^2v + \dots + \alpha_l A^{l-1}v \\ &= (\alpha_1 + \alpha_2 A + \alpha_3 A^2 + \dots + \alpha_l A^{l-1})v = p(A)v \end{aligned}$$

onde $p(A) = \alpha_1 + \alpha_2 A + \alpha_3 A^2 + \dots + \alpha_l A^{l-1}$ é um polinômio de grau não superior a $l - 1$. \square

Definição 2.5. *Um polinômio $p(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n$, onde $\alpha_n = 1$, é chamado polinômio mônico de grau n . O polinômio mônico não-nulo p de menor grau que satisfaz $p(A)v = 0$ é conhecido como polinômio mínimo do vetor v com relação à matriz A .*

Definição 2.6. *Seja E um \mathbb{R} -espaço vetorial de dimensão n e $T : E \rightarrow E$ um operador linear. O subespaço vetorial $S \subseteq E$ é denominado **subespaço vetorial invariante pelo operador T** ou **subespaço T -invariante** quando $T(S) \subseteq S$, sendo $T(S) = \{T(s); s \in S\}$, ou seja, a imagem de T restrito a S .*

Exemplo 2.7. Seja $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tal que $T(x, y) = (3x, 8x - y)$.

O subespaço $S = \{(x, 2x); x \in \mathbb{R}\}$ é T -invariante, já que $T(x, 2x) = (3x, 6x) \in S$.

Por outro lado, O subespaço $W = \{(x, 0); x \in \mathbb{R}\}$ não é T -invariante, já que $T(1, 0) = (3, 8) \notin W$.

Proposição 2.8. *Seja μ o grau do polinômio mínimo de v com relação a A . Então $\mathcal{K}_\mu(A, v)$ é um espaço invariante em relação à matriz A e $\mathcal{K}_l(A, v) = \mathcal{K}_\mu(A, v)$, para todo $l \geq \mu$.*

Demonstração. Seja $u \in \mathcal{K}_\mu(A, v)$, logo, $u = \alpha_1 v + \alpha_2 Av + \dots + \alpha_\mu A^{\mu-1}v$. Temos então que

$$Au = \alpha_1 Av + \alpha_2 A^2 v + \dots + \alpha_\mu A^\mu v. \quad (2.4)$$

Mas, como μ é o grau do polinômio mínimo de v com relação a A , existe um polinômio $p(x) = \beta_0 + \beta_1 x + \dots + \beta_\mu x^\mu$, com $\beta_\mu \neq 0$, tal que $p(A)v = 0$, o que nos dá:

$$A^\mu v = - \left(\frac{\beta_0}{\beta_\mu} v + \frac{\beta_1}{\beta_\mu} Av + \dots + \frac{\beta_{\mu-1}}{\beta_\mu} A^{\mu-1} v \right). \quad (2.5)$$

Substituindo-se (2.5) em (2.4), pode-se concluir que $Au \in \mathcal{K}_\mu(A, v)$, para todo $u \in \mathcal{K}_\mu(A, v)$ e, portanto, que $\mathcal{K}_\mu(A, v)$ é invariante sobre A . Para prova a segunda parte, basta notar que, como $\mathcal{K}_\mu(A, v)$ é invariante sobre A , tem-se que $\mathcal{K}_\mu(A, v) = \mathcal{K}_{\mu+1}(A, v)$ e um argumento indutivo leva-nos à conclusão. \square

Proposição 2.9. *Se μ é o grau do polinômio mínimo de v com relação a A , então $\dim(\mathcal{K}_\mu(A, v)) = \mu$.*

Demonstração. Suponha por absurdo que $\dim(\mathcal{K}_\mu(A, v)) < \mu$. Se $v = 0$ temos que o polinômio mínimo tem grau 0 e a contradição é óbvia. Se $v \neq 0$, o conjunto $\{v, Av, A^2v, \dots, A^{\mu-1}v\}$ não seria linearmente independente e existiria $(\alpha_1, \alpha_2, \dots, \alpha_\mu) \neq 0$ tal que $\alpha_1 v + \alpha_2 Av + \dots + \alpha_\mu A^{\mu-1}v = 0$ e, neste caso, teríamos o grau do polinômio mínimo de v com relação a A menor ou igual a $\mu - 1$, o que contradiz a hipótese. \square

Corolário 2.10. *Se o grau do polinômio mínimo de v com respeito a A é μ , então*

$$\dim(\mathcal{K}_l(A, v)) = \min\{l, \mu\}. \quad (2.6)$$

Demonstração. Se $l \geq \mu$ o resultado segue diretamente das Proposições 2.8 e 2.9. Agora se $l < \mu$ o resultado segue da Proposição 2.9. \square

2.3 MÉTODO GMRES

O método GMRES (*Generalized Minimal Residual*), proposto em 1986 por Saad e Schultz (SAAD, 2003), é um método iterativo para a resolução de sistemas lineares

$$As = b,$$

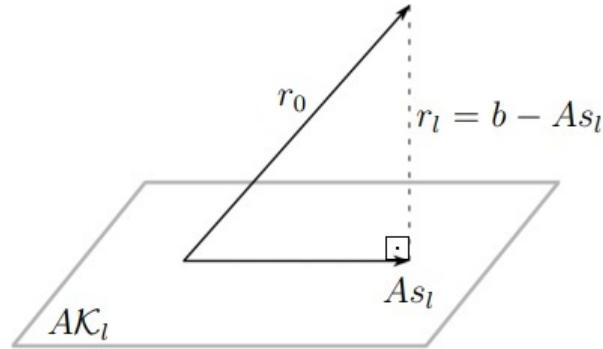
onde A é uma matriz $n \times n$, não-singular.

Considerando s_0 a aproximação inicial, cada iteração do método GMRES consiste na minimização da norma do resíduo sobre o subespaço de Krylov determinado pela matriz A e o resíduo inicial ($r_0 = b - As_0$), ou seja, se estivermos na l -ésima iteração do método de GMRES, devemos encontrar o passo s_l que minimiza a norma do resíduo $r = b - As$ no subespaço afim $s_0 + \mathcal{K}_l(A, r_0)$ (KELLEY, 1995).

O problema de minimização que devemos resolver a cada iteração pode ser escrito como:

$$\begin{aligned} \min \|b - As\| \\ \text{sujeito a } s \in s_0 + \mathcal{K}_l(A, r_0). \end{aligned} \quad (2.7)$$

Figura 14 – Projção do resíduo inicial sobre o subespaço $A\mathcal{K}_l(A, r_0)$



Fonte – Adaptado de (BEGIATO, 2007)

Além disso, para que a norma do resíduo $r_l = b - As_l$ seja mínima é necessário que a projeção de r_l com o subespaço $A\mathcal{K}_l(A, r_0)$ seja ortogonal. Tal constatação pode ser facilmente verificada, admitindo-se, por absurdo, a existência de r_l^* não ortogonal com o subespaço $A\mathcal{K}_l(A, r_0)$ com norma mínima (menor que a norma de r_l). No entanto, nestas condições, seria possível definir um triângulo retângulo com os vetores r_l e r_l^* , o que implicaria em uma contradição pelo Teorema de Pitágoras, o que significa dizer que o resíduo r_l é mínimo quando a projeção de r_l com o subespaço $A\mathcal{K}_l(A, r_0)$ for ortogonal. Sendo assim, teremos um problema de quadrados mínimos. Convém transformar a base atual do subespaço $\mathcal{K}_l(A, r_0)$ em uma base ortonormal e para isso usaremos o processo de Arnoldi, que é um processo de ortonormalização de uma base do subespaço de Krylov.

2.3.1 ITERAÇÃO DE ARNOLDI

A iteração de Arnoldi consiste em um processo semelhante ao Gram-Schmidt, no sentido que gera uma base ortonormal para o subespaço de Krylov. Considerando a matriz A , o resultado do processo de Arnoldi é a obtenção de uma matriz unitária e semelhante à matriz A , tal que $A = VHV^T$ ($\iff AV = VH$). Este processo fornecerá uma versão mais simples para o Problema 2.7, como veremos a seguir.

Queremos uma base ortonormal $\{v_1, v_2, \dots, v_l\}$ do subespaço de Krylov do resíduo inicial (r_0) sobre a matriz A , tal que $\text{span}\{r_0, Ar_0, \dots, A^{l-1}r_0\} = \text{span}\{v_1, v_2, \dots, v_l\}$. Sendo assim, iniciamos o processo com o vetor r_0 , e a partir dele obtemos $v_1 = \frac{r_0}{\|r_0\|}$. O vetor v_2 deve ser obtido de modo que $\text{span}\{v_1, v_2\} = \text{span}\{v_1, Av_1\}$. Além disso, v_2 deve ter norma igual a 1 e deve ser ortogonal a v_1 . Segue daí que:

$$Av_1 = h_{11}v_1 + h_{21}v_2 \iff Av_1 = \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{21} \end{pmatrix}.$$

Impondo condição da ortogonalidade e aplicando produto interno em ambos os lados por v_1 obtemos

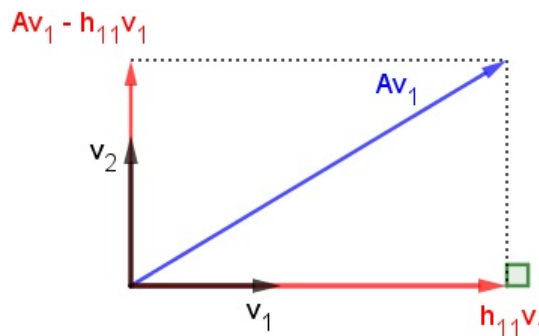
$$\begin{aligned} Av_1 = h_{11}v_1 + h_{21}v_2 &\Rightarrow v_1^T Av_1 = v_1^T h_{11}v_1 + v_1^T h_{21}v_2 \\ &\Rightarrow v_1^T Av_1 = h_{11}v_1^T v_1 + h_{21}v_1^T v_2 \\ &\Rightarrow h_{11} = v_1^T Av_1. \end{aligned}$$

Além disso

$$Av_1 = h_{11}v_1 + h_{21}v_2 \implies h_{21}v_2 = Av_1 - h_{11}v_1.$$

Como v_2 deve ter norma igual a 1, temos $\hat{v} = Av_1 - h_{11}v_1$ e fazendo $h_{21} = \|\hat{v}\|$, decorre que $v_2 = \frac{\hat{v}}{h_{21}}$.

Figura 15 – Construção de uma base ortonormal $\{v_1, v_2, \dots, v_l\}$ do subespaço de Krylov do resíduo inicial (r_0) sobre a matriz A por Arnoldi



Fonte – Autoria própria (2019)

Dado que o objetivo é construir uma base ortonormal para o subespaço, iniciamos a etapa seguinte com o vetor v_2 e calculamos Av_2 , uma vez que $v_2 \in \mathcal{K}_2(A, r_0)$. Continuando este processo, por indução temos que v_1, v_2, \dots, v_j uma base ortonormal para o subespaço de Krylov $\mathcal{K}_j(A, r_0)$. Levando em consideração que cada subespaço de Krylov de dimensão $j + 1$ pode ser obtido a partir da base ortogonal do subespaço de Krylov de dimensão j , usando o conjunto gerador $\{v_1, v_2, \dots, v_j, Av_j\}$, temos que $\mathcal{K}_{j+1}(A, r_0) = \text{span}\{v_1, v_2, \dots, v_j, Av_j\}$.

Formalizamos agora um algoritmo para o procedimento de Arnoldi.

Algoritmo 2.11. *Processo de Arnoldi para Ortonormalização sobre o subespaço de Krylov*

1. $v_1 = \frac{r_0}{\|r_0\|}$;
2. Para $j = 1, 2, \dots, l$, faça:
 - (a) $\hat{v} = Av_j$;
 - (b) Para $i = 1, 2, \dots, j$ faça:

- $h_{ij} = v_i^T \hat{v}$;
- $\hat{v} = \hat{v} - \sum_{i=1}^j h_{ij} v_i$;

(c) $h_{j+1,j} = \|\hat{v}\|$; Se, $h_{j+1,j} = 0$, então, o processo deve ser interrompido (pela impossibilidade da realização do próximo passo).

(d) $v_{j+1} = \frac{\hat{v}}{h_{j+1,j}}$.

Evidentemente pode-se concluir pelo Corolário 2.10 que, se o polinômio mínimo de r_0 com relação à matriz A for de grau $j < n$ não será possível construir uma base ortogonal com dimensão maior que j .

Proposição 2.12. *O Algoritmo 2.11 falha na j -ésima iteração (isto é, $h_{j+1,j} = 0$ no passo 2d) se, e somente se, o polinômio mínimo de r_0 com respeito à matriz A tem grau j .*

Demonstração. Se o Algoritmo 2.11 falha na j -ésima iteração, temos que:

$$\begin{aligned} h_{j+1,j} = 0 &\Leftrightarrow \|\hat{v}\| = 0 \\ &\Leftrightarrow \|Av_j - \sum_{i=1}^j h_{ij}v_i\| = 0 \\ &\Leftrightarrow Av_j = \sum_{i=1}^j h_{ij}v_i \\ &\Leftrightarrow Av_j = h_{1j}v_1 + h_{2j}v_2 + \dots + h_{jj}v_j \end{aligned}$$

O que significa dizer que $Av_j \in \text{span}\{v_1, v_2, \dots, v_j\}$ e, conseqüentemente, o polinômio mínimo de r_0 com respeito à matriz A tem grau j . \square

Exemplo 2.13. Considere o sistema cuja forma matricial é dada por

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_s = \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}}_b. \quad (2.8)$$

Tomando $s_0 = \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ arbitrário como possível candidato a solução, temos que o resíduo correspondente é

$$r_0 = b - As_0 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Queremos uma base ortonormal do subespaço de Krylov do resíduo inicial $r_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

sobre a matriz $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. Assim, aplicando o Algoritmo 2.11 (processo de ortonormalização de Arnoldi), temos:

$$1. v_1 = \frac{r_0}{\|r_0\|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix};$$

2. Para $j = 1$

$$(a) \hat{v} = Av_1 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix};$$

(b) Para $i = 1$:

$$\bullet h_{11} = v_1^T \hat{v} = (0 \ 0 \ 1) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = 1;$$

$$\bullet \hat{v} = \hat{v} - h_{11}v_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix};$$

$$(c) h_{21} = \|\hat{v}\| = \left\| \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{2};$$

$$(d) v_2 = \frac{\hat{v}}{h_{21}} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

3. Para $j = 2$

$$(a) \hat{v} = Av_2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix};$$

(b) Para $i = 1, 2$:

$$\bullet h_{12} = v_1^T \hat{v} = (0 \ 0 \ 1) \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = -\frac{1}{\sqrt{2}};$$

$$\bullet h_{22} = v_2^T \hat{v} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 0;$$

$$\bullet \hat{v} = \hat{v} - h_{12}v_1 - h_{22}v_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} - \left(-\frac{1}{\sqrt{2}}\right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 0 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix};$$

$$\text{(c) } h_{32} = \|\hat{v}\| = \left\| \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\| = 1;$$

$$\text{(d) } v_3 = \frac{\hat{v}}{h_{32}} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

4. Para $j = 3$

$$\text{(a) } \hat{v} = Av_3 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix};$$

(b) Para $i = 1, 2, 3$:

$$\bullet h_{13} = v_1^T \hat{v} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = -\frac{1}{\sqrt{2}};$$

$$\bullet h_{23} = v_2^T \hat{v} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 1;$$

$$\bullet h_{33} = v_3^T \hat{v} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 0;$$

$$\bullet \hat{v} = \hat{v} - h_{13}v_1 - h_{23}v_2 - h_{33}v_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} - \left(-\frac{1}{\sqrt{2}}\right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} - 0 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix};$$

$$(c) \quad h_{43} = \|\hat{v}\| = \left\| \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\| = 0;$$

(d) Não realizado.

Portanto, uma base ortonormal do subespaço de Krylov do resíduo inicial $r_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

sobre a matriz $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ é dada pelo conjunto

$$\alpha = \{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\}.$$

Perceba que, de fato α é uma base, pois além de gerar o subespaço de Krylov do resíduo inicial r_0 sobre a matriz A , é linearmente independente, visto que

$$a \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} + c \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow a = b = c = 0.$$

Além disso, α é ortonormal, pois

$$\left\| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\| = 1$$

e,

$$\left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\rangle = 0.$$

Supondo que o algoritmo não apresentou falha de execução, após a l -ésima iteração, temos os vetores ortonormais v_1, v_2, \dots, v_{l+1} . Definimos então as matrizes $V_l : n \times l$ e $V_{l+1} : n \times (l+1)$, cujas colunas são dadas por esses vetores. Além disso, também podemos definir a matriz $H_l : (l+1) \times l$ Hessenberg superior, onde:

$$h_l(i, j) = \begin{cases} 0, & \text{se } i > j + 1 \\ h_{ij}, & \text{se } i \leq j + 1 \end{cases} \quad (2.9)$$

o que nos permite construir a seguinte relação:

$$AV_l = V_{l+1}H_l \quad (2.10)$$

Exemplificando,

$$A \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_l \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | & | \\ v_1 & v_2 & \dots & v_l & v_{l+1} \\ | & | & & | & | \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1l} \\ h_{21} & h_{22} & \dots & h_{2l} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix}$$

Exemplo 2.14. Considere novamente o sistema (2.8) exposto no Exemplo 2.13, dado por

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_s = \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}}_b.$$

com $s_0 = \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ arbitrário como possível candidato a solução, onde seu resíduo correspondente é $r_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Pelo exposto no desenvolvimento do referido exemplo, podemos definir as matrizes $V_2 : 3 \times 2$ e $V_3 : 3 \times 3$, cujas colunas são dadas pelos vetores ortonormais v_1, v_2, v_3 obtidos por Arnoldi. Além disso, conforme a Equação (2.9) vamos definir a matriz $H_2 : 3 \times 2$ Hessenberg superior, de modo que vale a relação $AV_2 = V_3H_2$ (2.10).

De fato, temos que

$$\begin{aligned} AV_2 &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -\frac{1}{\sqrt{2}} \\ 1 & -\frac{1}{\sqrt{2}} \\ 1 & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} \\ \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix} = V_3H_2. \end{aligned}$$

Considerando a Equação (2.7) onde temos que $s \in s_0 + \mathcal{K}_l(A, r_0)$ e o fato de as colunas de V_l formar uma base para o subespaço $\mathcal{K}_l(A, r_0)$, então existe $y \in \mathbb{R}^l$ tal que $s = s_0 + V_ly$. Assim, podemos escrever:

$$\begin{aligned} \|b - As\| &= \|b - A(s_0 + V_ly)\| \\ &= \|b - As_0 - AV_ly\| \\ &= \|r_0 - AV_ly\| \\ &\stackrel{\text{Equação 2.10}}{=} \|r_0 - V_{l+1}H_ly\|. \end{aligned}$$

Mas, pelo Algoritmo 2.11, $v_1 = \frac{r_0}{\|r_0\|}$ e, considerando-se $\beta = \|r_0\|$, teremos:

$$\begin{aligned} \|b - As\| &= \|r_0 - V_{l+1}H_l y\| \\ &= \|\beta v_1 - V_{l+1}H_l y\| \\ &= \|V_{l+1}\beta e_1 - V_{l+1}H_l y\| \\ &= \|V_{l+1}(\beta e_1 - H_l y)\|. \end{aligned}$$

Contudo, temos que V_{l+1} é uma matriz ortogonal, uma vez que as suas colunas são vetores ortonormais, logo preserva norma e podemos então reescrever o problema (2.7) como:

$$\begin{aligned} \min \|\beta e_1 - H_l y\| \\ \text{sujeito a : } y \in \mathcal{R}^l \end{aligned} \quad (2.11)$$

o que nos permite definir o seguinte algoritmo:

Algoritmo 2.15. GMRES

1. Defina $r_0 = b - As_0$, $p = \|r_0\|$, $\beta = p$, $v_1 = \frac{r_0}{\beta}$ e $k = 0$;
2. Enquanto $k \leq k_{max}$ e $p > \varepsilon\|b\|$, faça:
 - (a) $k = k + 1$;
 - (b) Executar o passo 2 do processo de Arnoldi (Algoritmo 2.11) com $j = k$;
 - (c) Defina $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^{k+1}$;
 - (d) Encontre $y \in \mathbb{R}^k$ que minimiza $\|\beta e_1 - H_k y\|$;
 - (e) $p = \|\beta e_1 - H_k y_k\|$.
3. $x^* = s_0 + V_k y_k$.

2.3.2 ROTAÇÕES DE GIVENS

Continuando com o objetivo de se obter um sistema computacionalmente mais simples de se resolver com relação ao sistema original, dado que a matriz H_l é Hessenberg superior, podemos encontrar facilmente a sua fatoração QR através de **rotações de Givens**. O nosso interesse para aplicar as rotações de Givens é transformar a matriz de Hessenberg numa triangular superior, anulando os elementos $h_{i+1,i}$, para $i = 1, \dots, l$.

Iremos fatorar a matriz de Hessenberg da seguinte maneira:

$$G_l \dots G_2 G_1 H_l = R_l \quad (2.12)$$

onde R_l é a matriz triangular superior resultante e G_i é a matriz de rotação de Givens para eliminar o elemento $h_{i+1,i}$, que deverá ter a seguinte forma:

$$G_i = \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ 0 & \ddots & \ddots & & & \\ & \ddots & 1 & 0 & & \\ & & 0 & c_i & -s_i & \vdots \\ \vdots & & & s_i & c_i & 0 \\ & & & 0 & 1 & \ddots \\ & & & & \ddots & \ddots & 0 \\ 0 & & \dots & & 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow \text{linha } i \\ \leftarrow \text{linha } i + 1 \end{array} \quad (2.13)$$

sendo $c_i = \cos \theta$, $s_i = -\sin \theta$ e

$$s_i = -\frac{h_{i+1,i}}{\sqrt{(h_{ii}^{(i-1)})^2 + (h_{i+1,i})^2}} \quad e \quad c_i = \frac{h_{ii}^{(i-1)}}{\sqrt{(h_{ii}^{(i-1)})^2 + (h_{i+1,i})^2}} \quad (2.14)$$

Essa multiplicação irá produzir uma rotação de ângulo $-\theta$, com relação ao eixo x , no vetor ao qual for aplicada. Note que usamos a notação $h_{ii}^{(i-1)}$, isso porque esse elemento já foi alterado, $i - 1$ vezes, pela aplicação das matrizes $G_1 \dots G_{i-1}$.

Exemplo 2.16. Vamos transformar a matriz de Hessenberg

$$H_l = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1l} \\ h_{21} & h_{22} & \dots & h_{2l} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix}$$

numa triangular superior utilizando as rotações de Givens.

Primeiramente valor zera o elemento h_{21} utilizando a matriz de rotação:

$$G_1 = \begin{pmatrix} c_1 & -s_1 & 0 & \dots & 0 \\ s_1 & c_1 & 0 & & \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix}$$

Assim, por (2.12)

$$\begin{aligned}
G_1 H_l &= \begin{pmatrix} c_1 & -s_1 & 0 & \dots & 0 \\ s_1 & c_1 & 0 & & \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1l} \\ h_{21} & h_{22} & \dots & h_{2l} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} \\
&= \begin{pmatrix} \frac{(h_{11})^2 + (h_{21})^2}{\sqrt{(h_{11})^2 + (h_{21})^2}} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ \frac{h_{11}h_{21} - h_{11}h_{21}}{\sqrt{(h_{11})^2 + (h_{21})^2}} & h_{22}^{(1)} & \dots & h_{2l}^{(1)} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} = \begin{pmatrix} \sqrt{(h_{11})^2 + (h_{21})^2} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ 0 & h_{22}^{(1)} & \dots & h_{2l}^{(1)} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} \\
&= \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ 0 & h_{22}^{(1)} & \dots & h_{2l}^{(1)} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix}.
\end{aligned}$$

Vamos zerar agora o elemento h_{32} utilizando a matriz de rotação:

$$G_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & c_1 & -s_1 & \\ \vdots & s_1 & c_1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Utilizando procedimento análogo, temos

$$\begin{aligned}
G_2 G_1 H_l &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & c_1 & -s_1 & \\ \vdots & s_1 & c_1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ 0 & h_{22}^{(1)} & \dots & h_{2l}^{(1)} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} \\
&= \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ 0 & \frac{(h_{22}^{(1)})^2 + (h_{32})^2}{\sqrt{(h_{22}^{(1)})^2 + (h_{32})^2}} & \dots & h_{2l}^{(2)} \\ 0 & \frac{h_{22}^{(1)}h_{32} - h_{22}^{(1)}h_{32}}{\sqrt{(h_{22}^{(1)})^2 + (h_{32})^2}} & \dots & h_{3l}^{(2)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} = \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & \dots & h_{1l}^{(1)} \\ 0 & h_{22}^{(2)} & \dots & h_{2l}^{(2)} \\ 0 & 0 & \dots & h_{3l}^{(2)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix}
\end{aligned}$$

Continuando o processo $l - 2$ vezes obteremos uma matriz triangular superior equivalente à matriz de Hessenberg

$$H_l = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1l} \\ h_{21} & h_{22} & \dots & h_{2l} \\ 0 & h_{32} & \dots & h_{3l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{l+1,l} \end{pmatrix} \Leftrightarrow \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1l} \\ 0 & h_{22} & \dots & h_{2l} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{ll} \\ 0 & \dots & 0 & 0 \end{pmatrix} = G_l \dots G_2 G_1 H_l = R_l.$$

Exemplo 2.17. Seja a matriz:

$$A = \begin{pmatrix} 6 & 5 & 0 \\ 5 & 1 & 4 \\ 0 & 4 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

Vamos obter uma matriz triangular superior equivalente utilizando as rotações de Givens. Sendo assim, primeiramente valor zerar o elemento $(2, 1)$ utilizando a matriz de rotação:

$$G_1 = \begin{pmatrix} c & -s & 0 & 0 \\ s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Assim,

$$G_1 A = A_1 = \begin{pmatrix} c & -s & 0 & 0 \\ s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 6 & 5 & 0 \\ 5 & 1 & 4 \\ 0 & 4 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

onde,

$$\begin{aligned} r &= \sqrt{6^2 + 5^2} \approx 7,8102, \\ c &= \cos \theta = \frac{6}{r} \approx 0,7682, \\ s &= \operatorname{sen} \theta = -\frac{5}{r} \approx -0,6402. \end{aligned}$$

Substituindo, temos:

$$A_1 = \begin{pmatrix} 7,8102 & 4,4813 & 2,5607 \\ 0 & -2,4327 & 3,0729 \\ 0 & 4 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

Agora vamos zerar o elemento (3, 2) utilizando um processo análogo.

Com a matriz de rotação

$$G_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & -s & 0 \\ 0 & s & c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

temos

$$G_2 A_1 = A_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & -s & 0 \\ 0 & s & c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 7,8102 & 4,4813 & 2,5607 \\ 0 & -2,4327 & 3,0729 \\ 0 & 4 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

com,

$$\begin{aligned} r &= \sqrt{(-2,4327)^2 + 4^2} \approx 4,6817, \\ c &= \cos \theta = \frac{-2,4327}{r} \approx -0,5196, \\ s &= \sin \theta = -\frac{4}{r} \approx -0,8544. \end{aligned}$$

Logo, substituindo

$$A_2 = \begin{pmatrix} 7,8102 & 4,4813 & 2,5607 \\ 0 & 4,6817 & 0,9664 \\ 0 & 0 & -4,1843 \\ 0 & 0 & 2 \end{pmatrix}.$$

Para finalizar o processo, vamos zerar o elemento (4, 3) utilizando um processo análogo.

Com a matriz de rotação

$$G_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & -s \\ 0 & 0 & s & c \end{pmatrix}$$

temos

$$G_3 A_2 = A_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & -s \\ 0 & 0 & s & c \end{pmatrix} \begin{pmatrix} 7,8102 & 4,4813 & 2,5607 \\ 0 & 4,6817 & 0,9664 \\ 0 & 0 & -4,1843 \\ 0 & 0 & 2 \end{pmatrix}$$

com,

$$\begin{aligned} r &= \sqrt{(-4,1843)^2 + 2^2} \approx 4,6377, \\ c &= \cos \theta = \frac{-4,1843}{r} \approx -0,9022, \\ s &= \sin \theta = -\frac{2}{r} \approx -0,4312. \end{aligned}$$

Logo, substituindo

$$A_3 = \begin{pmatrix} 7,8102 & 4,4813 & 2,5607 \\ 0 & 4,6817 & 0,9664 \\ 0 & 0 & 4,6377 \\ 0 & 0 & 0 \end{pmatrix}.$$

Para facilitar a resolução do problema (2.11), vamos definir $Q_l = G_1^T G_2^T \dots G_l^T$. Perceba que G_i é ortogonal para todo i , pois, sabendo que $c_i^2 + s_i^2 = 1$, temos que

$$\begin{aligned} G_i G_i^T &= \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ 0 & \ddots & \ddots & & & \\ & \ddots & 1 & 0 & & \\ & & 0 & c_i^2 + s_i^2 & s_i c_i - s_i c_i & \vdots \\ \vdots & & s_i c_i - s_i c_i & c_i^2 + s_i^2 & 0 & \\ & & & 0 & 1 & \ddots \\ & & & & \ddots & \ddots & 0 \\ 0 & & \dots & & & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & & \dots & 0 \\ 0 & \ddots & & & \\ & & 1 & \ddots & \vdots \\ \vdots & & \ddots & 1 & \\ & & & \ddots & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix} = I. \end{aligned}$$

Além disso, temos que $Q_l = G_1^T G_2^T \dots G_l^T$ também é ortogonal. Para provarmos esta afirmação, basta lembrarmos que $(AB)^T = B^T A^T$ e que $(A^T)^T = A$. Assim, inicialmente

$$\begin{aligned} (G_1^T G_2^T)(G_1^T G_2^T)^T &= G_1^T G_2^T (G_2^T)^T (G_1^T)^T \\ &= G_1^T G_2^T G_2 G_1 \\ &= G_1^T (G_2^T G_2) G_1 \\ &= G_1^T G_1 = I. \end{aligned}$$

O que significa dizer que $G_1^T G_2^T$ é ortogonal. Continuando o processo, temos que $G_1^T G_2^T G_3^T$ também é ortogonal, pois

$$\begin{aligned} ((G_1^T G_2^T) G_3^T)((G_1^T G_2^T) G_3^T)^T &= (G_1^T G_2^T) G_3^T (G_3^T)^T (G_1^T G_2^T)^T \\ &= (G_1^T G_2^T) G_3^T G_3 (G_1^T G_2^T)^T \\ &= (G_1^T G_2^T) (G_3^T G_3) (G_1^T G_2^T)^T \\ &= (G_1^T G_2^T) (G_1^T G_2^T)^T = I. \end{aligned}$$

Portanto, continuando este processo $l - 2$ vezes, temos que

$$Q_l Q_l^T = (G_1^T G_2^T \dots G_l^T)(G_1^T G_2^T \dots G_l^T)^T = I$$

provando o resultado.

Neste contexto, considerando que $Q_l H_l = R_l$ e que Q_l é ortogonal e também preserva norma, o problema (2.11) pode ser escrito como

$$\|b - As\| = \|\beta e_1 - H_l y\| = \|\beta Q_l e_1 - R_l y\|. \quad (2.15)$$

Para compreender as vantagens proporcionadas pelas rotações de Givens, convém observar os resultados do seguinte Teorema 2.18:

Teorema 2.18. *Considere $R_l = G_l H_l$ como definido anteriormente e $g = \beta Q_l e_1$. Denotemos por \hat{R}_l a matriz triangular superior $l \times l$, obtida eliminando-se a última linha de R_l e \hat{g} um vetor de dimensão l encontrado eliminando-se o elemento $g_{l+1,1}$ (a última linha) de g . Então podemos dizer que:*

1. *O posto de AV_l é igual ao posto de \hat{R}_l . Em particular, se $r_{ll} = 0$ então A deve ser singular;*
2. *o vetor y^* que é o minimizador da equação $\|\beta e_1 - H_l y\|$ é dado por*

$$y^* = \hat{R}_l^{-1} \hat{g}; \quad (2.16)$$

3. *a norma do resíduo no passo l será dada por $g_{l+1,1}$.*

Demonstração. Para provar a primeira parte considere que:

$$AV_l = V_{l+1} Q_l^t Q_l H_l \Rightarrow AV_l = V_{l+1} Q_l^t R_l. \quad (2.17)$$

Visto que tanto Q_l quanto V_{l+1} são unitárias (norma é igual a um), temos que posto de AV_l é igual ao posto de R_l , que é igual ao posto de \hat{R}_l , já que a última linha de R_l é nula. Como R_l é triangular superior, se r_{ll} então R_l é singular, e portanto A é singular. A segunda e a terceira parte saem da Equação (2.10). Pois:

$$\|\beta e_1 - H_l y\| = \|\beta Q_l e_1 - R_l y\| = |g_{l+1,1}| + \|\hat{g} - R_l y\|. \quad (2.18)$$

O resíduo mínimo é encontrado zerando o segundo elemento da parte esquerda da Equação 2.18. Logo $y^* = \hat{R}_l^{-1} \hat{g}$ e a norma do resíduo será dada por $g_{l+1,1}$. \square

Esse resultado é um fator estimulante para a implementação pois, garantida a não-singularidade da matriz A , a aproximação para a solução do sistema linear não precisaria ser obtida a cada iteração do GMRES, mas somente quando tivermos um resíduo suficientemente pequeno, já que a norma do resíduo será dada pelo módulo do último elemento do vetor $g = \beta Q_l e_1$.

Ou seja, podemos reescrever o Algoritmo 2.15:

Algoritmo 2.19. *GMRES com Rotações de Givens*

1. Defina $r_0 = b - Ax_0$, $p = \|r_0\|$, $\beta = p$, $v_1 = \frac{r_0}{\beta}$, $Q = I : n \times n$ e $k = 0$;
2. Enquanto $k \leq kmax$ e $p > \varepsilon\|b\|$, faça:
 - (a) $k = k + 1$;
 - (b) Executar o passo 2 do processo de Arnoldi (Algoritmo 2.11) com $j = k$;
 - (c) Se $k > 1$, aplique Q_k na k -ésima coluna de H_k ;
 - (d) Faça: $s_k = -\frac{h_{k+1,k}}{\sqrt{(h_{kk})^2 + (h_{k+1,k})^2}}$ e $c_k = \frac{h_{kk}}{\sqrt{(h_{kk})^2 + (h_{k+1,k})^2}}$;
 - (e) $h_{kk} = c_k h_{kk} - s_k h_{k+1,k}$ e $h_{k+1,k} = 0$;
 - (f) Defina G_k como em (2.13) e faça $g = G_k g$;
 - (g) Faça $Q = G_k Q$ e defina Q_k a matriz quadrada correspondente às k primeiras linhas de Q ;
 - (h) $p = |g_{k+1}|$.
3. Considere R a matriz quadrada correspondente às k primeiras linhas de H e w o vetor correspondente às k primeiras linhas de g .
4. Encontre y_k que resolve o sistema: $Ry = w$.
5. $x^* = x_0 + V_k y_k$.

O Algoritmo 2.19 pode apresentar falha no passo 2b. Conforme a Corolário 2.10, isso acontecerá se, e somente se, o polinômio mínimo de r_0 em relação à matriz A tiver grau $c < n$. A Proposição 2.20, citada a seguir, garante que, neste caso, teremos encontrado a solução exata do sistema, é o que chamam na literatura de *lucky breakdown*.

Proposição 2.20. *Seja A uma matriz não-singular, então o Algoritmo 2.19 irá falhar na l -ésima iteração (isto é, $h_{l+1,l} = 0$) se, e somente se, s_k é a solução exata do sistema $As = b$.*

Demonstração. Considere que $h_{l+1,l} = 0$. Da não-singularidade de A temos, pela primeira parte do Teorema 2.18 que $\hat{R}_{ll} = h_{ll}^{l-1} \neq 0$. Sendo assim por (2.14) temos que $s_l = 0$. O vetor g atual é obtido através do produto da matriz G_l pelo vetor g anterior aumentado em uma dimensão. Chamando de $\bar{g} = (\bar{g}_1, \dots, \bar{g}_l, 0)^T$ este vetor g anterior, temos que $g = (\bar{g}_1, \dots, c_l \bar{g}_l, s_l \bar{g}_l)$. Pela última parte do Teorema 2.18, tem-se que $\|r_l\| = |g_{l+1}| = |s_l \bar{g}_l| = 0$.

Para provar que a condição é suficiente, devemos notar que se a solução exata é encontrada no passo l e não no passo $l - 1$, logo vamos ter $\|r_l\|$ e então $|g_{l+1}| = |s_l \bar{g}_l| = 0$, o que implica que $s_l = 0$ e, por (2.14), $h_{l+1,l} = 0$. \square

Teorema 2.21. *Seja A uma matriz $n \times n$ não-singular. Então o Algoritmo 2.19 encontrará a solução para o problema linear $As = b$ em no máximo n iterações.*

Demonstração. Ver (KELLEY, 1995). □

Exemplo 2.22. Considere novamente o Sistema (2.8) trabalhado nos Exemplos 2.13 e 2.14 , dado por

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_s = \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}}_b,$$

com $s_0 = \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ arbitrário como possível candidato a solução, onde seu resíduo correspondente é $r_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Nosso objetivo agora é aplicar o método GMRES a fim de obter a solução desejada, com a aplicação das rotações de Givens.

No Exemplo 2.13 aplicamos o algoritmo de Arnoldi e obtivemos uma base ortonormal de r_0 em relação a matriz A . Na ocasião, verificou-se uma falha no passo 2d da terceira iteração. Consequentemente, ao aplicarmos o Algoritmo 2.19, haverá falha na terceira iteração no passo 2b, o que significa dizer, pela Proposição 2.20 e Teorema 2.21, que nesta terceira iteração obteremos a solução exata do sistema encontrando $y \in \mathbb{R}^3$ que minimiza $\|\beta Q_2 e_1 - R_3 y\|$.

Diante do exposto, fazendo a terceira iteração do Algoritmo 2.19, a fim de obter a solução x^* , tal que

$$\underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{x^*} = \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}}_b,$$

temos

$$1. \text{ Defina } r_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad p = \|r_0\| = \left\| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\| = 1, \quad \beta = 1, \quad v_1 = \frac{r_0}{\beta} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$v_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{e } k = 2;$$

2. • Defina a Matriz $H_3 : 4 \times 3$ de Hessenberg superior dada por

$$\begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \sqrt{2} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- Defina a Matriz \bar{H}_3 , dada pelo produto $G_2G_1H_3$, onde G_2 e G_1 são as matrizes de rotação para zerar respectivamente os elementos h_{32} e h_{21} da matriz H_3 .

$$\begin{aligned} \bar{H}_3 &= G_2G_1H_3 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 \\ -\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \sqrt{2} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 \\ -\frac{\sqrt{2}}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

- Defina

$$\begin{aligned} g &= \beta G_2G_1e_1 \\ &= 1 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 \\ -\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{\sqrt{2}}{2\sqrt{3}} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}. \end{aligned}$$

3. Considerando R a matriz quadrada correspondente às 3 primeiras linhas de \bar{H}_3 e w o vetor correspondente às 3 primeiras linhas de g , temos

$$R = \begin{pmatrix} \frac{3}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & -1 \end{pmatrix} \quad e \quad w = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{\sqrt{2}}{2\sqrt{3}} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

4. Resolvendo o sistema $Ry = w$, com $y \in \mathbb{R}^3$, temos

$$\begin{pmatrix} \frac{3}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{\sqrt{2}}{2\sqrt{3}} \\ \frac{\sqrt{2}}{2} \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

5. Por fim, a solução do problema é dada pelo vetor

$$\begin{aligned} x^* &= x_0 + V_3 y \\ &= \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

2.4 MÉTODO DOS GRADIENTES CONJUGADOS

O Método dos Gradientes Conjugados (*Conjugate Gradient Method*) foi criado na década de 50 do século passado por Hestenes e Steifel (HESTENES; STIEFEL, 1952) para resolver sistemas lineares $As = b$ onde A é uma matriz simétrica e definida positiva, isto é,

$$A = A^T \quad e \quad x^T Ax > 0 \quad \text{para todo } x \neq 0 \quad (2.19)$$

Inicialmente desenvolvido como um método direto, posteriormente foi verificada uma forte viabilidade de utilizá-lo também como um método iterativo, principalmente em problemas de grande porte.

Considere o sistema

$$As = b, \quad (2.20)$$

onde A é uma matriz simétrica e definida positiva e $x^* = A^{-1}b$ é solução. A k -ésima iteração do Método dos Gradientes Conjugados consiste em minimizar a função

$$f(x) = \frac{1}{2}x^T Ax - x^T b. \quad (2.21)$$

Note que se $f(\bar{x})$ é o valor mínimo (em \mathbb{R}^n) então

$$\nabla f(\bar{x}) = A\bar{x} - b = 0 \quad (2.22)$$

e portanto $\bar{x} = x_*$.

Definição 2.23. *Seja A uma matriz de ordem n simétrica e definida positiva e seja x um vetor não nulo de \mathbb{R}^n , a função $\|\cdot\|_A$, denominada A -norma em \mathbb{R}^n , é definida como*

$$\|x\|_A = \sqrt{x^T Ax}. \quad (2.23)$$

Lema 2.24. *Seja $S \subset \mathbb{R}^n$. Se x_k minimiza f então x_k também minimiza $\|x_* - x\|_A = \|r\|_{A^{-1}}$ sobre S , onde r é o resíduo. (2.1).*

Demonstração. Seja A simétrica e $Ax_* = b$. temos que:

$$\begin{aligned}
 \|x_* - x\|_A^2 &= (x_* - x)^T A(x_* - x) \\
 &= ((x_*)^T - x^T)A(x_* - x) \\
 &= (x_*)^T A(x_* - x) - x^T A(x_* - x) \\
 &= (x_*)^T Ax_* - (x_*)^T Ax - x^T Ax_* + x^T Ax \\
 &= (x_*)^T Ax_* - 2x^T Ax + x^T Ax \\
 &= 2f(x) + (x_*)^T Ax_*.
 \end{aligned}$$

Perceba que $(x_*)^T Ax_*$ independe de x , logo minimizar f é equivalente a minimizar $\|x_* - x\|_A^2$, o que consequentemente equivale minimizar $\|x_* - x\|_A$. Ademais, chamando $e = x_* - x$, temos que

$$\begin{aligned}
 \|e\|_A^2 = e^T Ae &= (x_* - x)^T A(x_* - x) \\
 &= (x_* - x)A^T(A^{-1}A)(x_* - x) \\
 &= (A(x_* - x))^T A^{-1}(A(x_* - x)) \\
 &= \|A(x_* - x)\|_{A^{-1}}^2 \\
 &= \|b - Ax\|_{A^{-1}}^2 = \|r\|_{A^{-1}}^2.
 \end{aligned}$$

e, portanto a A -norma do erro é também a A^{-1} -norma do resíduo. \square

Neste contexto, considerando o Sistema 2.20 com $x_* = A^{-1}b$ como sendo a solução e $e_n = x_* - x_n$ o erro no passo n , temos que a iteração do gradiente conjugado pode ser descrita como um sistema de fórmulas de recorrência que gera a sequência única de iteração $\{x_n \in \mathcal{K}_n\}$ com a propriedade que no passo n , $\|e_n\|_A$ é minimizado.

Exemplo 2.25. Considere o sistema $As = b$, onde

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \quad e \quad b = \begin{pmatrix} 2 \\ -8 \end{pmatrix}$$

Perceba que A é simétrica e definida positiva pois, se $s = \begin{pmatrix} x \\ y \end{pmatrix}$ então

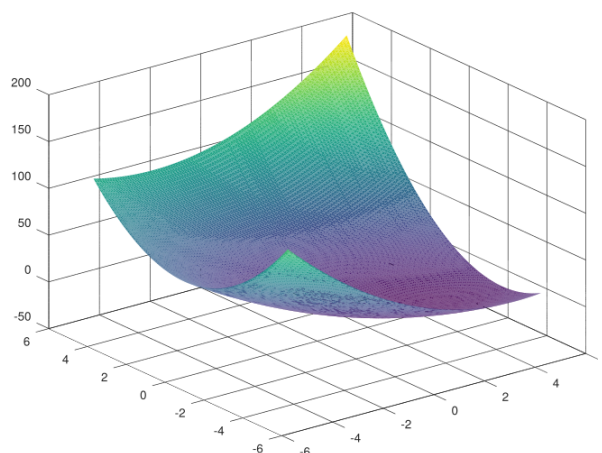
$$\begin{aligned}
 x^T Ax &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3x^2 + 4xy + 6y^2 \\
 &= 2x^2 + (x + y)^2 + 5y^2 > 0.
 \end{aligned}$$

Logo, obter a solução do sistema $Ax = b$ é o mesmo que encontrar as coordenadas (x, y) que minimizam a função

$$\begin{aligned}
 f(x) = \frac{1}{2}x^T Ax - x^T b &= \frac{1}{2} \left[\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right] - \begin{pmatrix} 2 & -8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\
 &= \frac{3}{2}x^2 + 2xy + 3y^2 - 2x + 8y
 \end{aligned}$$

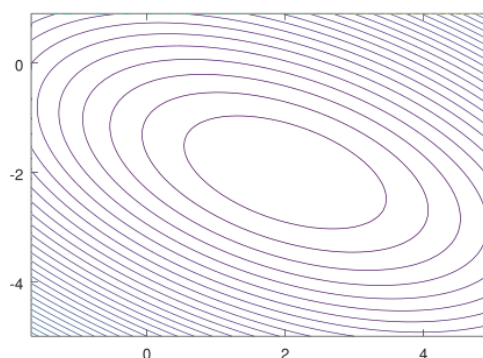
Veja as ilustrações abaixo (Figura 16) e (Figura 17):

Figura 16 – Gráfico de $f(x)$: o ponto mínimo dessa superfície é solução de $As = b$



Fonte – Autoria própria (2019)

Figura 17 – Curvas de nível: interseção da superfície com hiperplanos



Fonte – Autoria própria (2019)

A seguir vamos deduzir um conjunto de fórmulas de recorrência que compreendem o método dos gradientes conjugados. Para isto, seguiremos o mesmo raciocínio exposto por (LAGO, 2010) e (SHEWCHUK, 1994).

Suponha que dispomos de um conjunto de vetores ortogonais $S = \{p_0, p_1, \dots, p_{n-1}\}$, onde $p_i \in \mathbb{R}^n$. Chamaremos estes vetores de direções de busca. A princípio seria possível escrever

$$e_0 = x_0 - x_* = - \sum_{i=0}^{n-1} \alpha_i p_i. \quad (2.24)$$

Baseado nesta possibilidade e supondo que conhecemos os α_j , a cada iteração k daremos “um passo α_k ” na direção de busca p_k , obtendo

$$x_{k+1} = x_0 + \sum_{i=0}^k \alpha_i p_i = x_k + \alpha_k p_k. \quad (2.25)$$

Vamos encontrar agora o coeficiente α_k , que define o tamanho do passo a ser dado na direção p_k . Das Equações (2.24) e (2.25), temos que

$$e_k = x_k - x_* = \left(x_0 + \sum_{i=0}^{k-1} \alpha_i p_i \right) - x_* = \sum_{i=0}^{k-1} \alpha_i p_i - \sum_{i=0}^{n-1} \alpha_i p_i = - \sum_{i=k}^{n-1} \alpha_i p_i \quad (2.26)$$

e, devido à ortogonalidade de S , tem-se que $e_k \perp p_j$, para $j < k$. Para facilitar o cálculo de α_j vamos escrever o erro da seguinte forma:

$$e_{k+1} = e_0 + \sum_{i=1}^k \alpha_i p_i = e_k + \alpha_k p_k. \quad (2.27)$$

Forçando a condição de ortogonalidade $e_{k+1} \perp p_k$ e aplicando produto interno a Equação (2.27) teremos

$$0 = \langle p_k, e_{k+1} \rangle = \langle p_k, e_k + \alpha_k p_k \rangle = \langle p_k, e_k \rangle + \alpha_k \langle p_k, p_k \rangle,$$

e conseqüentemente

$$\alpha_k = - \frac{\langle p_k, e_k \rangle}{\langle p_k, p_k \rangle}. \quad (2.28)$$

Mas como não dispomos de e_k , pois caso tivéssemos teríamos resolvido o problema fazendo $x_k - e_k = x_*$. Uma solução seria exigir uma *A-ortogonalidade* das direções de busca, ou seja,

$$\langle p_i, Ap_j \rangle = 0, \quad i \neq j. \quad (2.29)$$

Isto faz com que o erro seja *A-ortogonal* às direções de busca, ao invés da ortogonalidade que mostramos em (2.26). Com isto, da Equação (2.28) e sabendo que $Ae_k = Ax_k - b = -r_k$ então temos que

$$\alpha_k = - \frac{\langle p_k, Ae_k \rangle}{\langle p_k, Ap_k \rangle} = \frac{\langle p_k, r_k \rangle}{\langle p_k, Ap_k \rangle}. \quad (2.30)$$

O que pode ser calculado, pois podemos obter o resíduo fazendo

$$r_{k+1} = -Ae_{k+1} = -A(e_k + \alpha_k p_k) = r_k - \alpha_k Ap_k = r_0 - \sum_{i=0}^k \alpha_i Ap_i. \quad (2.31)$$

Entretanto, cabe notar que, caso A seja não-singular, Ap_k pode ser zero o que impede o cálculo de α_k .

Vamos agora mostrar como encontrar o conjunto de direções de busca de modo que sejam mutuamente *A-ortogonais*. Uma ideia intuitiva seria partir de um conjunto de vetores ortogonais de fácil obtenção, como a base canônica, definir $p_k = e_k^n$ e então *A-ortogonalizar* p_k com os vetores anteriores. Entretanto, se tomarmos os resíduos ao invés da base canônica, ou seja, se fizermos

$$p_k = r_k + \sum_{i=0}^{k-1} \beta_{(k,i)} p_i \quad (2.32)$$

para algum conjunto de $\beta_{(k,j)}$, começando com $p_0 = r_0$, teremos uma facilidade muito grande para o cálculo dos $\beta_{(k,j)}$, como mostraremos a seguir. Primeiramente, consideremos dois teoremas do resíduo.

Teorema 2.26. *No Gradiente Conjugado o resíduo é ortogonal às direções de busca anteriores.*

Demonstração. Multiplicando (2.26) por $-A$ teremos uma expressão para r_k em função das direções de busca p_i :

$$r_k = \sum_{i=k}^{n-1} \alpha_i A p_i. \quad (2.33)$$

Portanto, fazendo o produto interno entre r_k e qualquer p_j onde $j < k$ teremos

$$\langle p_j, r_k \rangle = \left\langle p_j, \sum_{i=k}^{n-1} \alpha_i A p_i \right\rangle = \sum_{i=k}^{n-1} \alpha_i \langle p_j, A p_i \rangle = 0 \quad (2.34)$$

provando que o resíduo é ortogonal às direções de busca anteriores. \square

Teorema 2.27. *Os resíduos são mutuamente ortogonais.*

Demonstração. Vamos fazer o produto interno entre p_j e r_k , que já sabemos que é zero para todo $j < k$

$$0 = \langle p_j, r_k \rangle = \left\langle r_j + \sum_{i=0}^{j-1} \beta_{(j,i)} p_i, r_k \right\rangle = \langle r_j, r_k \rangle + \sum_{i=0}^{j-1} \beta_{(j,i)} \langle p_i, r_k \rangle. \quad (2.35)$$

Do Teorema 2.26 eliminaremos os $\langle p_i, r_k \rangle$, e ficaremos com

$$\langle r_j, r_k \rangle = 0. \quad (2.36)$$

Como $\langle r_j, r_k \rangle = \langle r_k, r_j \rangle = 0$, então temos que isto também será verdade para $j > k$, ou seja, $\langle r_k, r_j \rangle = 0$ para $j \neq k$. \square

De posse dos Teoremas 2.26 e 2.27, nosso objetivo será calcular os $\beta_{(k,j)}$ da Equação (2.32). Faremos o produto interno entre p_k e $A p_j$ (com $j < k$). Da Equação (2.32) obtemos

$$\begin{aligned} \langle p_k, A p_j \rangle &= \left\langle r_k + \sum_{i=0}^{k-1} \beta_{(k,i)} p_i, A p_j \right\rangle = \langle r_k, A p_j \rangle + \sum_{i=0}^{k-1} \beta_{(k,i)} \langle p_i, A p_j \rangle \\ 0 &= \langle r_k, A p_j \rangle + \beta_{(k,j)} \langle p_j, A p_j \rangle \end{aligned}$$

e, portanto

$$\beta_{(k,j)} = -\frac{\langle r_k, A p_j \rangle}{\langle p_j, A p_j \rangle}. \quad (2.37)$$

Agora vamos finalmente mostrar que o cálculo dos $\beta_{(k,j)}$ pode ser significativamente simplificado. Aplicando produto interno por r_k na Equação (2.31) temos

$$\begin{aligned} \langle r_k, r_{j+1} \rangle &= \langle r_k, r_j \rangle - \alpha_j \langle r_k, A p_j \rangle \\ \alpha_j \langle r_k, A p_j \rangle &= \langle r_k, r_j \rangle - \langle r_k, r_{j+1} \rangle. \end{aligned}$$

Pela ortogonalidade mostrada no Teorema 2.26 concluímos que, se $j = k$

$$\langle r_k, Ap_k \rangle = \frac{\langle r_k, r_k \rangle - \langle r_k, r_{k+1} \rangle}{\alpha_k} = \frac{\langle r_k, r_k \rangle}{\alpha_k}, \quad (2.38)$$

e que caso $j = k - 1$, teremos

$$\langle r_k, Ap_{k-1} \rangle = \frac{\langle r_k, r_{k-1} \rangle - \langle r_k, r_k \rangle}{\alpha_{k-1}} = -\frac{\langle r_k, r_k \rangle}{\alpha_{k-1}} \quad (2.39)$$

e todos os demais $\langle r_k, Ap_j \rangle$ serão zero. Como o cálculo de $\beta_{(k,j)}$ só é efetuado para $j < k$, então $\beta_{(k,j)}$ só será diferente de zero para $j = k - 1$. Para simplificar a notação faremos $\beta_{(k,k-1)} = \beta_k$. Assim,

$$p_k = r_k + \beta_k p_{k-1}. \quad (2.40)$$

Com o objetivo de simplificar ainda mais o cálculo de β_k , isolando β_k e aplicando produto interno por Ap_{k-1} na Equação (2.40) temos

$$\begin{aligned} \beta_k p_{k-1} &= p_k - r_k \\ \beta_k \langle p_{k-1}, Ap_{k-1} \rangle &= \langle p_k - r_k, Ap_{k-1} \rangle = \langle p_k, Ap_{k-1} \rangle - \langle r_k, Ap_{k-1} \rangle \\ \beta_k &= -\frac{\langle r_k, Ap_{k-1} \rangle}{\langle p_{k-1}, Ap_{k-1} \rangle}. \end{aligned} \quad (2.41)$$

De posse das Equações (2.39) e (2.30), podemos simplificar a equação (2.41):

$$\begin{aligned} \beta_k &= -\frac{\langle r_k, Ap_{k-1} \rangle}{\langle p_{k-1}, Ap_{k-1} \rangle} = \langle r_k, Ap_{k-1} \rangle \left(-\frac{1}{\langle p_{k-1}, Ap_{k-1} \rangle} \right) \\ &= \left(-\frac{\langle r_k, r_k \rangle}{\alpha_{k-1}} \right) \left(-\frac{1}{\langle p_{k-1}, Ap_{k-1} \rangle} \right) \\ &= \frac{1}{\alpha_{k-1}} \frac{\langle r_k, r_k \rangle}{\langle p_{k-1}, Ap_{k-1} \rangle} \\ &= \frac{\langle p_{k-1}, Ap_{k-1} \rangle}{\langle p_{k-1}, r_{k-1} \rangle} \frac{\langle r_k, r_k \rangle}{\langle p_{k-1}, Ap_{k-1} \rangle} = \frac{\langle r_k, r_k \rangle}{\langle p_{k-1}, r_{k-1} \rangle}. \end{aligned}$$

Mas, aplicando produto interno por r_k na Equação (2.40) temos

$$\langle p_k, r_k \rangle = \langle r_k, r_k \rangle + \beta_k \langle p_{k-1}, r_k \rangle = \langle r_k, r_k \rangle, \quad (2.42)$$

portanto

$$\beta_k = \frac{\langle r_k, r_k \rangle}{\langle r_{k-1}, r_{k-1} \rangle} = \frac{\|r_k\|_2^2}{\|r_{k-1}\|_2^2} \quad (2.43)$$

e

$$\alpha_k = \frac{\langle r_k, r_k \rangle}{\langle p_k, Ap_k \rangle} = \frac{\|r_k\|_2^2}{\langle p_k, Ap_k \rangle}. \quad (2.44)$$

De modo mais geral, as etapas do método gradiente conjugado podem ser divididas em duas partes:

- **Passo inicial:** Selecionar uma estimativa inicial arbitrária x_0 e calcular o resíduo r_0 e a direção p_0 pelas fórmulas (1.a) e (1.b).
- **Rotina geral:** Tendo determinado a estimativa s_i , o resíduo r_i e a direção p_i , calcular s_{i+1} , r_{i+1} e p_{i+1} sucessivamente pelas fórmulas (2.c), (2.d), (2.e), (2.f) e (2.g).

Algoritmo 2.28. *Gradiente Conjugado*

1. Defina A , b , s_0

(a) $r_0 = b - As_0$;

(b) $p_0 = r_0$.

2. Para $i = 0, 1, 2, \dots$ faça:

(c) $\alpha_i = \frac{\|r_i\|^2}{\langle p_i, Ap_i \rangle}$;

(d) $s_{i+1} = x_i + \alpha_i p_i$;

(e) $r_{i+1} = r_i - \alpha_i Ap_i$;

(f) $\beta_{i+1} = \frac{\|r_{i+1}\|^2}{\|r_i\|^2}$;

(g) $p_{i+1} = r_{i+1} + \beta_{i+1} p_i$.

Exemplo 2.29. Considere o sistema

$$\underbrace{\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_s = \underbrace{\begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix}}_b. \quad (2.45)$$

Nosso objetivo é obter a solução do sistema através da aplicação do método dos Gradientes Conjugados.

Perceba inicialmente que A é simétrica (visto facilmente que $A = A^T$) e definida positiva pois, se $x = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ então

$$\begin{aligned} x^T Ax &= \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= 2x^2 - xy - xy + 2y^2 - yz - yz + 2z^2 \\ &= x^2 + x^2 - 2xy + y^2 + y^2 - 2yz + z^2 + z^2 \\ &= x^2 + (x - y)^2 + (y - z)^2 + z^2 > 0. \end{aligned}$$

Logo, aplicando o Algoritmo 2.28,

1. Defina $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, $b = \begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix}$, $s_0 = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}$.

(a) $r_0 = b - As_0 = \begin{pmatrix} -3 \\ 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$;

(b) $p_0 = r_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

2. Para $i = 0$:

(c) $\alpha_0 = \frac{\|r_0\|^2}{\langle p_0, Ap_0 \rangle} = \frac{\left\| \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\|^2}{\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle} = \frac{1}{2}$;

(d) $s_1 = s_0 + \alpha_0 p_0 = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$;

(e) $r_1 = r_0 - \alpha_0 Ap_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}$;

(f) $\beta_1 = \frac{\|r_1\|^2}{\|r_0\|^2} = \frac{\left\| \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \right\|^2}{\left\| \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\|^2} = \frac{2}{1} = \frac{1}{2}$;

(g) $p_1 = r_1 + \beta_1 p_0 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

3. Para $i = 1$:

$$(c) \alpha_1 = \frac{\|r_1\|^2}{\langle p_1, Ap_1 \rangle} = \frac{\left\| \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} \right\|^2}{\left\langle \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right\rangle} = 1;$$

$$(d) s_2 = s_1 + \alpha_1 p_1 = \begin{pmatrix} -\frac{3}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} + 1 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix};$$

$$(e) r_2 = r_1 - \alpha_1 Ap_1 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} - 1 \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

E portanto, temos que a solução do Sistema (2.45) é dado pelo vetor $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

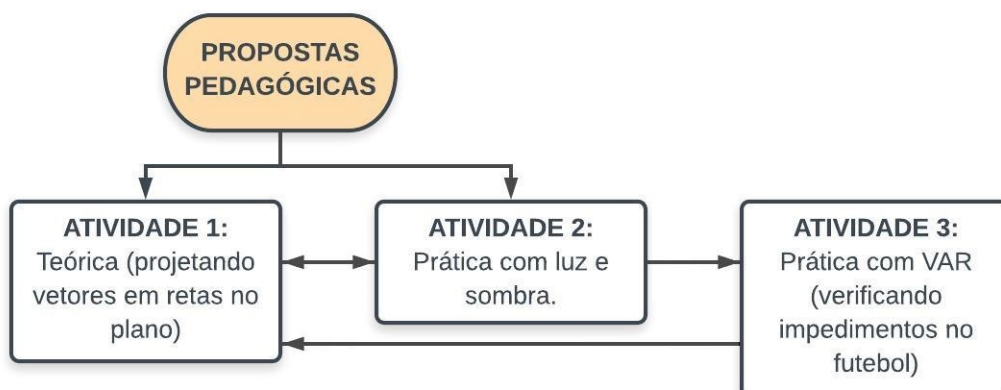
3 PROPOSTAS DE ATIVIDADES PARA USO NO ENSINO MÉDIO

Com base nas Orientações Curriculares para o Ensino Médio (BRASIL, 2006a), a forma do professor trabalhar os conteúdos da disciplina de matemática deve sempre agregar um valor formativo no que diz respeito ao desenvolvimento do pensamento matemático. Isto significa dizer que além do estudante estar envolvido diretamente em sua aprendizagem, é dever do professor colocar os alunos em um processo de aprendizagem que valorize o raciocínio matemático em todos os seus aspectos, seja na formulação de questões e hipóteses, ou mesmo na explicação e dedução de fórmulas utilizadas.

Neste contexto, pretende-se propor atividades para o uso no ensino médio, durante o conteúdo referente à Geometria Analítica, em que o tema *Projeção* seja tratado em diferentes momentos, com o intuito de enriquecer a formação dos alunos. As atividades visam trabalhar projeção em \mathbb{R}^2 e \mathbb{R}^3 , onde serão priorizados conhecimentos geométricos relacionados ao estudo do triângulo retângulo, como as relações trigonométricas e Teorema de Pitágoras, com uma introdução a linguagem vetorial, sobre o plano cartesiano.

São três as atividades a serem propostas (Figura 18): Uma teórica, onde a intenção é verificar a projeção de vetores em retas no \mathbb{R}^2 , e duas práticas, sendo a primeira relacionada ao problema de calcular o comprimento da sombra de um objeto projetada ortogonalmente por uma lâmpada puntiforme presa ao teto e a segunda relacionada com um exemplo de aplicação real utilizado no esporte, mais precisamente no futebol, através da marcação de impedimentos, onde é utilizada a tecnologia VAR (*video assistant referee*) ou *árbitro assistente de vídeo*.

Figura 18 – Fluxograma das atividades



Fonte – Autoria própria (2019)

Pelo exposto no fluxograma acima, percebe-se que duas atividades são independentes, isto é, o professor pode escolher a ordem que lhe seja conveniente para o trabalho com sua turma.

3.1 MODELAGEM MATEMÁTICA E INTERDISCIPLINARIDADE

Duas características importantes advindas das atividades práticas que constituem importante fundamento na aprendizagem qualitativa são a modelagem matemática e a interdisciplinaridade.

Levando em consideração as propostas de atividades, a modelagem matemática se aplica muito bem a este trabalho, uma vez que para todo problema proposto é nossa intenção fazer o aluno pensar sobre formas de resolvê-lo. Sobre a modelagem matemática as Orientações Curriculares para o Ensino Médio discorrem que

A modelagem matemática, percebida como estratégia de ensino, apresenta fortes conexões com a idéia de resolução de problemas (...). Ante uma situação problema ligada ao “mundo real”, com sua inerente complexidade, o aluno precisa mobilizar um leque variado de competências: selecionar variáveis que serão relevantes para o modelo a construir; problematizar, ou seja, formular o problema teórico na linguagem do campo matemático envolvido; formular hipóteses explicativas do fenômeno em causa; recorrer ao conhecimento matemático acumulado para a resolução do problema formulado, o que, muitas vezes, requer um trabalho de simplificação quando o modelo originalmente pensado é matematicamente muito complexo; validar, isto é, confrontar as conclusões teóricas com os dados empíricos existentes; e eventualmente ainda, quando surge a necessidade, modificar o modelo para que esse melhor corresponda à situação real, aqui se revelando o aspecto dinâmico da construção do conhecimento. (BRASIL, 2006a)

A interdisciplinaridade também constitui um elemento muito importante que pode ser agregado a este trabalho. Conforme os Parâmetros Curriculares Nacionais (BRASIL, 2000) "Na perspectiva escolar, a interdisciplinaridade não tem a pretensão de criar novas disciplinas ou saberes, mas de utilizar os conhecimentos de várias disciplinas para resolver um problema concreto ou compreender um determinado fenômeno sob diferentes pontos de vista".

Neste contexto, as atividades podem ser organizadas na forma de projeto interdisciplinar. Por exemplo, considerando a terceira atividade relacionada a prática com VAR (verificação de impedimentos), pode-se desenvolver uma proposta interdisciplinar envolvendo as disciplinas de Matemática, Física e Educação Física.

Na disciplina de Física os alunos trabalham com aspectos vetoriais no que diz respeito a definição e propriedades; na disciplina de Matemática o professor trabalha aspectos relacionados a projeção; enquanto na disciplina de Educação Física os alunos buscam entender e verificar a definição de impedimento numa perspectiva prática. No somatório desta atividade, o problema de se verificar corretamente impedimento acaba sendo compreendido pelo aluno de forma muito mais completa, ao mesmo tempo em que uma atividade conjunta acaba sendo muito mais atrativa para o estudante.

3.2 ATIVIDADE 1: PROJETANDO VETORES EM RETAS NO PLANO

Público alvo: 3º ano do Ensino Médio

Tempo previsto: 2 aulas

Pré-requisitos: Congruência de triângulos; Teorema de Pitágoras; Equação de retas no plano cartesiano

Recursos: Material de escrita

3.2.1 OBJETIVOS

- Definir vetores em \mathbb{R}^2 ;
- Relacionar uma reta no plano cartesiano com um vetor dado;
- Utilizar propriedades relacionadas ao triângulo retângulo na resolução dos problemas propostos, com o intuito de compreender a definição de projeção ortogonal.

3.2.2 DESENVOLVIMENTO DA PROPOSTA

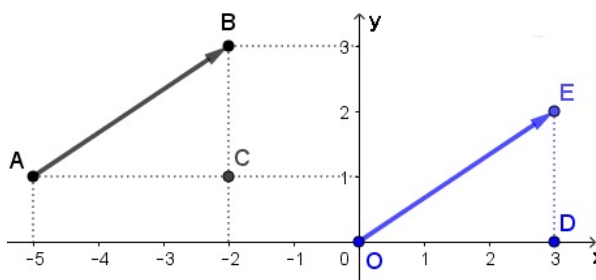
3.2.2.1 LINGUAGEM VETORIAL

Inicialmente faz-se necessário que o professor dê uma noção introdutória sobre a definição de vetores. Para tanto, é sugerido que o professor trabalhe primeiramente com a ideia de **segmento orientado** para vetores. Tal abordagem é mais simples de ser absorvida pelo aluno, uma vez que os mesmos tem maior familiaridade com o termo "segmento", e a noção de vetores vai sendo construída ao longo da atividade. Ao final o professor pode concretizar a definição de vetores, como o conjunto segmentos orientados equipolentes.

Exemplo 3.1. Dados os pontos $A = (-5, 1)$ e $B = (-2, 3)$, vamos definir o segmento orientado que tem A como ponto de partida e B como ponto de chegada. Segue que

$$\overrightarrow{AB} = B - A = (-2, 3) - (-5, 1) = (-2 - (-5), 3 - 1) = (3, 2).$$

Figura 19 – Representação geométrica do vetor definido pelos pontos A e B

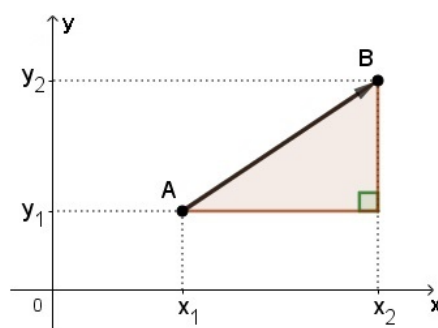


Fonte – Autoria própria (2019)

Com a representação geométrica o professor pode mostrar ao aluno que $\vec{AB} = \vec{OE}$, através da congruência dos triângulos ABC com OED , e justificar a vantagem do trabalho com vetores. Em seguida, sugere-se que o professor repita a atividade, variando os pontos, de modo que os estudantes consigam definir um vetor tendo dois pontos quaisquer.

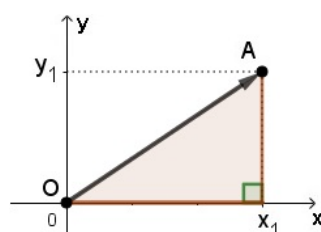
Como estamos trabalhando basicamente com triângulo retângulo e levando em consideração a abordagem de segmento orientado para vetores, o professor tem condições de mostrar que o comprimento de um vetor (**norma**) é dado basicamente pela aplicação do teorema de Pitágoras, o que se resume ao cálculo da distância entre dois pontos (os pontos que definem o vetor em estudo).

Figura 20 – $\|\vec{AB}\| = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$



Fonte – Autoria própria (2019)

Figura 21 – Vetor com extremidade na origem: $\|v\| = d(0, A) = \sqrt{x_1^2 + y_1^2}$



Fonte – Autoria própria (2019)

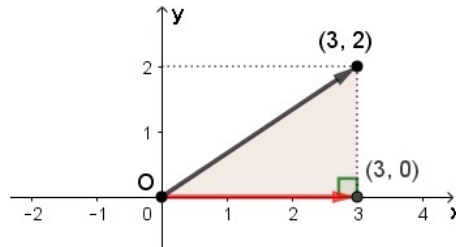
3.2.2.2 PROJEÇÃO DE VETORES SOBRE EIXOS COORDENADOS

Neste momento, no qual os estudantes estão familiarizados com a ideia de vetores, o professor conduzirá um estudo sobre como projetar vetores nos eixos coordenados. Para tanto, após a definição do vetor que se quer trabalhar e o eixo coordenado, é definido um triângulo retângulo, e a partir deste triângulo o professor construirá a ideia de projeção com utilização de propriedades geométricas.

A seguir, alguns exemplos de exercícios que o professor pode realizar com seus alunos.

Exemplo 3.2. A projeção do vetor $(3, 2)$ no eixo x é dada pelo vetor $(3, 0)$. Note que, como a projeção ortogonal é sobre o eixo x , temos que a componente y da coordenada da projeção é zero. Logo, a projeção é definida pela componente x da coordenada do vetor escolhido.

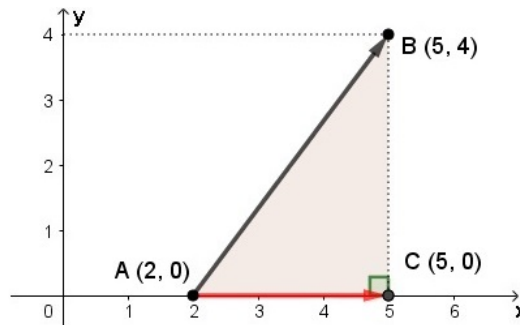
Figura 22 – Representação geométrica da projeção do vetor $(3, 2)$ sobre eixo x



Fonte – Autoria própria (2019)

Exemplo 3.3. Vamos obter a projeção ortogonal do vetor definido pelos pontos $A = (2, 0)$ e $B = (5, 4)$ sobre o eixo x . Neste caso, temos que a projeção desejada é o vetor definido pelos pontos $A = (2, 0)$ e $C = (5, 0)$, isto é, $\overrightarrow{AC} = C - A = (5, 0) - (2, 0) = (3, 0)$.

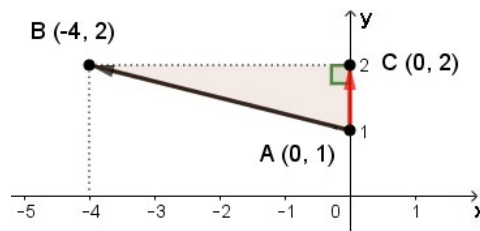
Figura 23 – Representação geométrica da projeção do vetor \overrightarrow{AB} sobre eixo x



Fonte – Autoria própria (2019)

Exemplo 3.4. A projeção do vetor \overrightarrow{AB} definido pelos pontos $A = (0, 1)$ e $B = (-4, 2)$ sobre o eixo y é dada pelo vetor $\overrightarrow{AC} = C - A = (0, 2) - (0, 1) = (0, 1)$.

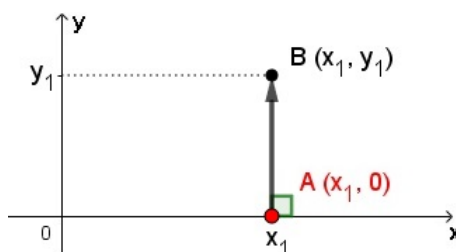
Figura 24 – Representação geométrica da projeção do vetor \overrightarrow{AB} sobre eixo y



Fonte – Autoria própria (2019)

Exemplo 3.5. A projeção ortogonal de qualquer vetor paralelo ao eixo y em relação ao eixo x é o vetor nulo, isto é, apenas um ponto.

Figura 25 – Representação geométrica da projeção de vetor paralelo ao eixo y em relação ao eixo x



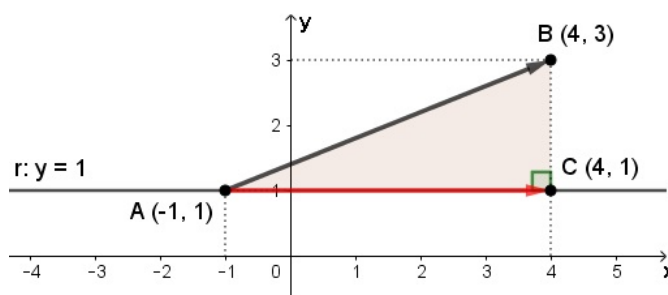
Fonte – Autoria própria (2019)

3.2.2.3 PROJEÇÃO DE VETORES SOBRE RETAS NO PLANO

A ideia aqui é similar à anterior, no entanto agora a intenção é projetar um determinado vetor sobre uma reta qualquer no plano. Para isso, sugerimos que inicialmente se trabalhe alguns exemplos de projeções de vetores sobre retas paralelas aos eixos coordenados, para em seguida formalizar projeções sobre as demais situações de retas. A seguir alguns exemplos desta aplicação.

Exemplo 3.6. Vamos projetar o vetor determinado pelos pontos $A(-1, 1)$ e $B(4, 3)$ sobre a reta $y = 1$. Neste caso, conforme Figura 26 temos que a projeção desejada é o vetor definido pelos pontos $A(-1, 1)$ e $C(4, 1)$, ou seja, $\overrightarrow{AC} = C - A = (4, 1) - (-1, 1) = (5, 0)$.

Figura 26 – Representação geométrica da projeção de vetor determinado pelos pontos $A(-1, 1)$ e $B(4, 3)$ sobre reta $y = 1$



Fonte – Autoria própria (2019)

Note que em todos os exemplos acima, podemos mostrar que o resultado obtido pela análise da representação geométrica é válida através da aplicação do Teorema de Pitágoras. Esta é uma observação importante a ser feita, pois mostra uma relação direta do quanto uma representação geométrica (quando possível) contribui no entendimento de algumas definições importantes na matemática, em especial a definição de projeção que é nosso objeto de estudo.

Exemplo 3.7. (Retomando o Exemplo 3.6) Vamos mostrar que o vetor $\overrightarrow{AC} = C - A = (4, 1) - (-1, 1) = (5, 0)$ definido pelos pontos $A(-1, 1)$ e $C(4, 1)$ é projeção ortogonal do vetor determinado pelos pontos $A(-1, 1)$ e $B(4, 3)$ sobre a reta $y = 1$.

Se o ponto $C(4, 1)$ é tal que \overrightarrow{AC} é ortogonal a \overrightarrow{BC} . Logo, aplicando o Teorema de Pitágoras ao triângulo ABC , temos

$$\begin{aligned} \|\overrightarrow{AB}\|^2 &= \|\overrightarrow{AC}\|^2 + \|\overrightarrow{BC}\|^2 \\ \Leftrightarrow d(A, B)^2 &= d(A, C)^2 + d(B, C)^2 \\ \Rightarrow (\sqrt{5^2 + 2^2})^2 &= (\sqrt{5^2 + 0^2})^2 + (\sqrt{0^2 + 2^2})^2 \\ \Rightarrow 29 &= 29, \end{aligned}$$

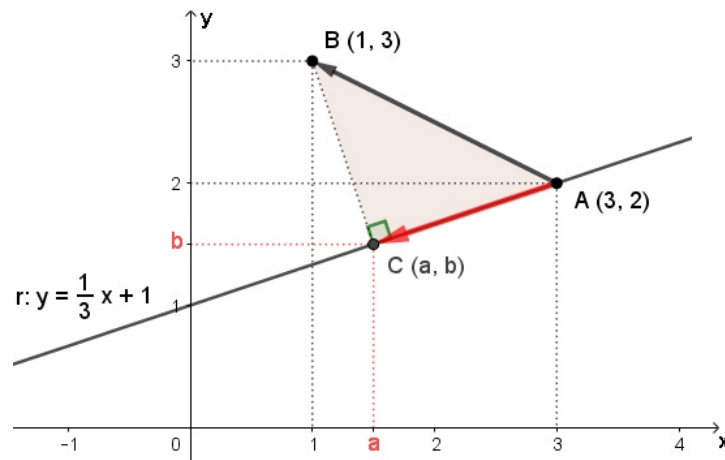
provando o resultado.

Vamos agora trabalhar com um exemplo onde se quer projetar um vetor v sobre uma reta oblíqua aos eixos coordenados.

Exemplo 3.8. Como projetar o vetor definido pelos pontos $A(3, 2)$ e $B(1, 3)$ sobre a reta $r : y = \frac{1}{3}x + 1$?

Geometricamente, temos a seguinte situação:

Figura 27 – Representação geométrica da projeção de vetor definido pelos pontos $A(3, 2)$ e $B(1, 3)$ sobre a reta $r : y = \frac{1}{3}x + 1$



Fonte – Autoria própria (2019)

Nosso objetivo é determinar (a, b) tal que \overrightarrow{BC} seja ortogonal ao vetor \overrightarrow{AC} . Assim, pelo Teorema de Pitágoras, temos

$$\begin{aligned} \|\overrightarrow{AB}\|^2 &= \|\overrightarrow{AC}\|^2 + \|\overrightarrow{BC}\|^2 \\ \Leftrightarrow d(A, B)^2 &= d(C, A)^2 + d(B, C)^2 \\ \Rightarrow 5 &= (3 - a)^2 + (2 - b)^2 + (a - 1)^2 + (b - 3)^2. \end{aligned} \quad (3.1)$$

Mas, (a, b) pertence a reta r , logo

$$b = \frac{1}{3}a + 1. \quad (3.2)$$

Substituindo a equação 3.2 em 3.1, obtemos

$$\begin{aligned}
 5 &= (3-a)^2 + (2-b)^2 + (a-1)^2 + (b-3)^2 \\
 &= (3-a)^2 + \left(2 - \frac{1}{3}a - 1\right)^2 + (a-1)^2 + \left(\frac{1}{3}a + 1 - 3\right)^2 \\
 &= (3-a)^2 + \left(1 - \frac{1}{3}a\right)^2 + (a-1)^2 + \left(\frac{1}{3}a - 2\right)^2 \\
 &= 9 - 6a + a^2 + 1 - \frac{2}{3}a + \frac{1}{9}a^2 + a^2 - 2a + 1 + \frac{1}{9}a^2 - \frac{4}{3}a + 4 \\
 &= \frac{20}{9}a^2 - 10a + 15.
 \end{aligned}$$

Logo, basta verificar as raízes da equação

$$\begin{aligned}
 \frac{20}{9}a^2 - 10a + 15 &= 0 \\
 \Leftrightarrow 20a^2 - 90a + 90 &= 0 \\
 \Leftrightarrow 2a^2 - 9a + 9 &= 0.
 \end{aligned}$$

Assim, pelo teorema de Pitágoras,

$$a = \frac{9 \pm \sqrt{(-9)^2 - 4 \cdot 2 \cdot 9}}{2 \cdot 2} = \frac{9 \pm \sqrt{9}}{4} = \frac{9 \pm 3}{4},$$

e, conseqüentemente segue que $a = 3$ ou $a = \frac{3}{2}$. Mas, $a = 3$ contraria nossa hipótese inicial de ortogonalidade. Então, tomando $a = \frac{3}{2}$, obtemos, $b = \frac{1}{3} \cdot \frac{3}{2} + 1 = \frac{3}{2}$, e portanto obtemos que $C\left(\frac{3}{2}, \frac{3}{2}\right)$.

Neste contexto, o vetor correspondente a projeção de \overrightarrow{AB} sobre a reta $r : y = \frac{1}{3}x + 1$ é o vetor $\overrightarrow{AC} = C - A = \left(\frac{3}{2}, \frac{3}{2}\right) - (3, 2) = \left(-\frac{3}{2}, -\frac{1}{2}\right)$.

3.2.3 AVALIAÇÃO

Acreditando que o processo avaliativo se dá como um todo e acontece processualmente durante as aulas, a avaliação da presente proposta de atividade se dará por participação e envolvimento dos alunos com questionamentos e contribuições e o envolvimento da turma na resolução dos problemas propostos durante toda a atividade, podendo o professor avaliar também todo o material produzido pela turma neste processo. Contudo, muito além dos registros feitos pelos estudantes, está a compreensão de cada etapa do processo que foi desenvolvido com a aplicação das etapas desta atividade. Essa compreensão, ainda que não apareça nos registros, é o que deve determinar a conclusão do processo avaliativo.

3.3 ATIVIDADE 2: PRÁTICA COM LUZ E SOMBRA

Público alvo: 3º ano do Ensino Médio

Tempo previsto: 3 aulas

Pré-requisitos: Semelhança de triângulos; Trigonometria no triângulo retângulo; Teorema de Pitágoras

Recursos: Caixa octante, régua, transferidor e material de escrita

3.3.1 OBJETIVOS

- Definir vetores em \mathbb{R}^2 e introduzir noção de vetores em \mathbb{R}^3 ;
- Utilizar propriedades relacionadas ao triângulo retângulo na resolução dos problemas propostos, com o intuito de compreender a definição de projeção ortogonal;
- Diferenciar projeções ortogonais e não ortogonais.

3.3.2 CAIXA OCTANTE

Com o objetivo de tornar prático a visualização dos conceitos preliminares relacionados a vetores, assim como a construção da definição de projeção com os estudantes, foi produzido um material concreto para ser utilizado, a qual a chamamos de "*caixa octante*". Tal material também será útil na aplicação da 3ª atividade que veremos mais adiante. Além disso, espera-se que este material forneça um estímulo maior aos estudantes, uma vez que os mesmos podem experimentar uma situação prática e interativa de aprendizagem.

Tal material representa um octante do \mathbb{R}^3 (para fins de cálculos, utilizaremos a caixa octante para representar especificamente o 1º octante). Uma vantagem desse material concreto é que sua construção leva materiais de fácil acesso e de baixo custo, viabilizando reprodução. Ademais, a montagem desse objeto leva em consideração aspectos práticos de transporte e manejo, facilitando o trabalho de locomoção do professor bem como de armazenagem de tal objeto. Veja (Figura 28).

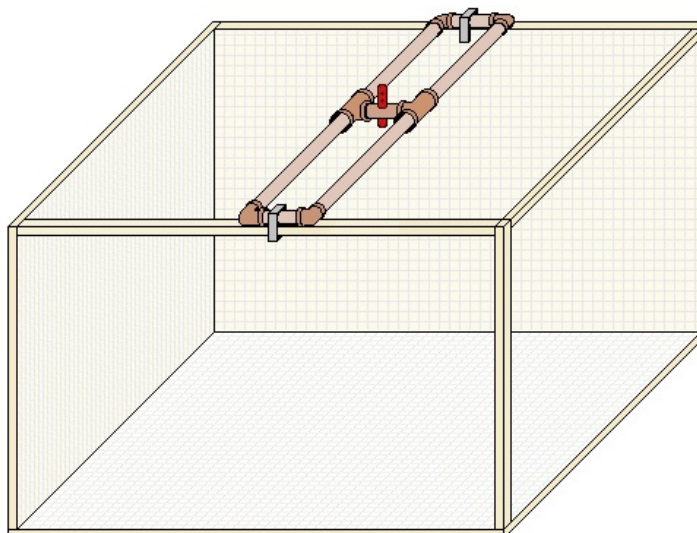
Uma característica muito importante da caixa octante é que ela pode ser adaptada (com inclusão de acessórios) para que qualquer atividade a ser realizada na mesma possa também ser aplicada a alunos com deficiência visual. Tal adaptação favorece a inclusão de qualquer aluno, independente de sua condição.

Materiais estruturais básicos utilizados:

- 3 painéis de metal: 2 peças de 50cm × 35cm × 2cm e 1 peça de 50cm × 52cm × 0,001cm;
- 1 placa de mdf (expositor com furação): 50cm × 52cm × 0,5cm.
- 1 metro de cano pvc com 4 "joelhos" e 2 "T" compatíveis;

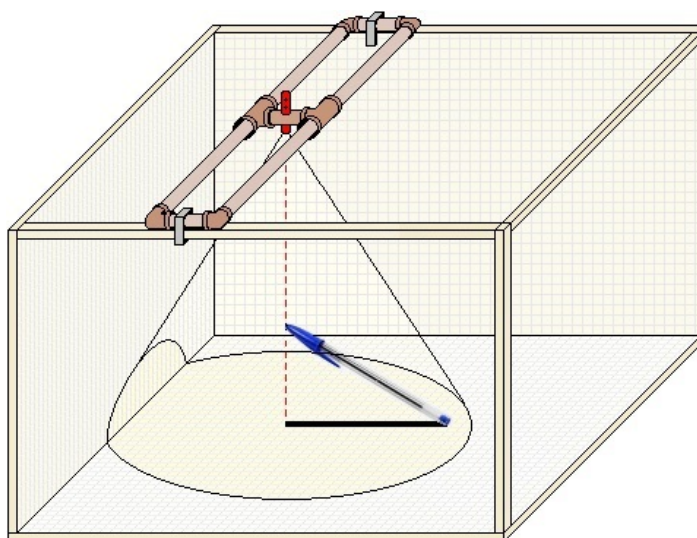
- 1,2 m de mdf de 2cm × 2cm;
- 25 parafusos com porca e arruela compatíveis e uma lanterna pequena com laser.

Figura 28 – Caixa octante



Fonte – Autoria própria (2019)

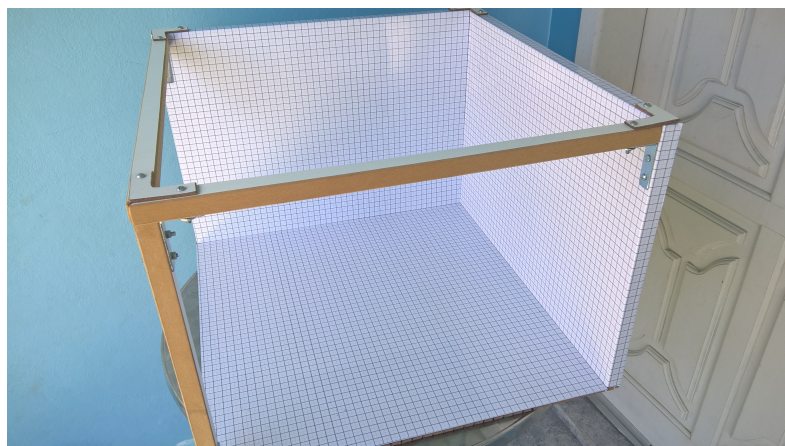
Figura 29 – Exemplo de atividade a ser realizada na caixa octante



Fonte – Autoria própria (2019)

Conforme podemos visualizar na (Figura 29), além da caixa octante possibilitar ao professor trabalhar características de pontos e principalmente de vetores no espaço \mathbb{R}^3 , ela permitirá o trabalho de calcular o comprimento da sombra de uma luz projetada ortogonalmente sobre um objeto (que é o objetivo desta 2ª atividade). A seguir algumas fotos da caixa octante montada:

Figura 30 – Caixa octante montada I



Fonte – Autoria própria (2019)

Figura 31 – Caixa octante montada II



Fonte – Autoria própria (2019)

3.3.3 DESENVOLVIMENTO DA PROPOSTA

3.3.3.1 LINGUAGEM VETORIAL

Assim como na primeira atividade, faz-se necessário que o professor inicie o trabalho inserindo uma linguagem vetorial básica, através da definição de vetores. No entanto, caso o professor já tenha desenvolvido a 1ª atividade (levando em consideração que as atividades são independentes, e que o professor tem a liberdade para aplicar as atividades na ordem que lhe for conveniente), basta que o mesmo complemente tal etapa inserindo uma noção introdutória do \mathbb{R}^3 , que abordaremos em seguida.

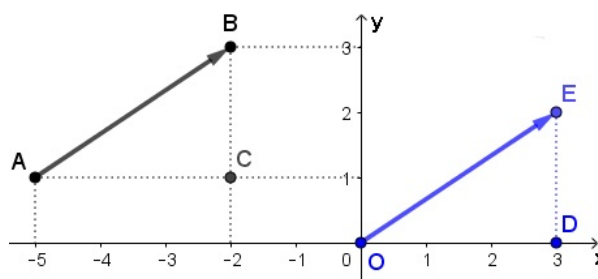
Sobre a construção da definição de vetores em \mathbb{R}^2 , o desenvolvimento é análogo ao exposto na primeira atividade. Inicialmente, é sugerido que o professor trabalhe com a ideia de **segmento orientado** para vetores. Tal abordagem é mais simples de ser absorvida pelo aluno, uma vez que os mesmos têm maior familiaridade com a definição de segmento. Neste contexto,

a noção de vetores vai sendo construída à medida que a atividade vai se desenvolvendo. Ao final da atividade o professor pode concretizar a definição de vetores como o conjunto segmentos orientados equipolentes.

Exemplo 3.9. Dados os pontos $A = (-5, 1)$ e $B = (-2, 3)$, vamos definir o segmento orientado que tem A como ponto de partida e B como ponto de chegada. Segue que

$$\vec{AB} = B - A = (-2, 3) - (-5, 1) = (-2 - (-5), 3 - 1) = (3, 2).$$

Figura 32 – Representação geométrica do vetor definido pelos pontos A e B

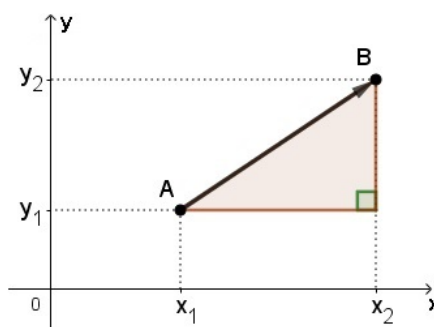


Fonte – Autoria própria (2019)

Com esta representação geométrica o professor pode mostrar ao aluno que $\vec{AB} = \vec{OE}$, através da congruência dos triângulos ABC com OED e justificar a vantagem do trabalho com vetores. Em seguida, sugere-se que o professor repita a atividade exposta no exemplo anterior, variando os pontos, de modo que os estudantes consigam definir um vetor tendo dois pontos quaisquer.

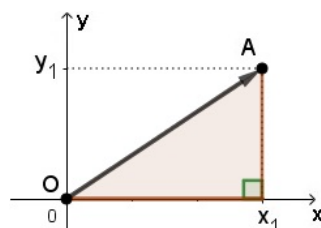
Como estamos trabalhando basicamente com triângulo retângulo e levando em consideração a abordagem de segmento orientado para vetores, o professor tem condições de mostrar que o comprimento de um vetor (**norma**) é dado basicamente pela aplicação do teorema de Pitágoras, o que se resume ao cálculo da distância entre dois pontos (os pontos que definem o vetor em estudo).

Figura 33 – $\|\vec{AB}\| = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$



Fonte – Autoria própria (2019)

Figura 34 – Vetor com extremidade na origem: $\|v\| = d(0, A) = \sqrt{x_1^2 + y_1^2}$



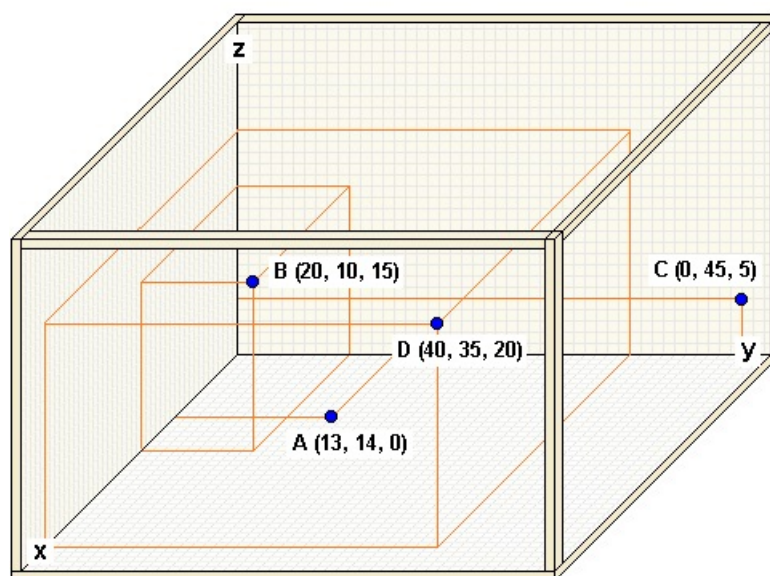
Fonte – Autoria própria (2019)

Para introduzir a definição de vetores no espaço em \mathbb{R}^3 , inicialmente é necessário que o professor contextualize ao estudante o que vem a ser o \mathbb{R}^3 , inserindo principalmente sua composição e noção de coordenadas. Para tanto, o professor pode fazer uso da caixa octante, tanto na exposição das principais características e definição quanto no entendimento da composição das coordenadas no espaço. Isso permitirá ao estudante uma melhor visualização e contextualização do conteúdo.

Exemplo 3.10. Segue na Figura 35 a representação na caixa octante dos pontos:

$$A(13, 14, 0), \quad B(20, 10, 15), \quad C(0, 45, 5) \quad \text{e} \quad D(40, 35, 20)$$

Figura 35 – Representação dos pontos A, B, C e D



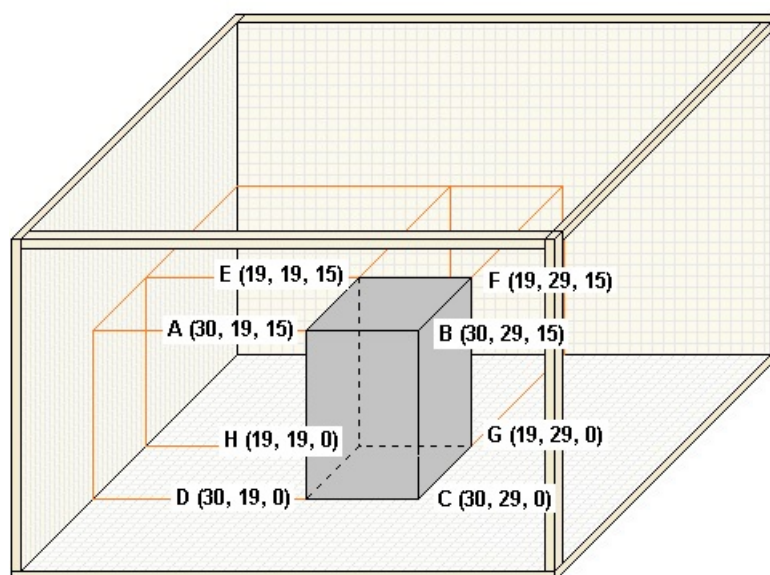
Fonte – Autoria própria (2019)

Outra forma interessante de abordar as coordenadas no \mathbb{R}^3 é sob o ponto de vista de geometria espacial. O professor pode inserir algum sólido geométrico (caixa de leite, caixa de cereal entre outros) na caixa octante e instigar seus alunos a expor as coordenadas de seus vértices. Mesmo que o professor não tenha trabalhado com a turma os sólidos geométricos a

atividade ainda é viável, levando em consideração que o objetivo é mostrar a localização do sólido escolhido com base nas coordenadas de seus vértices.

Exemplo 3.11. Veja na Figura 36 as coordenadas de um paralelepípedo posicionado de forma arbitrária na caixa octante.

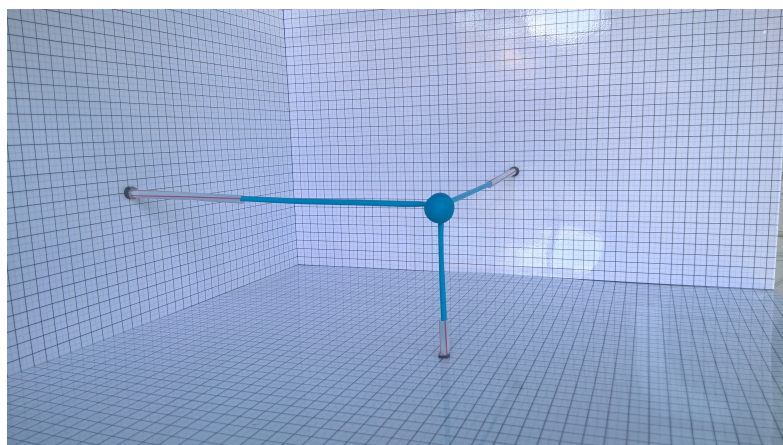
Figura 36 – Representação das coordenadas do paralelepípedo ABCDEFGH



Fonte – Autoria própria (2019)

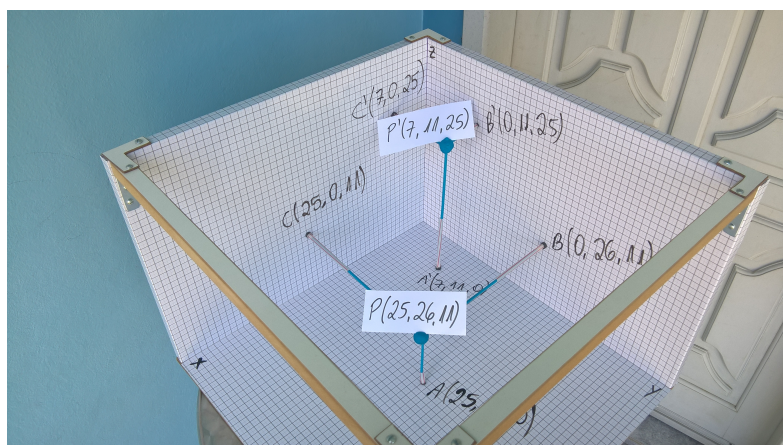
A seguir algumas fotos de exemplos práticos para ilustrar a marcação de pontos com a utilização da caixa octante. Para a marcação de pontos foram utilizados alguns acessórios, dentre eles, massinha de modelar, palito de churrasco e ima.

Figura 37 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante I



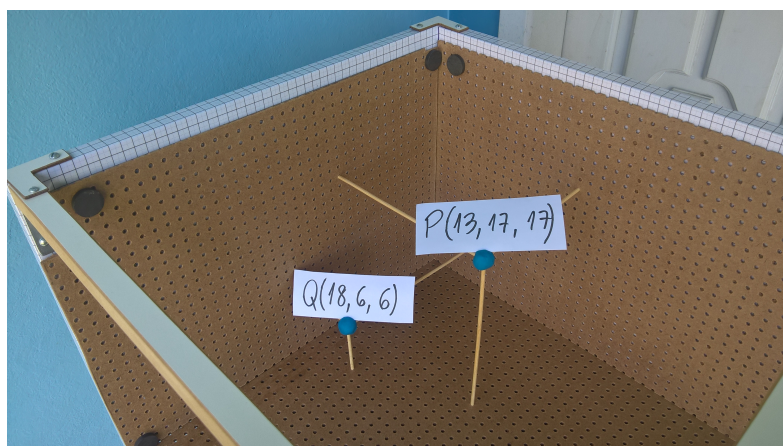
Fonte – Autoria própria (2019)

Figura 38 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante II



Fonte – Autoria própria (2019)

Figura 39 – Coordenadas de pontos em \mathbb{R}^3 na caixa octante com adaptação para alunos com deficiência visual



Fonte – Autoria própria (2019)

Sobre a construção da definição de vetores em \mathbb{R}^3 , o desenvolvimento é análogo ao exposto em \mathbb{R}^2 . Continuaremos com a ideia de **segmento orientado** para vetores, de modo que no decorrer da atividade a noção de vetores vai sendo construída.

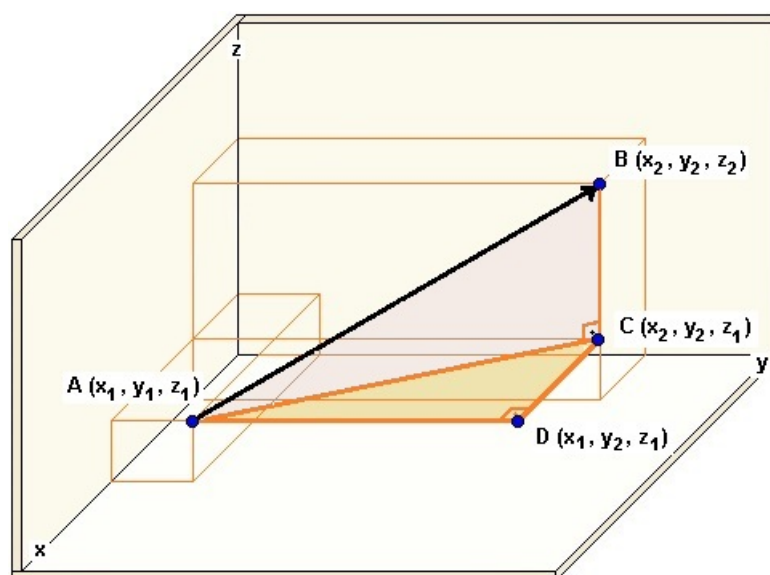
Exemplo 3.12. Dados os pontos $A = (26, 4, 5)$ e $B = (5, 32, 20)$, vamos definir o segmento orientado que tem A como ponto de partida e B como ponto de chegada. Segue que

$$\begin{aligned} \overrightarrow{AB} &= B - A \\ &= (5, 32, 20) - (26, 4, 5) \\ &= (5 - 26, 32 - 4, 20 - 5) \\ &= (-21, 28, 15). \end{aligned}$$

Em seguida, sugere-se que o professor repita a atividade exposta no exemplo anterior, variando os pontos, de modo que os estudantes consigam definir um vetor tendo dois pontos quaisquer.

Perceba que como estamos trabalhando basicamente com triângulo retângulo também em \mathbb{R}^3 e levando em consideração a abordagem de segmento orientado para vetores, o professor também tem condições de mostrar que o comprimento de um vetor (**norma**) é dado basicamente pela aplicação do teorema de Pitágoras (duas vezes), o que se resume ao cálculo da distância entre dois (os pontos que definem o vetor em estudo).

Figura 40 – $\|\vec{AB}\| = d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$



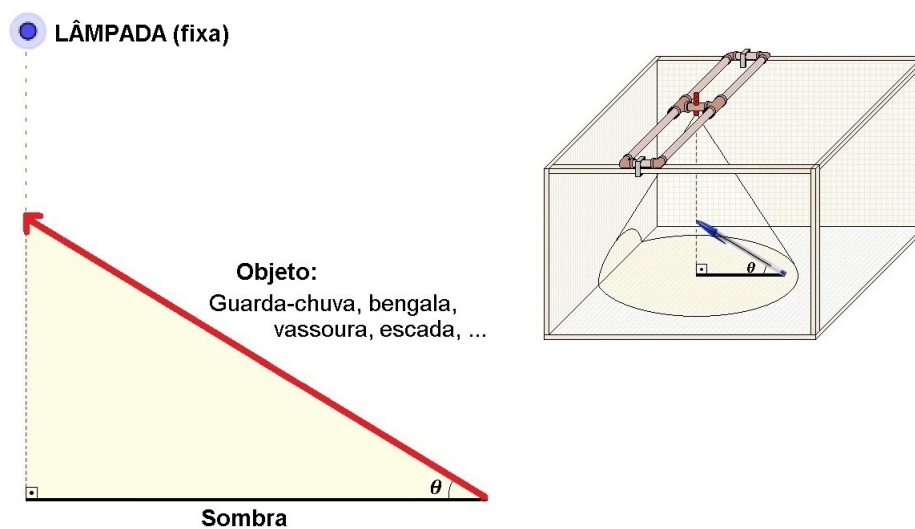
Fonte – Autoria própria (2019)

3.3.3.2 ESCOLHA DE UM OBJETO PARA EXPERIMENTO

Como nosso objetivo é obter o comprimento da sombra de um objeto projetada ortogonalmente por uma lâmpada puntiforme presa ao teto, precisamos inicialmente definir um objeto de trabalho. Para isso, pedir para turma, em geral, selecionar um objeto, como por exemplo lápis, caneta, régua, entre outros. A ideia é que sejam selecionados objetos que caibam na caixa octante, tendo em vista que é de interesse utilizar esse material concreto.

Neste instante cabe ao professor mostrar à turma que o experimento pode ser ampliado para um objeto de qualquer tamanho simplesmente utilizando semelhança de triângulo envolvendo algum objeto "pequeno" com o objeto escolhido pelos alunos (Figura 41).

Figura 41 – Exemplo de atividade realizada na caixa octante com semelhança de triângulos

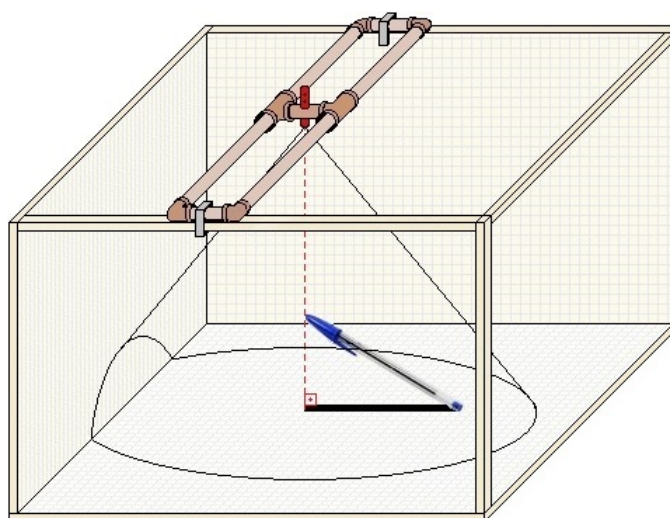


Fonte – Autoria própria (2019)

3.3.3.3 EXPERIMENTO COM OBJETO

De posse do objeto, pedir para algum estudante primeiramente posicionar o objeto na caixa octante de modo que o mesmo possua alguma inclinação em relação a base da caixa. Em seguida, ligar a lanterna (com laser) de modo que a mesma esteja ortogonal à base da caixa, em relação à extremidade do objeto escolhido, conforme Figura 42. Para fins de cálculo, nesta 2ª atividade todos os objetos dos exemplos foram posicionados na caixa paralelos a um dos planos laterais.

Figura 42 – Sombra de um objeto projetada ortogonalmente por uma lâmpada puntiforme fixa

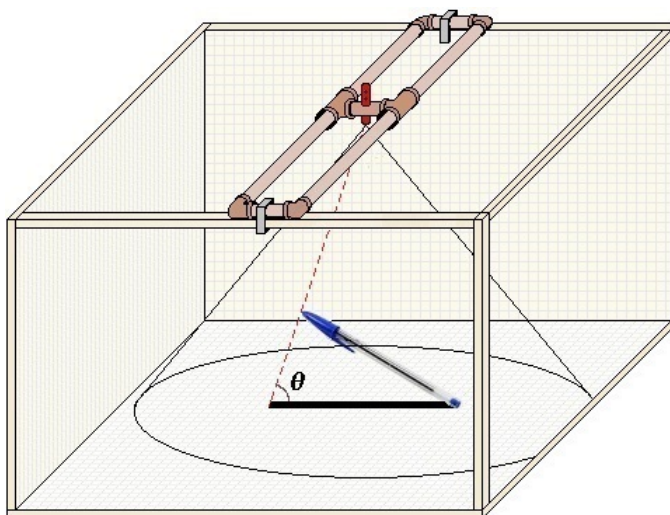


Fonte – Autoria própria (2019)

Como estamos interessados em fazer um estudo de projeção ortogonal, vamos modelar o experimento de calcular o comprimento da sombra como um triângulo retângulo. Neste contexto,

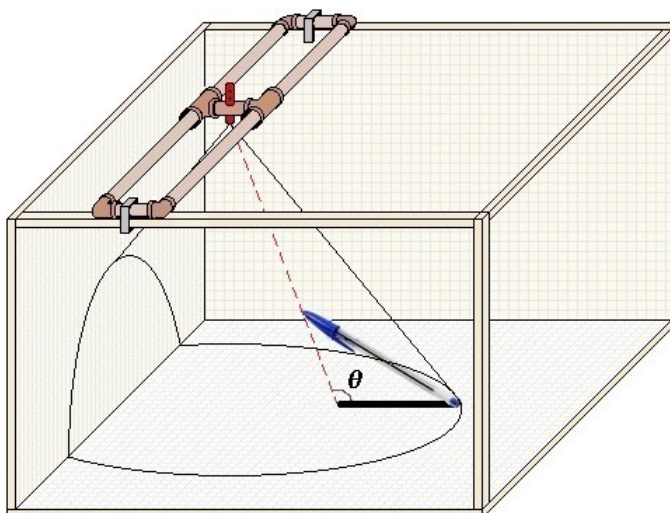
precisamos garantir a existência de um ângulo reto, que é dado pelo correto posicionamento da lanterna em relação ao objeto (Figuras 43 e 42). Para isso, basta o professor mostrar ao aluno que a medida do comprimento da sombra do objeto varia conforme posicionamento da luz e que para garantir a ortogonalidade precisamos utilizar algum instrumento de medição de ângulo, como por exemplo o transferidor.

Figura 43 – Projeção não ortogonal com $\theta < 90^\circ$



Fonte – Autoria própria (2019)

Figura 44 – Projeção não ortogonal com $\theta > 90^\circ$



Fonte – Autoria própria (2019)

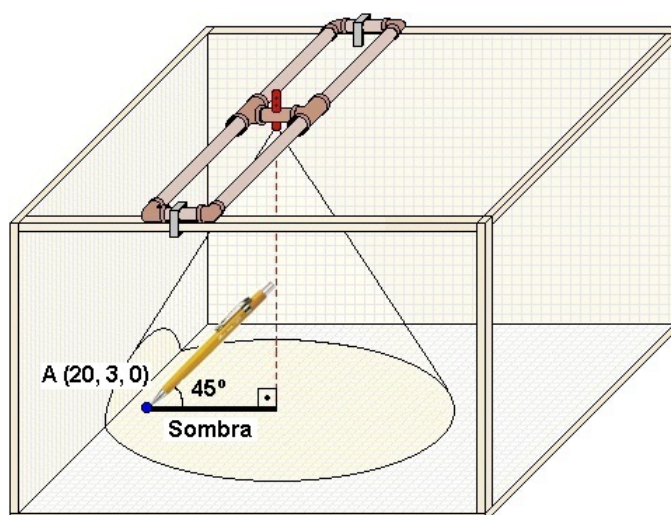
Em seguida, propor aos estudantes que pensem em formas para se calcular a medida do comprimento da sombra projetada ortogonalmente pela luz fixa em relação a base da caixa pelo objeto posicionado. Apesar de parecer natural a ideia de medir com uma régua o comprimento solicitado, a intenção desta proposta é que os alunos utilizem algum argumento matemático para

obter a medida solicitada, ou mesmo provar o resultado obtido pela simples medição com uma régua.

Contudo, valem alguns considerações sobre as hipóteses deste nosso problema. Primeiramente, a partir do momento que um objeto foi escolhido pelos alunos e conseqüentemente posicionado de forma arbitrária na caixa octante, partimos da ideia de que "conhecemos" este objeto e sua localização, ou seja, é de nosso conhecimento o comprimento deste objeto, o ângulo de inclinação deste objeto com relação a base da caixa (medido com transferidor), assim como sua localização em termos das coordenadas do ponto onde tal objeto foi fixado.

Exemplo 3.13. Vamos obter o comprimento da sombra de uma lapiseira de 14 cm posicionada no ponto $A(20, 3, 0)$ e projetada ortogonalmente com relação à base da caixa octante com inclinação de 45° (Figura 45).

Figura 45 – Obtendo comprimento da sombra de uma lapiseira posicionada no ponto $A(20, 3, 0)$ com inclinação de 45° com relação à base da caixa



Fonte – Autoria própria (2019)

Do triângulo retângulo modelado pelo problema, temos que:

$$\cos 45^\circ = \frac{\text{Comprimento da sombra}}{\text{Comprimento da lapiseira}} = \frac{\text{Comprimento da sombra}}{14}$$

e, portanto, o comprimento da sombra é de aproximadamente 9,9 cm.

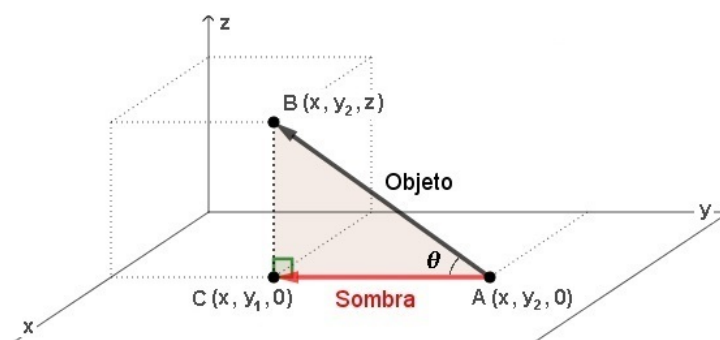
Neste momento, sugere-se que o professor repita a atividade exposta no exemplo anterior, variando o objeto, sua localização e inclinação, de modo que os estudantes consigam ao final verificar a relação entre a definição de projeção ortogonal com o simples exercício de calcular a sombra de objetos.

Em seguida, a partir da inserção de uma linguagem vetorial apropriada, vamos obter o vetor que corresponde à sombra de tal objeto em função das coordenadas do ponto onde o objeto

escolhido foi posicionado, além do ângulo de inclinação e comprimento do objeto escolhido. Contudo, consideraremos que o objeto foi posicionado paralelamente a uma das laterais da caixa octante, no caso, paralelo ao plano $x = 0$. Nos demais casos, o desenvolvimento é análogo ao que será exposto a seguir. Tal exercício será útil na realização da 3ª atividade que será discutida na próxima seção.

Exemplo 3.14. Seja \overrightarrow{AB} correspondente ao objeto escolhido, onde $A(x, y_1, 0)$ e $B(x, y_2, z)$ com $x, y_1, y_2, z \in \mathbb{R}$ (sendo A o ponto de contato deste objeto com a caixa octante). Considere ainda que $\|\overrightarrow{AB}\| = k$ e que o ângulo de inclinação deste objeto com relação a base da caixa é θ (Figura 46).

Figura 46 – Sombra de um objeto qualquer posicionado em ponto arbitrário e com inclinação de θ com relação à base da caixa



Fonte – Autoria própria (2019)

Queremos obter $C(x, y_2, 0)$ em função das coordenadas de A , do comprimento do objeto e do ângulo de inclinação de tal objeto, tal que o \overrightarrow{AC} seja ortogonal a \overrightarrow{BC} . Desta forma, aplicando definição de cosseno ao triângulo retângulo ABC , temos que

$$\cos \theta = \frac{\|\overrightarrow{AC}\|}{\|\overrightarrow{AB}\|} = \frac{\sqrt{(x-x)^2 + (y_2 - y_1)^2 + (0-0)^2}}{k} = \frac{y_2 - y_1}{k}$$

e, portanto

$$y_1 = y_2 - k \cos \theta. \quad (3.3)$$

Exemplo 3.15. Retomando o exemplo 3.13, vamos obter o vetor \overrightarrow{AC} correspondente à sombra de uma lapiseira de 14 cm posicionada no ponto $A(20, 3, 0)$ e projetada ortogonalmente com relação à base da caixa octante com inclinação de 45° (Figura 45).

Desta forma, aplicando a equação 3.3, temos que $y = 3 + 14 \cdot \cos 45^\circ \approx 3 + 9,9 = 12,9$. Consequentemente, temos que $C(20; 12,9; 0)$, e portanto, $\overrightarrow{AC} = C - A = (20; 12,9; 0) - (20, 3, 0) = (0; 9,9; 0)$.

Em seguida, como forma de fixação e formalização da definição de projeção ortogonal, proporcionado pelo exemplo anterior, é sugerido que o professor repita a atividade exposta acima, variando o objeto, sua localização e inclinação.

3.3.4 AVALIAÇÃO

Acreditando que o processo avaliativo se dá como um todo e acontece processualmente durante as aulas, a avaliação da presente proposta de atividade se dará por participação e envolvimento dos alunos com questionamentos e contribuições e o envolvimento da turma na resolução dos problemas propostos durante toda a atividade, podendo o professor avaliar também todo o material produzido pela turma neste processo. Contudo, muito além dos registros feitos pelos estudantes, está a compreensão de cada etapa do processo que foi desenvolvido com a aplicação das etapas desta atividade. Essa compreensão, ainda que não apareça nos registros, é o que deve determinar a conclusão do processo avaliativo.

3.4 ATIVIDADE 3: VERIFICANDO IMPEDIMENTOS NO FUTEBOL

Público alvo: 3º ano do Ensino Médio

Tempo previsto: 3 aulas

Pré-requisitos: 2ª atividade

Recursos: Caixa octante, régua, transferidor e material de escrita

3.4.1 OBJETIVOS

- Fixar definições de vetores em \mathbb{R}^2 e \mathbb{R}^3 ;
- Verificar aplicação direta da definição de projeção ortogonal no futebol, mais precisamente no que diz respeito a avaliação de uma das regras de jogo: o impedimento.

A ideia geral desta atividade tem como motivação levar os estudantes a um pleno entendimento da regra do impedimento por meio do estudo de projeções minimizando a proposição de críticas e reclamações por vezes decorrentes do desentendimento dessa regra. Além disso, esta atividade lúdica também objetiva instigar e motivar o aluno no estudo da disciplina de Matemática, ao mesmo tempo em que busca aproximar teoria e prática de tópicos relacionados com geometria analítica através de uma aplicação bem popular no esporte.

3.4.2 DESENVOLVIMENTO DA PROPOSTA

3.4.2.1 REGRA DE IMPEDIMENTO

Conforme as regras da IFAB (*The international football association board*), que é o órgão responsável pela formulação e manutenção das regras que compõem o futebol, um jogador estará em posição de impedimento quando qualquer parte de sua cabeça, corpo ou pés (exceto mãos e braços, inclusive dos goleiros) estiver no campo adversário e mais próximo da linha de meta adversária do que a bola e o penúltimo adversário. Ademais, um jogador em posição de impedimento somente será punido se no momento em que a bola for jogada ou tocada por um

companheiro de equipe (considerando o momento do primeiro ponto de contato com a bola) participar ativamente do jogo, seja interferindo no jogo, interferindo num adversário ou mesmo ganhando vantagem de sua posição de impedimento (IFAB, 2019).

O responsável por esta observação simultânea (observação do jogador e da bola no momento do passe, para quaisquer posições da bola e do jogador no campo de futebol) é o árbitro (juiz) assistente, também conhecido como “bandeirinha”. Este árbitro se locomove em uma das laterais do campo de futebol procurando sempre a melhor posição para marcar, entre outras infrações do futebol, o impedimento (DELFINI; JESUS, 2011).

Vejamos alguns exemplos.

Figura 47 – Exemplo de impedimento I



Fonte – Adaptado de (IFAB, 2019)

Figura 48 – Exemplo de impedimento II



Fonte – Adaptado de (IFAB, 2019)

Figura 49 – Exemplo de impedimento III



Fonte – Adaptado de (IFAB, 2019)

Figura 50 – Exemplo de impedimento IV



Fonte – Adaptado de (IFAB, 2019)

Apesar da popularidade do futebol, não são todos os alunos que conhecem a regra do impedimento numa partida de futebol. Sendo assim, num primeiro momento o professor conduzirá um trabalho (simples) a fim de que todos os alunos entendam a regra. Neste momento, caso haja necessidade de maior compreensão e disponibilidade de computador, sugere-se que o professor exiba a sua turma alguns vídeos de situações de impedimentos.

3.4.2.2 FUNCIONAMENTO DA TECNOLOGIA VAR

Por muito tempo a forma como esta regra foi "avaliada" durante uma partida de futebol foi tema de longos debates por especialistas devido a erros cometidos pela equipe de arbitragem, seja na marcação equivocada de um impedimento, ou mesmo na não marcação de lances

claros de impedimentos, os quais interferiam diretamente nos resultados das partidas. Uma das principais causas destes erros de marcação está relacionado com o fato de o árbitro assistente ser humano, sujeito a erros. Nem todos os lances de impedimentos são claros (visualmente), o que consequentemente dificulta o trabalho dos árbitros.

Muitas críticas a marcações de impedimentos aos árbitros assistentes, por parte das pessoas que assistem ao jogo (seja num estádio de futebol ou mesmo numa televisão), têm como referência imagens diferentes das que um árbitro assistente (bem posicionado) tem num jogo, e acabam tirando conclusões precipitadas sem levar em conta seu ponto de visão (Figuras 51 e 52).

Figura 51 – Posicionamento de câmera para verificação de impedimento I



Fonte – Adaptado de (IFAB, 2019)

Figura 52 – Posicionamento de câmera para verificação de impedimento II



Fonte – Adaptado de (IFAB, 2019)

Para tanto, recentemente foi aprovada pela FIFA (do francês: *Fédération Internationale de Football Association*) a tecnologia VAR, (*video assistant referee*) ou *árbitro assistente de vídeo*, com o objetivo de extinguir estes erros de arbitragem. Aqui não estamos fazendo nenhuma crítica à forma como era avaliada situações de impedimento, mas mostrar as vantagens advindas com esta tecnologia no que diz respeito à marcação de impedimentos, relacionando com o estudo de projeções.

Um árbitro assistente de vídeo (VAR) é um árbitro de partida que possui acesso independente às imagens gravadas da partida, o qual poderá auxiliar o árbitro na eventualidade de um “erro claro e óbvio”, como por exemplo a marcação de impedimento (IFAB, 2019).

Para análise de um lance de impedimento, levando em conta que a regra do impedimento (mencionada na etapa anterior) leva em consideração a menor distância entre o atacante e o defensor rival mais próximos em relação à linha de meta, o sistema primeiramente traça a projeção ortogonal da parte do corpo mais próxima da linha de meta destes dois jogadores em questão (exceto mãos e braços, inclusive dos goleiros) em relação ao campo. Em seguida é definido uma reta paralela à linha de gol para cada projeção, isto é, para cada jogador, para por fim ser analisada a menor distância entre as retas paralelas e a linha de gol. A visualização destas retas paralelas para cada jogador torna este sistema prático para a arbitragem e ao mesmo tempo didático. Vejamos um exemplo de participação do VAR (Figura 53):

Figura 53 – Lance capturado por emissora de TV



Fonte – (O DOCUMENTO, 2019)

Tal aplicação ocorreu numa partida de futebol no dia 12/06/2019 válida pela série A do campeonato brasileiro de futebol, entre Internacional e Bahia. Na ocasião houve muitas reclamações por conta da não marcação de um suposto impedimento no 1º gol do internacional (Figuras 54 e 55).

Figura 54 – Reação de torcedores



Fonte – (TORCEDORES, 2019)

Figura 55 – Reação de dirigentes de um dos clubes envolvidos



Fonte – (GLOBOESPORTE, 2019)

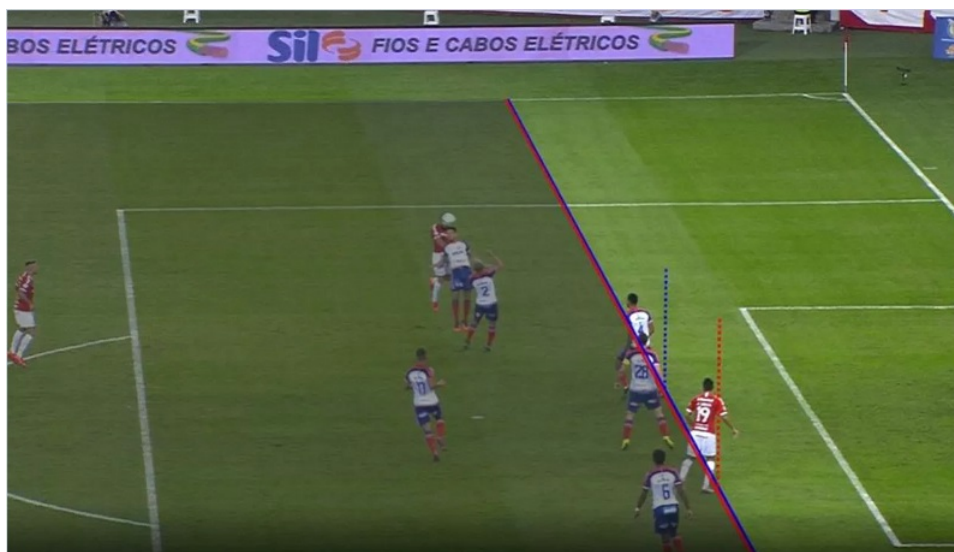
Com o objetivo de minimizar as críticas, o chefe de arbitragem da CBF (Confederação Brasileira de Futebol) divulgou as imagens analisadas pelo VAR e explicou o funcionamento do VAR, justificando a correta decisão tomada pelos árbitros (Figuras 56, 57 e 58).

Figura 56 – Declaração do chefe de arbitragem



Fonte – (SPORTV, 2019)

Figura 57 – Imagem capturada e analisada pelo VAR com projeções



Fonte – (SPORTV, 2019)

Figura 58 – Funcionamento do VAR

REDAÇÃO SPORTV

– Nós tomamos a decisão de fazer a divulgação dessas imagens pela credibilidade do projeto. Não temos autorização para mostrar essas imagens em todos os momentos, mas acho que nesse momento é importante mostrar que o gol foi legal e por que a equipe do árbitro de vídeo chegou a essa conclusão. O que o pessoal vê em casa é a imagem em diagonal. A imagem da transmissão não estava exatamente na linha de impedimento e, quando a gente olha a imagem em diagonal, a gente tem uma distorção visual. O que a gente utiliza na cabine do VAR? É feita uma calibragem de todas câmeras, é parada a imagem no momento do toque do jogador de cabeça e, a partir daí, a gente trabalha com a linha vertical. A gente puxa essa linha em relação à linha de fundo, vê qual o ponto do jogador, nesse caso o ponto dele mais próximo da linha de fundo é o ombro, faz uma projeção ao solo e vai até o jogador do Bahia, pois com a imagem em diagonal a gente não consegue ver o pé do jogador, que está dando condição – explicou Gaciba.

Fonte – (SPORTV, 2019)

É muito importante que os estudantes realmente entendam a regra e o funcionamento da VAR, principalmente no que diz respeito à "distância" do jogador com a linha do gol, que está relacionado à distância da projeção da parte do corpo do jogador mais próximo da linha de gol com relação ao campo até a linha do gol.

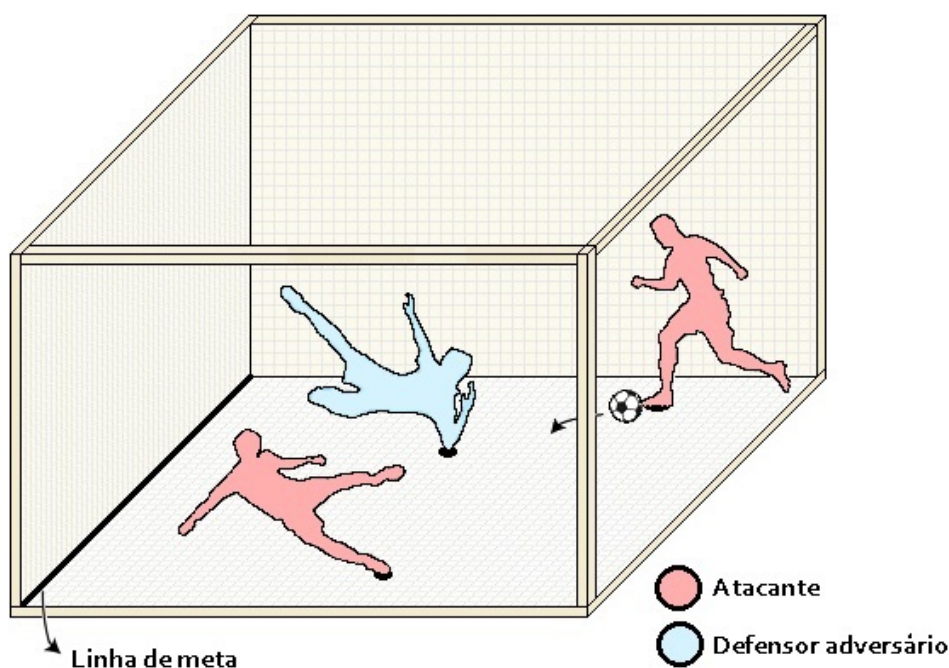
3.4.2.3 EXPERIMENTO COM VAR

A ideia desta atividade é fazer os alunos experimentarem na prática o funcionamento do VAR por meio de um experimento controlado, e mostrar a relação direta com o estudo de projeção que propomos. Resumidamente, a ideia é simular situações de impedimento na caixa octante com base em análise de projeções.

Primeiramente utilizando a caixa octante, pedir para os estudantes selecionarem dois (ou mais) objetos, que representarão os jogadores, e posicionarem lado a lado na base da caixa voltados para um dos planos que compõe os lados da caixa octante. Defina a reta de interseção da base da caixa com o plano escolhido como lado como sendo a linha de meta.

Para que a atividade seja mais atrativa para os estudantes, o professor pode sugerir a utilização de bonecos em miniaturas de jogadores de diversos formatos, ficando a critério dos estudantes a escolha. A utilização de bonecos em miniatura também é interessante para mostrar aos alunos as vantagens do VAR em relação à avaliação de um lance apenas pelo árbitro assistente, principalmente no que se refere às partes do corpo analisadas nestes lances. Para um maior realismo do experimento é interessante que também seja feita uma escolha entre qual objeto será o atacante e qual o defensor adversário (Figura 59).

Figura 59 – Experimento simulado do VAR

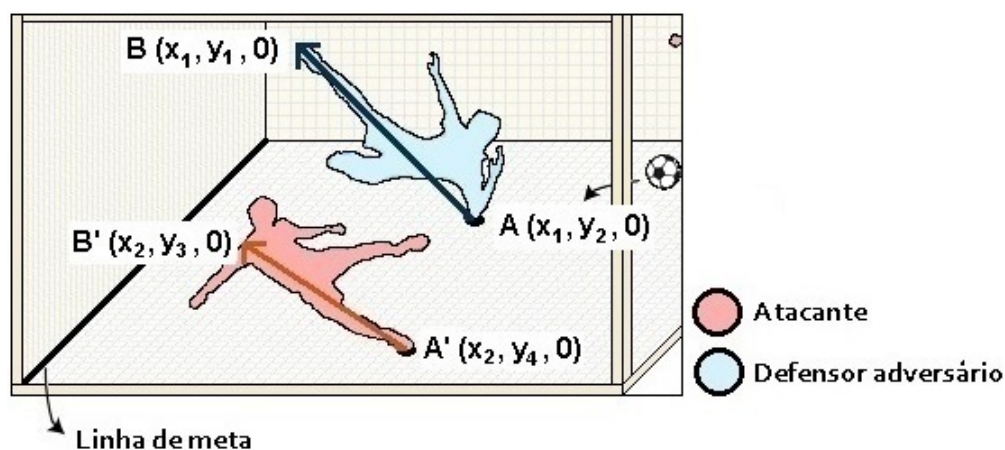


Fonte – Autoria própria (2019)

Após os objetos serem posicionados pelos estudantes, inicia-se o processo de análise de existência ou não de impedimento. O processo que buscaremos desenvolver é análogo ao exposto no Exemplo 3.14. Vale mencionar que como estamos buscando ver o jogador mais próximo da linha do gol (meta) por meio da análise das suas respectivas projeções, vamos trabalhar com vetores ortogonais à linha de gol.

Inicialmente vamos definir um vetor correspondente para cada jogador envolvido com norma fixada inicialmente. Estes vetores deverão ser definidos de modo que seus pontos iniciais devem pertencer ao plano da base e seus pontos finais deverão ser respectivamente o ponto do corpo de cada jogador mais próximo da linha de gol (lembrando que devem ser excluídos mãos e braços) (Figura 60).

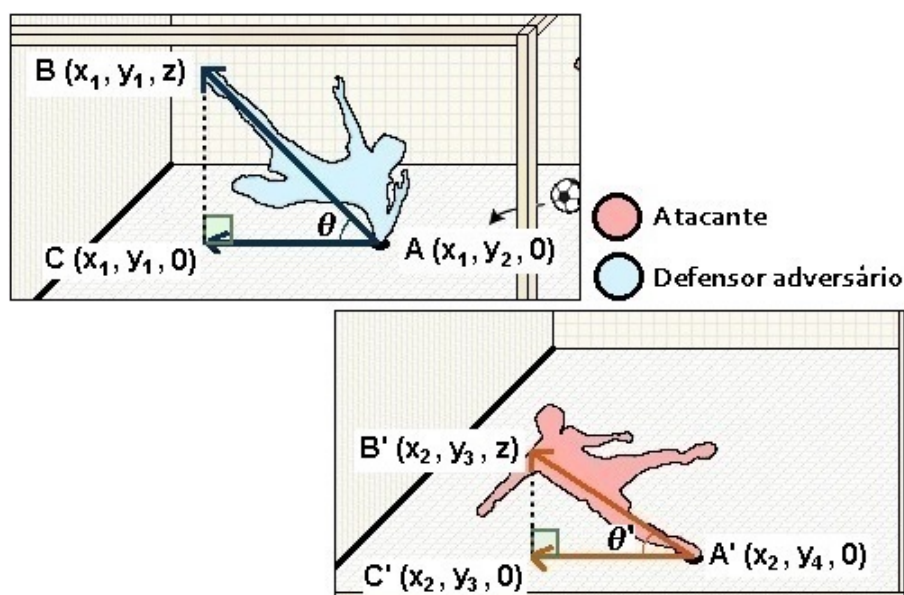
Figura 60 – Definição de vetores correspondentes a cada jogador envolvido na jogada



Fonte – Autoria própria (2019)

Como a definição da regra de impedimento está vinculada à definição de projeção ortogonal do ponto do corpo de cada jogador mais próximo da linha de gol, vamos verificar a projeção ortogonal desses vetores correspondente aos jogadores envolvidos. Para isso, definamos um triângulo retângulo para cada vetor (Figura 61).

Figura 61 – Definição de triângulo retângulo correspondente a cada jogador com relação à base da caixa (campo)



Fonte – Autoria própria (2019)

Analisando a ilustração acima, perceba que pela forma como definimos a linha de gol (eixo x) e vetores correspondentes aos jogadores (ortogonais à linha de gol), basta que verifiquemos a componente y dos pontos C e C' , que definem respectivamente os vetores \vec{AC} e $\vec{A'C'}$ que são as projeções ortogonais dos vetores de cada jogador envolvido.

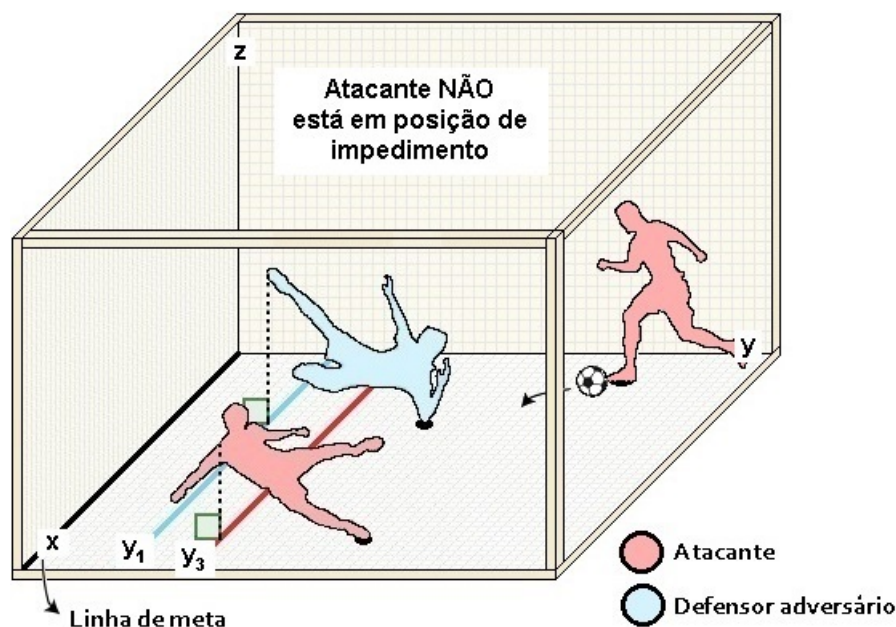
Além disso, como inicialmente fixamos a norma de cada vetor bem como os posicionamos na caixa octante em nosso experimento, temos que é de nosso conhecimento as coordenadas dos pontos onde tais vetores foram fixados na caixa, o ângulo de inclinação desses vetores com relação a base da caixa (verificado com auxílio de transferidor) além da norma de cada vetor.

Seguindo desenvolvimento análogo ao realizado no Exemplo 3.14, temos que nosso objetivo é obter $C(x_1, y_1, 0)$ em função das coordenadas de A , da norma do vetor definido e do ângulo de inclinação de tal vetor, tal que o \vec{AC} seja ortogonal a \vec{BC} e $C'(x_2, y_3, 0)$ em função das coordenadas de A' , da norma do vetor definido e do ângulo de inclinação de tal vetor, tal que o $\vec{A'C'}$ seja ortogonal a $\vec{B'C'}$. Desta forma, aplicando definição de cosseno aos triângulos retângulos ABC e $A'B'C'$, temos que as componentes y dos pontos C e C' são respectivamente

$$y_1 = y_2 - \|\vec{AB}\| \cos \theta \quad e \quad y_3 = y_4 - \|\vec{A'B'}\| \cos \theta'. \quad (3.4)$$

Assim, a verificação da existência de impedimento pode ser feita com base na análise das retas da base paralelas a linha do gol passando pelos pontos y_1 e y_3 (Figura 62).

Figura 62 – Verificação da não existência de impedimento



Fonte – Autoria própria (2019)

Tendo o entendimento do processo, é sugerido ao professor que estimule sua turma a variar o posicionamento dos bonecos de jogadores com intuito de que os alunos consigam comprovar analiticamente a existência ou não de impedimentos.

A seguir uma foto relacionada a um exemplo de definição do vetor correspondente a um jogador em estudo na caixa octante.

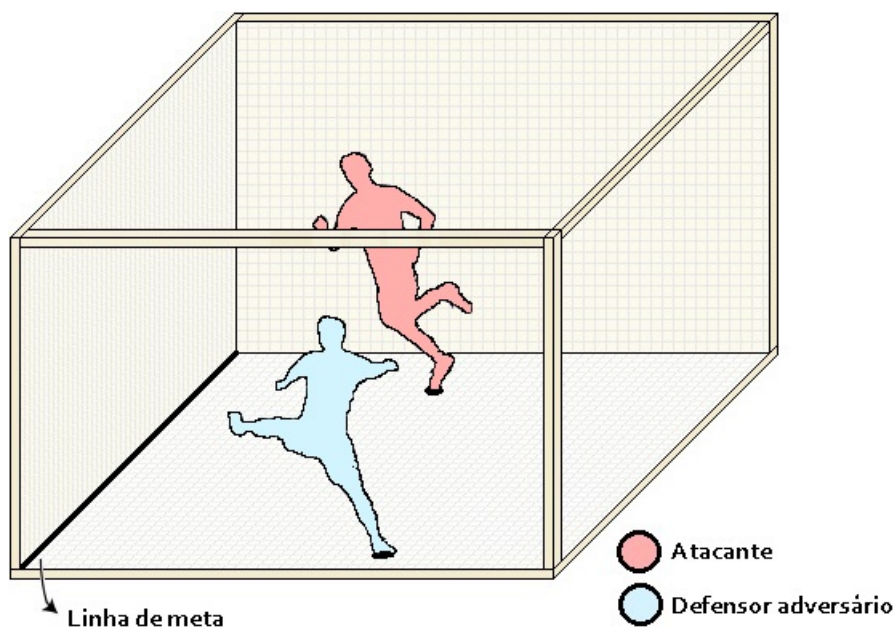
Figura 63 – Exemplo de definição de vetor correspondente a um jogador



Fonte – Autoria própria (2019)

Exemplo 3.16. Vamos verificar se o atacante da Figura 64 está em posição de impedimento.

Figura 64 – Verificando se atacante está em posição de impedimento

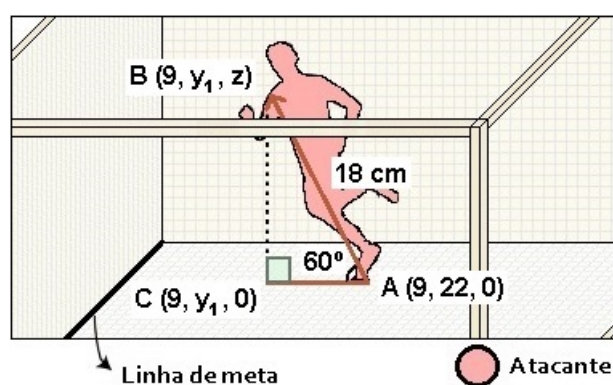


Fonte – Autoria própria (2019)

Inicialmente precisamos definir os vetores que corresponderão ao atacante e ao defensor, assim como as informações de norma, localização na base da caixa em termos das coordenadas do ponto de cada vetor pertencente ao plano da base, assim como suas respectivas inclinações

(com utilização de régua para medição do comprimento do vetor e de transferidor para medição do ângulo de inclinação dos vetores). Em seguida, modelamos o problema como um triângulo retângulo correspondente para cada jogador definido pelo vetor do jogador com sua respectiva projeção ortogonal em relação à base da caixa octante para assim aplicar definição de cosseno em ambos os triângulos (Figuras 65 e 66).

Figura 65 – Verificando projeção ortogonal do atacante

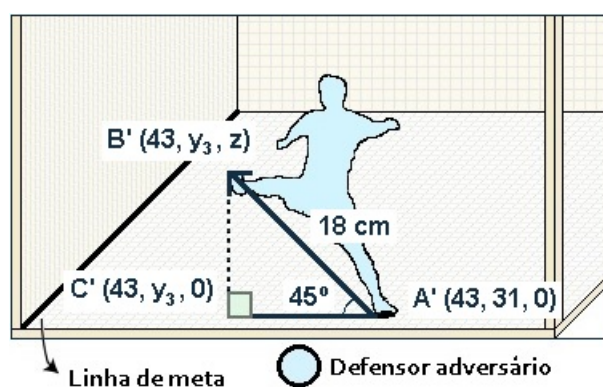


Fonte – Autoria própria (2019)

Neste caso, da equação 3.4, segue que $y_1 = 22 - 18 \cdot \cos 60^\circ = 22 - 9 = 13$.

Por outro lado, analisando o defensor temos que $y_3 = 31 - 18 \cdot \cos 45^\circ \approx 18,27$ (Figura 66).

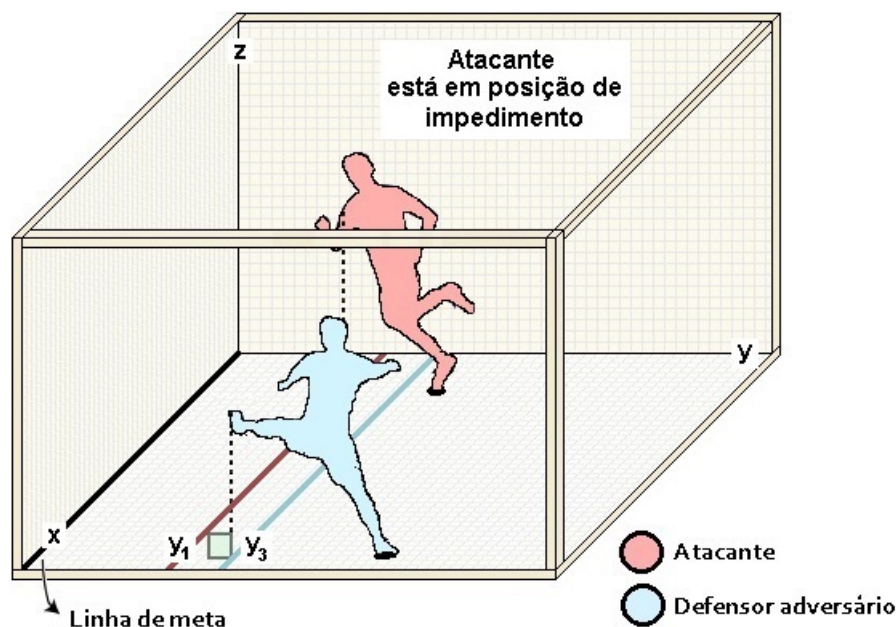
Figura 66 – Verificando projeção ortogonal do defensor



Fonte – Autoria própria (2019)

Portanto, como $y_1 = 13 < 18,27 \approx y_3$ concluímos que o atacante está em posição de impedimento. Neste caso, podemos dizer que o atacante está aproximadamente 5,27 cm mais adiantado que o último defensor (Figura 67).

Figura 67 – Verificação da existência de impedimento



Fonte – Autoria própria (2019)

3.4.3 AVALIAÇÃO

Acreditando que o processo avaliativo se dá como um todo e acontece processualmente durante as aulas, a avaliação da presente proposta de atividade se dará por participação e envolvimento dos alunos com questionamentos e contribuições e o envolvimento da turma na resolução dos problemas propostos durante toda a atividade, podendo o professor avaliar também todo o material produzido pela turma neste processo. Contudo, muito além dos registros feitos pelos estudantes, está a compreensão de cada etapa do processo que foi desenvolvido com a aplicação das etapas desta atividade. Essa compreensão, ainda que não apareça nos registros, é o que deve determinar a conclusão do processo avaliativo.

4 CONCLUSÃO

Este trabalho propôs uma série de atividades para o Ensino Médio onde através do estudo de projeções pretende-se reforçar tópicos de Geometria Analítica, principalmente, Trigonometria. Ademais, desenvolveu-se um material concreto para o trabalho (caixa octante) cujo custo é altamente compatível com a realidade das escolas nacionais, podendo inclusive ser produzido na própria escola e pelos próprios alunos se houver material e espaço para este fim. Cabe destacar que devido a suas características, a utilização do material concreto elaborado neste trabalho é bem ampla, não restringindo às atividades aqui propostas, podendo atender também alunos com deficiência visual.

Embora o fato do tema proposto (projeções) não estar no currículo, ele permite trabalhar com conteúdos curriculares, nomeadamente Trigonometria e Geometria Analítica, além de conteúdos de outras disciplinas de maneira interdisciplinar. Dessa maneira, pode-se afirmar que o que se propôs neste texto é a utilização do tema projeções de maneira motivacional para a aprendizagem de outros temas, estes sim curriculares.

Ao propor as atividades para o Ensino Médio, buscamos ter um cuidado especial em mostrar ao aluno de forma concreta e estimulante um conteúdo essencialmente abstrato com utilização de recursos concretos e dinâmicos, sem deixar de lado importantes fundamentos teóricos. Neste contexto, o estudante não só tem um envolvimento com o estudo de projeções como também o faz de forma lúdica, propiciando um ambiente que julgamos favorecer o processo de ensino-aprendizagem.

Sobre a proposta de trazer atividades diversificadas neste trabalho, com possibilidade de serem aplicadas de forma interdisciplinar, vamos ao encontro das Orientações Educacionais Complementares aos Parâmetros Curriculares Nacionais.

É importante uma preocupação consciente e explícita para atender adequadamente todos os alunos de uma classe heterogênea, propondo o trabalho diversificado na sala de aula e o trabalho coletivo dos diversos professores de um mesmo aluno. O trabalho diversificado pressupõe o reconhecimento de que a situação normal em uma sala de aula é a diferença de ritmo, de motivação e de formação, e de que queremos respeitar o direito de todos de acesso ao conhecimento. Finalmente, é importante lembrar que o desafio de fazer com que todos aprendam não é tarefa para um só professor, mas pressupõe o trabalho coletivo dos diferentes professores desses alunos e do envolvimento da escola em um projeto pedagógico comum. A Matemática tem papel relevante nessa ação coletiva porque frequentemente ela é mitificada por sua pretensa dificuldade. É importante deixar claro que todos podem aprendê-la (BRASIL, 2006b).

Esperamos que esta proposta possa ser uma ferramenta útil, para professores e alunos, no complemento do estudo de Geometria Analítica. Que os alunos possam perceber que mesmo conteúdos abstratos, como o de projeções por exemplo, possuem aplicações diversas, inclusive no

esporte, e que o correto entendimento dos conceitos matemáticos envolvidos acabam permitindo uma melhor compreensão do problema e contribuindo no desenvolvimento de uma solução adequada.

REFERÊNCIAS

- BEGIATO, R. G. **Um método Newton-inexato com estratégia híbrida para globalização**. Dissertação (Mestrado) — Universidade Estadual de Campinas, 2007. 54, 58
- BRASIL. Secretaria de Educação Básica. **Parâmetros curriculares nacionais: Ensino médio**. Brasília: Ministério da Educação, Secretaria de Educação Básica, 2000. 16, 85
- BRASIL. Secretaria de Educação Básica. **Orientações curriculares para o Ensino Médio: Ciências da natureza, matemática e suas tecnologias**. Brasília: Ministério da Educação, Secretaria de Educação Básica, 2006. 15, 84, 85
- BRASIL. Secretaria de Educação Básica. **Orientações educacionais complementares aos parâmetros curriculares nacionais (PCN+): Ciências da natureza, matemática e suas tecnologias**. Brasília: Ministério da Educação, Secretaria de Educação Básica, 2006. 117
- COELHO, F. U.; LOURENÇO, M. L. **Um curso de álgebra linear**. São Paulo, SP: EDUSP, 2013. 18
- DELFIM, T. F.; JESUS, V. L. B. de. O problema da simultaneidade na lei do impedimento do futebol. **Revista Brasileira de Ensino de Física**, Rio de Janeiro, v. 33, n. 4, 4308, dez. 2011. 105
- DEMMEL, J. W. **Applied numerical linear algebra**. Philadelphia: SIAM, 1997. 54
- GLOBOESPORTE. **Após gol polêmico do Inter, Bellintani pede imagens do VAR e descarta anulação da partida**. 2019. Publicado: 13 de jun. 2019 por Por GloboEsporte.com. Disponível em: <<https://globoesporte.globo.com/ba/futebol/times/bahia/noticia/apos-gol-polemico-do-inter-bellintani-pede-imagens-do-var-e-descarta-anulacao-da-partida.ghml>>. Acesso em: 23 jul. 2019. 109
- GREENBAUM, A. **Iterative methods for solving linear systems**. Philadelphia: SIAM, 1997. 54
- HEFEZ, A.; FERNADEZ, C. d. S. **Introdução à álgebra linear**. Rio de Janeiro: SBM, 2016. (Coleção PROFMAT). 18
- HESTENES, M. R.; STIEFEL, E. Methods of conjugate gradients for solving linear systems. **Journal of Research of the National Bureau of Standards**, Los Angeles, v. 49, n. 6, 2379, p. 409–436, dec. 1952. 54, 75
- IFAB. **Laws of the game: 2019/20**. Zurich: The International Football Association Board, 2019. 16, 105, 106, 107, 108
- KELLEY, C. T. **Iterative methods for linear and Nonlinear Equations**. Philadelphia: SIAM, 1995. 15, 54, 57, 73
- LAGO, R. F. **Estudos sobre os métodos iterativos de Krylov para solução de sistemas de equações lineares**. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, 2010. 54, 77

LIMA, E. L. **Álgebra linear**. 8. ed. Rio de Janeiro: IMPA, 2014. (Coleção matemática universitária). 18

NASCIMENTO, M. C. **Álgebra linear**. Bauru: UNESP, 2013. 18

O DOCUMENTO. **Presidente do Bahia critica VAR, mas descarta anulação de jogo contra o Inter**. 2019. Publicado: 13 de jun. 2019 por Da Redação. Disponível em: <<https://odocumento.com.br/presidente-do-bahia-critica-var-mas-descarta-anulacao-de-jogo-contra-o-inter/>>. Acesso em: 23 jul. 2019. 108

SAAD, Y. **Iterative methods for sparse Linear Systems**. [S.l.]: SIAM, 2003. 54, 57

SANTOS, R. d. J. **Álgebra linear e aplicações**. Belo Horizonte: Imprensa universitária da UFMG, 2006. 18, 40

SHEWCHUK, J. R. An introduction to the conjugate gradient method without the agonizing pain. School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213, Pensilvânia, p. 2–3, aug. 1994. 77

SPORTV. **Chefe de arbitragem da CBF divulga imagem e explica lance polêmico do VAR em Inter x Bahia**. 2019. Publicado: 13 de jun. 2019 por Por SporTV.com. Disponível em: <<https://sportv.globo.com/site/programas/redacao-sportv/noticia/chefe-de-arbitragem-da-cbf-divulga-imagem-e-explica-lance-polemico-do-var-em-inter-x-bahia.ghtml>>. Acesso em: 23 jul. 2019. 109, 110

TEIXEIRA, R. C. **Álgebra linear: exercícios e soluções**. 3. ed. Rio de Janeiro: IMPA, 2015. (Coleção matemática universitária). Soluções dos exercícios do livro Álgebra Linear de Elon Lages Lima. 18

TORCEDORES. **Internautas questionam uso polêmico do VAR na vitória do Inter sobre o Bahia**. 2019. Publicado: 13 de jun. 2019 por Andressa Fischer. Disponível em: <<https://www.torcedores.com/noticias/2019/06/internautas-questionam-var-inter-x-bahia>>. Acesso em: 23 jul. 2019. 109

TREFETHEN, L. N.; BAU, D. **Numerical linear algebra**. Philadelphia: SIAM, 1997. 54