UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE EDUCAÇÃO E SAÚDE

UNIDADE ACADÊMICA DE EDUCAÇÃO

Curso de Graduação em Licenciatura em Matemática

José Filho do Nascimento

TEOREMA ESPECTRAL:

Operadores Autoadjuntos e Aplicação.

Cuité-PB

2013

José Filho do Nascimento

TEOREMA ESPECTRAL:

Operadores Autoadjuntos e Aplicação.

TCC apresentado ao curso Graduação em Licenciatura em Matemática do Centro de Educação e Saúde da Universidade Federal de Campina Grande em cumprimento às exigências do Componente Curricular Trabalho Acadêmico Orientado, para obtenção do grau de Graduado em Licenciatura em Matemática.

Orientadora: Maria Gisélia Vasconcelos
Coorientadora: Márcia Cristina Silva Brito

Cuité-PB

2013

FICHA CATALOGRÁFICA ELABORADA NA FONTE
Responsabilidade Jesiel Ferreira Gomes – CRB 15 – 256

N244t Nascimento, José Filho do.

Teorema espectral: operadores autoadjuntos e aplicação. / José Filho do Nascimento – Cuité: CES, 2013.

60 fl.

Monografia (Curso de Licenciatura em Matemática) – Centro de Educação e Saúde / UFCG, 2013.

Orientadora: Maria Gisélia Vasconcelos.
Co-orientadora: Márcia Cristina Silva Brito.

1. Teorema espectral. 2. Operador autoadjunto. 3. Máximo e mínimo local. I. Título.

CDU 51

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE – UFCG
CENTRO DE EDUCAÇÃO E SAÚDE - CES
UNIDADE ACADÊMICA DE EDUCAÇÃO – UAE

José Filho do Nascimento

Teorema Espectral: operadores autoadjunto e aplicações

Monografia de Trabalho de Conclusão de Curso submetida à banca examinadora como parte dos requisitos necessários a obtenção do grau de Graduação em Licenciatura em Matemática.

A citação de qualquer trecho deste trabalho é permitida, desde que seja feita de conformidade com as normas de ética científica.

Trabalho de Conclusão de Curso (TCC) aprovado em 17 de setembro de 2013.

Banca Examinadora

Prof. Marco Aurélio Soares Souto

Profª. Maria Giselia Vasconcelos
(Orientadora)

Profª. Márcia Cristina Silva Brito
(Coorientadora)

Agradecimentos

Agradeço em primeiro lugar a Deus, que tem me dado forças para superar cada obstáculo enfrentado até aqui.

Agradeço a toda minha família, principalmente aqueles que me incentivam de forma direta ou indiretamente para que eu concluísse essa etapa de minha vida.

Agradeço a Escola Estadual de Ensino Fundamental e Médio Orlando Venâncio dos Santos, em nome da Professora Maria Aparecida Dantas (Cida), que me concebeu os Estágios Supervisionados I, II e III em suas salas de aula, na referida escola.

Agradeço ao Subprojeto - PIBID/Matemática, em nome do Coordenador de Área o Professor Alecxandro Alves Vieira. Projeto este, que serviu-me de incentivo, fortalecendo ainda mais o meu desejo de exercer esta profissão.

Agradeço a todos os meus colegas de turma, especialmente a Bosoerg, Elizângela, Fagner e Leonardo, pois estes tiveram uma fundamental importância para que eu chegasse até aqui.

Agradeço aos meus conterrâneos Bruno Linhares, Edclebeson Berto e Natália Ferreira, pelos os incentivos, as horas de descontração e por terem participado dessa caminhada junto comigo.

Agradeço a Família Maneiros, em nome de Jeferson Rafael, pelo acolhimento nesta cidade, e por tudo que fizeram para com a minha pessoa.

Agradeço ao Professor Marco Aurélio Soares Souto, pela disponibilidade de está presente na mesa examinadora deste trabalho.

E por último agradeço a todos os Professores do Centro de Educação e Saúde (CES), da Universidade Federal de Campina Grande (UFCG), Campus - Cuité, que tive a oportunidade de estudar com eles, em especial as Professoras Maria Gisélia Vasconcelos e Márcia Cristina Silva Brito, por terem confiado em meu trabalho e por tudo que fizeram para que eu conseguisse concluir essa etapa de minha vida.

Dedicatória

A toda minha família, principalmente aos meus pais
José Francisco do Nascimento e Francisca Sebastiana do Nascimento.

"A matemática se revela em mentes sensíveis, capazes de ver espirais em um girassol, ângulos em uma estrela e Deus no infinito."

Manoel Rodrigues Paiva

Resumo

Este trabalho é dedicado ao estudo de condições para diagonalização de Operadores lineares, definidos em espaços vetoriais reais ou complexos. O resultado principal é o teorema Espectral para Operadores Lineares, que dá condições para a diagonalização. Definiremos os conceitos prévios sobre: Espaço Vetorial, Base, Dimensão, Produto Interno, Autovalores e Autovetores e Operadores Autoadjuntos. Conceitos estes usados no Teorema Espectral, onde o mesmo nos diz que "Para todo operador autoadjunto T definido de V em V, num espaço vetorial de dimensão finita munido do produto interno, existe uma base ortonormal de vetores contida em V formada por autovetores de T". Como aplicação, usamos o conceito de máximo e mínimo local e uma forma quadrática para classificação de pontos de uma função de várias variáveis.

Palavras-chave: Operador autoadjunto. Teorema Espectral. Máximo e Mínimo Local.

Abstract

This work is dedicated to the study of conditions for diagonalization of Linear Operators, defined in real or complex vector spaces. The main result is the Spectral Theorem for Linear Operators, which gives conditions for diagonalization. We define the concepts about: Vector Spaces, Basis, Dimension, Internal Product, Eigenvalues and Eigenvectors and Selfadjoint Operators. These concepts used in the Spectral Theorem, where it tells us that "For every selfadjoint operator T defined of V on V, in a vector space of finite dimensional equipped with the inner product, there is an orthonormal basis of vectors contained in V consisting of eigenvectors of T". As an application, we used the concept of maximum and local minimum and a quadratic form to classify points of a function of several variables.

Keywords: Selfadjoint Operator. Spectral Theorem. Maximum and Minimum Local.

Sumário

Introdução

Teorema Espectral. Origens e evolução.

A evolução da teoria espectral é um dos capítulos mais informativos na história da matemática moderna. O tema central do teorema espectral é que certos operadores lineares de dimensão infinita podem ser representado na forma diagonal. Embora o teorema tem raízes profundas no passado, a teoria matemática é um fenômeno do século XX. Todo aluno de matemática estuda o teorema espectral extraído de seu contexto histórico, mas incorporado no contexto do curso. Este esquema, embora pedagogicamente eficiente, muitas vezes obscurece o fato do teorema espectral fornecer teorias matemáticas apropriadas para vários fenômenos físicos.

As raízes históricas do teorema espectral pode ser encontrado em três áreas distintas, tais como geometria analítica (teorema dos eixos principais), análise (equações integrais) e álgebra (sistemas infinitos de equações lineares).

O espectro de um operador na matemática contemporânea surge da tentativa de compreender vários problemas concreto de álgebra envolvendo soluções das equações lineares e suas generalizações a dimensão infinita. Esta visão é o resultado da evolução matemática de uma série de ideias ao longo de quase 300 anos.

Teoria Espectral

O tema dos autovalores aparece quando Euler, no primeiro terço do século XVIII estudou sistematicamente a equação do segundo grau geral de duas e três variáveis no plano e no espaço, respectivamente. Ele demonstrou que existem eixos perpendiculares onde a expressão da cônica ou quádrica é particularmente simples. Posteriormente em

1760 em seu livro *Recherches sur la courbure des surfaces*, estudando as secções normais de uma superfície em um ponto encontrou que existem dois planos mutuamente ortogonais cujas secções determinam as curvaturas máxima e mínima das curvas. Posteriormente, verificou-se que estas duas situações são casos particulares do fato de uma matriz simétrica ser ortogonalmente diagonalizável. A noção de polinômio característico aparece explicitamente na obra de Lagrange sobre os sistemas de equações diferenciais em 1774 e nos trabalhos de Laplace (1749-1827) em 1775.

Cauchy reconheceu o problema do autovalor no trabalho de Euler, Lagrange e Laplace. Em 1826, considerou o problema da redução de uma forma quadrática de três variáveis e mostrou que a equação característica é invariante, para mudança dos eixos retangulares, em linguagem moderna, se A é uma matriz quadrada e se S é inversível, então

$$det(A - \lambda I) = det(SAS^{-1} - \lambda I).$$

Em 1829 Cauchy provou que os autovalores de uma matriz simétrica são reais. As matrizes Hermitianas ($A = \overline{A}^t$) foram introduzidas por Hermite (1822-1901). Frobenius em 1878, garantiu diagonalização de matrizes ortogonais, estendendo em 1883 a demonstração a matrizes unitárias ($A\overline{A}^t = I$). O teorema espectral para matrizes normais ($A\overline{A}^t = \overline{A}^t A$) é devido ao Toeplitz (1881 - 1940).

Jacobi (1804-1851) deu a solução do sistema de equações diferenciais $Y' = AY$, sendo A uma matriz diagonalizável. Jordan resolveu o caso que pode ser diagonalizável usando conceitos de matrizes semelhantes (duas matrizes A e B são semelhantes se existe uma matriz invertível S tal que $A = SBS^{-1}$) e de equação característica. No livro *Traité des substituions* (1870) demonstrou que uma matriz pode ser transformado numa forma canônica agora chamado de forma canônica de Jordan.

Um passo simultâneo em direção ao conceito de autovalor e autovetor num espaço vetorial abstrato é o trabalho de Sturm e Liouville que estudaram equações que agora levam seu nome. Os problemas estudados por Sturm e Liouville pode ser divididos em três tipos: 1) As propriedades dos autovalores, 2) comportamento qualitativo das funções (ou funções próprias) e 3) a expansão de funções arbitrárias em termos destas funções. Dessas três questões, os dois primeiros foram estudados por Sturm e o terceiro por Liouville. Eles observaram que, se ϕ é um operador diferencial, então existe uma sequência de valores λ_n de tal modo que existem funções e y_n não nulas ortogonais entre si verificando $\phi(y_n) = \lambda_n y_n$.

De 1904-1910, Hilbert estudou a equação integral $u(x) = \lambda \int_a^b K(x,y)u(y)dy$. Supondo que K é simétrica e definiu o que é um operador de autoadjunto para um espaço de função, permitindo fazer uso das propriedades das matrizes simétricas no caso finito. Em particular, mostrou que o operador

$$\phi(u)(x) = \int_a^b K(x,y)u(y)dy$$

é autoadjunto. As autofunções associados a dois autovalores distintos são perpendiculares dois a dois. Com estes resultados Hilbert demonstrou o conhecido teorema dos eixos principais para espaços de dimensão infinita. Hilbert usou um processo de passar limite que permitiu generalizar resultados sobre sistemas finitos equações lineares. Com base nisso observou que o tratamento das formas quadrática infinitas "iria complementar de uma maneira geral a bem conhecida teoria das formas quadrática em um número finito às variáveis".

O período entre 1740 e 1840 é um período particularmente importante no estabelecimento de teoria espectral. Mas isso não quer dizer que é impossível encontrar teoria espectral antes de 1740. É possível, considerar os problemas da teoria espectral como um ponto de partida para o desenvolvimento da matemática e traçar suas origens até os séculos V e VI aC. com o trabalho realizado pelos pitagóricos na corda vibrante. Acreditamos que este problema está claramente ligado ao que temos discutido, não é parte do que consideramos como a história da teoria espectral enquanto teoria matemática.

Observamos também que, após o impressionante trabalho feito por Sturm, sem dúvida, tinha muito a fazer na teoria espectral. A teoria de Sturm e Liouville demorou a ser aceite, devido a genialidade da mesma, ou seja, a natureza qualitativa dos resultados e verificou-se que enfrentou um forte problema, pois houve uma falta geral de teoremas de existência em meados do século XIX. Diante disso, só no final do século XX é que a teoria de Sturm- Liouville foi aceito e continuada.

Além disso, entre 1836 e 1900 houve grandes avanços em outros campos desta teoria, avanços de um certo ponto de vista permitindo a comunidade matemática finalmente amadurecer e apreciar o teoria criada por Sturm e Liouville.

O trabalho de Cauchy em 1829 foi crucial para o desenvolvimento da Teoria

Espectral, não só pelos resultados encontrados e a motivação que levou à grande avanços na teoria espectral de matrizes em 1870, com base no trabalho desenvolvido por Cayley, Hermite, Sylvester (que foi o primeiro a usar a palavra matriz, em 1850), Frobenius e Weierstrass a quem devemos a criação da teoria da matriz.

Finalmente , podemos dizer que a criação da teoria espectral culmina no início do século XX, quando o objeto de estudo, o **espectro** é finalmente definido por Hilbert.[1] A Teoria espectral, passou a ser utilizado 1930, cerca de um século depois do aparecimento das memórias Sturm.

Aplicações da teoria espectral

Cauchy percebeu a estreita relação entre autovalores e autovetores de uma matriz simétrica com as direções principais e o comprimento dos eixos da cônica associada a esta matriz simétrica.

Uma das primeiras aplicações do teoria dos autovalores e autovetores foi o estudo das sequências dada por recorrência linear, por exemplo, a sequência de Fibonacci.

$$a_1 = 1, a_2 = 1, a_{n+2} = a_{n+1} + a_n \quad \forall \quad n \geq 1.$$

A técnica que ainda é usado hoje em dia reduz ao cálculo de uma matriz de energia.

Markov (1856-1922) foi o primeiro a estudar processos estocásticos não dependentes de tempo, chamado hoje cadeias de Markov. Uma cadeia de Markov é um sucessão de variáveis dependentes $X(t_i) = (x_1(t_i), \ldots, x_n(t_i))$ identificados por valores discreto crescentes de t_i (normalmente, o tempo), com a propriedade de que qualquer previsão de $X(t_i)$ é apenas uma função de $X(t_{i-1})$. Isto é, o valor futuro da variável X depende apenas o valor presente e n ao dos valores passado. Usando a teoria da diagonalização de matrizes Markov estudou completamente as cadeias de Markov onde a relação entre $X(t_i)$ e $X(t_{i-1})$ é linear. Seu trabalho também foi aplicado na biologia. Em 1945, Leslie, introduziu um certo tipo de matrizes (chamadas matrizes de Leslie), a fim de estudar os problemas de Evolução das populações de animais.

[1]Hilbert utiliza a palavra grande **Spektrum** no seu trabalho de 1906 para falar sobre formas quadráticas em um número infinito de variáveis.

Mecânica Quântica

A **mecânica quântica** é a teoria física que obtêm sucesso no estudo dos sistemas físicos cujas dimensões são próximas ou abaixo da escala atômica, tais como **moléculas, átomos, elétrons, prótons** e de outras **partículas subatômicas**, muito embora também possa descrever fenômenos macroscópicos em diversos casos.

Muitos fenômenos quânticos difíceis de se imaginar concretamente podem ser compreendidos com um pouco de abstração matemática. Há três conceitos fundamentais da matemática - mais especificamente da álgebra linear - que são empregados constantemente pela mecânica quântica. São estes: (1) o conceito de operador; (2) de autovetor; e (3) de autovalor.

As **formulações matemáticas da mecânica quântica** são os formalismos matemáticos que permitam uma descrição rigorosa da mecânica quântica. Estas, por sua vez, se distinguem do formalismo matemático da mecânica clássica, (antes do início de 1900) pelo uso de estruturas matemáticas abstratas, tais como **espaços de Hilbert** de dimensão infinita e operadores sobre estes espaços. Muitas destas estruturas são retiradas da **análise funcional**, uma área de pesquisa dentro matemática que foi influenciada, em parte, pelas necessidades da mecânica quântica. Em resumo, os valores observáveis físicos, tais como **energia e momento** já não eram considerados como valores de funções em espaço de fase, mas como autovalores, mais precisamente: como valores espectrais de **operadores lineares** no espaço de Hilbert.

Em 1925-1926, W. Heisenberg e E. Schrödinger criou teoria mecânica quântica em Göttingen. Na teoria de Heisenberg certas observações atômicas não podem serem feitas simultaneamente isto foi interpretado matematicamente como a n ao comutatividade das operações que os representam.

Como a álgebra matricial não é comutativa Heisenberg juntos com M. Born e P. Jordan representaram cada grandeza física por uma matriz apropriada. O conjunto de valores possíveis da referida quantidade física era o espectro da transformação. Em contraste, Schrödinger iniciou uma teoria baseada na equação da onda. Depois da surpresa inicial, que a mecânica da onda de Schrödinger e mecânica matricial de Heisenberg - essencialmente duas teorias com pressupostos diferentes - devem produzir os mesmos resultados, Schrödinger unificou as duas abordagens provando que os autovalores do operador diferencial na equação de onda determina a correspondente matriz de Heisenberg, resultados também obtidos simultaneamente por P. Dirac. Foi encontrado que a energia de um átomo pode ser expresso por uma forma quadrática, cujo espectro é observável por um espectroscópio.

Tudo isso renovou o interesse pela a teoria espectral. O próprio Hilbert ficou espantado que os espectros de suas formas quadráticas poderiam ser interpretada como espectros atômicos. Como ele disse: *"Quando eu desenvolvi minha teoria de variáveis infinitas não previa mais tarde encontrar aplicações para espectro da física"*.

Ficou clara que a teoria espectral de Hilbert era a base matemática apropriada para as nova mecânica. Matrizes finitas e infinitas foram interpretados como operadores no espaço de Hilbert e quantidades físicas são representados por esses operadores.

A maquinaria matemática da mecânica quântica tornou-se a ser a análise espectral e a renovada atividade precipitou a publicação por A. Wintner (1929), do primeiro livro dedicado à teoria espectral.

Embora esta abordagem poderia explicar com sucesso a nova mecânica quântica naquele momento, representou uma extensão para o Teorema espectral e realmente ajudou a estender a teoria Hilbert a operadores ilimitados.

Capítulo 1

Tópicos de Álgebra Linear

O objetivo deste capítulo é revisar algumas noções básicas de Álgebra Linear (espaços vetoriais, subespaços, bases, dimensão, transformações lineares, etc.), que vamos precisar no decorrer deste trabalho. Em particular, não faremos as demonstrações de alguns resultados aqui enunciados. Indicamos os textos [7] , [8] e [15] para maiores detalhes.

1.1 Espaços Vetoriais

O corpo $\mathbb{R}$ ou o corpo $\mathbb{C}$ serão denotados por $\mathbb{K}$.

Definição 1.1. *Um conjunto não vazio V é um* **espaço vetorial** *sobre o corpo $\mathbb{K}$ (denotamos também como $\mathbb{K}$ - espaço vetorial) se seus elementos (chamados* **vetores***) podem ser somados e multiplicados por escalares (elementos do corpo). Ou seja, dizemos que V é um espaço vetorial sobre o corpo $\mathbb{K}$ se estiverem definidas as seguintes duas operações:*

$$\begin{array}{rccc} (+): & V \times V & \longrightarrow & V \\ & (u,v) & \longmapsto & u+v \end{array} \qquad \begin{array}{rccc} (\cdot): & \mathbb{K} \times V & \longrightarrow & V \\ & (k,u) & \longmapsto & k \cdot v \end{array}$$

tal que $\forall\ u, v, w \in V$ e $\forall\ \alpha, \beta \in \mathbb{K}$ temos as seguintes propriedades:

$(A1)$ $u + v = v + u$ *(comutativa)*

$(A2)$ $(u + v) + w = u + (v + w)$ *(associativa)*

$(A3)$ $\exists\ 0 \in V$, *tal que* $u + 0 = u$ *(vetor nulo)*

(A4) $\forall\ u \in V,\ \exists\ -u \in V$ *tal que* $u + (-u) = 0$ *(vetor oposto)*

(P5) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$ *(associativa)*

(P6) $1 \cdot v = v$ *(1 é o elemento identidade de* $\mathbb{K}$*)*

(P7) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ *(distributiva)*

(P8) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ *(distributiva)*

Vejamos abaixo alguns exemplos de espaços vetoriais.

Exemplo 1.1. *O conjunto* $\mathbb{K}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{K}(i = 1, 2, \ldots, n)\}$ *com as definições usuais de adição e multiplicação por escalar é um espaço vetorial.*

Exemplo 1.2. *O* $\mathbb{R}^n$ *com as operações:*

$$\begin{aligned} (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ k \cdot (x_1, x_2, \ldots, x_n) &= (k \cdot x_1, k \cdot x_2, \ldots, k \cdot x_n) \end{aligned}$$

é um espaço vetorial pelo exemplo acima.

Exemplo 1.3. *O conjunto das matrizes reais de ordem* $m \times n$*, como as operações usuais é um espaço vetorial, tal que o elemento neutro da adição é a matriz nula.*

Definição 1.2. *Seja* V *um espaço vetorial sobre* $\mathbb{K}$*. Seja um subconjunto* $W \neq \emptyset$ *de* V*. Dizemos que* W *é* **subespaço vetorial** *de* V *se e somente se são válidas as seguintes condições:*

(i) $0 \in W$;

(ii) se $u, v \in W$ *então* $u + v \in W$;

(iii) se $\lambda \in \mathbb{K}$ *e* $v \in W$ *então* $\lambda \cdot v \in W$.

A restrição das operações de V *e* W *torna esse subconjunto um* $\mathbb{K}$ *- espaço vetorial.*

Todo espaço vetorial V admite pelo menos dois subespaços: o próprio espaço V e o conjunto $\{0_v\}$, chamado subespaço nulo. Estes dois subespaços são denominados **subespaços triviais** ou **impróprios** de V. Os demais subespaços de V são chamados de subespaços próprios de V.

Definição 1.3. *Sejam W_1 e W_2 dois subespaços vetoriais de um espaço vetorial V. Dizemos que a soma $W_1 + W_2$ é* **direta** *se $W_1 \cap W_2 = \{0\}$ e, neste caso, escrevemos $W_1 \oplus W_2$.*

Definição 1.4. *Seja V um $\mathbb{K}$ - espaço vetorial e sejam W_1, W_2 subespaços de V. Dizemos que V é a* **soma direta** *de W_1 e W_2 se $V = W_1 \oplus W_2$.*

Proposição 1.1. *Seja $V = W_1 \oplus W_2$ um $\mathbb{K}$ - espaço vetorial. Então todo elemento $v \in V$ se escreve de maneira única como uma soma $v = w_1 + w_2$ com $w_1 \in W_1$ e $w_2 \in W_2$.*

Demonstração. Ver [8]. □

Nesta seção vamos discutir um dos conceitos mais importantes envolvendo a estrutura de um espaço vetorial. Antes de definir o conceito de base, iremos definir os seguintes conceitos: combinação linear, conjunto gerador, independência linear e dependência linear. A partir do conceito de base definiremos dimensão.

Definição 1.5. *Seja V um espaço vetorial sobre $\mathbb{K}$.*

(*i*) *Um vetor $v \in V$ é uma* **combinação linear** *dos vetores $v_1, \ldots, v_n \in V$ se existirem escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que:*

$$v = \alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n = \sum_{i=1}^{n} \alpha_i \cdot v_i.$$

(*ii*) *Seja B um subconjunto de V. Dizemos que B é um* **conjunto gerador** *de V (ou que B* **gera** *V) se, para todo $v \in V$, existirem (finitos) elementos $v_1, \ldots, v_n \in B$ e escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que $v = \alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n$. Denotamos por $[B] = V$.*

Definição 1.6. *Seja V um $\mathbb{K}$ - espaço vetorial e B um subconjunto de V.*

(*i*) *Dizemos que B é* **linearmente independente** *(L.I.) se $\alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n = 0$, para $v_i \in B$ e $\alpha_i \in \mathbb{K}, i = 1, \ldots, n$, implica que $\alpha_1 = \ldots = \alpha_n = 0$ é a única solução.*

(*ii*) *O conjunto B é* **linearmente dependente** *(L.D.) se não for L.I., ou seja, existe $\alpha_i \in \mathbb{K}^*$, $i = 1, \ldots, n$ tal que $\alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n = 0$.*

Definição 1.7. *Seja V um $\mathbb{K}$ - espaço vetorial. Dizemos que um conjunto $\beta \subset V$ é uma* **base** *de V se são válidas as condições seguintes:*

(*i*) *β gera V ($[\beta] = V$)*

(*ii*) *β for L.I.*

Definição 1.8. *Um espaço vetorial V é de dimensão finita se e somente se V possui uma base finita. Ou seja, o número n de elementos de uma base finita de V chama-se* **dimensão** *de V, onde denotaremos por $dim(V)$. Caso contrário, dizemos que V tem dimensão infinita.*

Exemplo 1.4. *Em $\mathbb{R}^2$, $\beta = \{(1,0),(0,1)\}$ é uma base de $\mathbb{R}^2$, também chamada de* **base canônica** *do espaço $\mathbb{R}^2$. Como essa base possui dois elementos, então $dim(\mathbb{R}^2) = 2$. Em geral, $dim(\mathbb{R}^n) = n$. Uma base para o $\mathbb{R}^2$ pode ser a base canônica $\beta = \{e_1, e_2, \dots, e_n\}$, onde $e_1 = (1,0,0,\dots,0)$, $e_2 = (0,1,0,\dots,0), \dots, e_n = (0,0,0,\dots,1)$.*

Exemplo 1.5. *Exemplos de dimensão:*

(*a*) *$dim(\mathbb{K}^n) = n$ ($\mathbb{K}^n$ é um $\mathbb{K}$ - espaço vetorial).*

(*b*) *$dim(\mathbb{C}^n) = n$ quando $\mathbb{C}^n$ é um $\mathbb{C}$ - espaço e $dim(\mathbb{C}^n) = 2n$, quando for um $\mathbb{R}$ - espaço.*

(*c*) *$dim(\mathbb{M}_{m\times n}(\mathbb{C})) = m \cdot n$, quando $\mathbb{M}_{m\times n}(\mathbb{C})$ é um $\mathbb{C}$ - espaço e $dim(\mathbb{M}_{m\times n}(\mathbb{C})) = 2 \cdot m \cdot n$, quando for um $\mathbb{R}$ - espaço.*

Proposição 1.2. *Sejam V um espaço vetorial e U e W dois subespaços vetoriais de V, ambos de dimensão fiita. Então*

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

Demonstração. Ver [8]. □

1.2 Transformações Lineares

As aplicações mais notáveis entre espaços vetoriais são aquelas que preservam sua estrutura. Tais aplicações se denominam lineares.

Definição 1.9. *Sejam U e V $\mathbb{K}$ - espaços vetoriais. Uma aplicação $T : U \longrightarrow V$ é uma* **transformação linear** *se são válidas as condições:*

(i) $T(u_1 + u_2) = T(u_1) + T(u_2)$, $\forall\ u_1, u_2 \in U$.

(ii) $T(\lambda \cdot u) = \lambda \cdot T(u)$, $\forall\ \lambda \in \mathbb{K}$ e $\forall\ u \in U$.

Teorema 1.1. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$, com $dim\ V = n$ e $\beta = \{v_1, \ldots, v_n\}$ uma base ordenada para V e seja $w_1, \ldots, w_n$ elementos arbitrários de W. Então, existe uma única transformação linear $T : V \longrightarrow W$ tal que*

$$T(v_j) = w_j \ \ para \ \ j = 1, \ldots, n.$$

Demonstração. Ver [8]. □

Observação 1.1. *Sejam U e V $\mathbb{K}$ - espaços vetoriais. Seja $T : U \longrightarrow V$ uma transformação linear.*

1. *Se W é um subespaço vetorial de U, então a imagem de W por T é um subespaço de V;*

2. *Se $U = V$ então T é chamado de* **operador linear***;*

3. *Se $V = \mathbb{K}$ então T é chamado de* **funcional linear***;*

4. *Se T for uma bijeção, dizemos que T é um isomorfismo e que os espaços U e V são* **isomorfos**.

Definição 1.10. *Seja $T : U \longrightarrow V$ uma transformação linear. Definimos a* **imagem** *de T (denotada por $Im(T)$) por*

$$Im(T) = \{v \in V; T(u) = v, u \in U\}.$$

Definimos o **núcleo** *de T (denotado por $ker(T)$) por*

$$ker(T) = \{u \in U; T(u) = 0\}.$$

Proposição 1.3. *Sejam U e V espaços vetoriais sobre $\mathbb{K}$ e $T : U \longrightarrow V$ uma transformação linear. Então*

(i) *$ker(T)$ é um subespaço de U e a $Im(T)$ é um subespaço de V;*

(ii) *T é injetora se e somente se* $ker(T) = \{0\}$.

Demonstração. Ver [7]. □

Teorema 1.2. *(Do Núcleo e da Imagem). Sejam U e V espaços vetoriais sobre o corpo* $\mathbb{K}$ *com* $dim_{\mathbb{K}}U$ *finita e* $T : U \longrightarrow V$ *uma transformação linear. Então*

$$dim_{\mathbb{K}}(U) = dim_{\mathbb{K}}ker(T) + dim_{\mathbb{K}}Im(T).$$

Demonstração. Ver [8]. □

Corolário 1.1. *Sejam U e V espaços vetoriais de mesma dimensão. Seja* $T : U \longrightarrow V$ *uma transformação linear. Então as seguintes afirmações são equivalentes:*

(i) *T é injetiva;*

(ii) *T é sobrejetiva;*

(iii) *T é bijetiva;*

(iv) *T leva bases em bases, ou seja, se* β *é uma base de U então* $T(\beta)$ *é uma base de V.*

Demonstração. Ver [7]. □

Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$. Vamos considerar $\beta = \{v_1, \ldots, v_n\}$ uma base ordenada para V e $\gamma = \{w_1, \ldots, w_m\}$ uma base ordenada para W. Seja T : V $\longrightarrow$ W uma aplicação linear, pelo Teorema 1.1 sabemos que T fica bem determinada pelo seu efeito sobre os elementos da base β de V. Assim, cada elemento $T(v_j) \in W$ pode ser escrito de modo único da forma:

$$T(v_j) = \sum_{i=1}^{m} \alpha_{ij} \cdot w_i \quad para \quad j = 1, \ldots, n,$$

onde os escalares $a_{1j}, \ldots, a_{mj} \in \mathbb{K}$ são as coordenadas do elemento $T(v_j)$ com relação à base ordenada de W. Os coeficientes $a_{ij}, i = 1, \ldots, n, j = 1, \ldots, m$, podem ser organizados como uma matriz $m \times n$

$$A = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

cuja j-ésima coluna são as coordenadas do elemento $T(v_j)$ com relação à base ordenada de W.

Definição 1.11. *A matriz* $A = (a_{ij})_{i,j} \in \mathbb{M}_{m\times n}(\mathbb{M})$ *definida acima é chamada de* **matriz da transformação linear** T *com relação a base* β *e* γ *e é denotada por* $[T]_{\gamma}^{\beta} = [a_{ij}]$.

No caso em que V e W e as bases β e γ sejam iguais denotamos $[T]_{\gamma}^{\beta}$ simplesmente por $[T]_{\beta}$.

Seja V um $\mathbb{K}$ - espaço de dimensão $n \geq 1$ e sejam $\beta = \{u_1, \dots, u_n\}$ e $\beta' = \{v_1, \dots, v_n\}$ duas bases de V. Considere a matriz $M = (a_{ij})_{i,j} = [Id]_{\beta'}^{\beta}$ associada à transformação identidade com relação às bases β e β', isto é, a matriz dada pelos coeficientes

$$\begin{cases} u_1 &= a_{11}v_1 + a_{21}v_2 + \ldots + a_{n1}v_n &= \sum_{i=1}^{n} \alpha_{i1}v_i \\ \vdots & & \vdots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \ldots + a_{nn}v_n &= \sum_{i=1}^{n} \alpha_{in}v_i \end{cases}$$

Com isso, se $v \in V$ e escrevendo $v = (\alpha_1, \dots, \alpha_n)_{\beta} = (\lambda_1, \dots, \lambda_n)_{\beta'}$ as coordenadas de v com relação às bases β e β', teremos

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}_{\beta} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}_{\beta'}$$

isto é, a multiplicação de M pelas coordenadas de v na base β fornece as coordenadas de v na base β'. Tal matriz é chamada de **matriz de mudança de bases** de β' para β.

Observação 1.2. *Seja* $T : V \longrightarrow V$ *uma transformação linear, e sejam* β *e* β' *duas bases de* V *(assuma* $dim_{\mathbb{K}}V = n \geq 1$*). Se* P *é a matriz de mudança de bases de* β *para* β'*, teremos:*

$$[T]_{\beta} = P^{-1} \cdot [T]_{\beta'} \cdot P. \tag{1.1}$$

Duas matrizes M e N são ditas **semelhantes** se existir uma matriz invertível P tal que $M = P^{-1}NP$. Pelo que acabamos de ver, as matrizes $[T]_\beta$ e $[T]_{\beta'}$ são semelhantes.

1.3 Espaços com Produto Interno

O conceito de produto interno em um espaço vetorial real (complexo) é uma generalização do conceito de produto escalar definido em $\mathbb{R}^n$. Um conceito central que estudaremos em particular é o de ortogonalidade.

Definição 1.12. *Seja V um $\mathbb{K}$ - espaço vetorial, onde $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. E uma aplicação $\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{K}$ que satisfaz as seguintes propriedades:*

$(P1)$ $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle; \forall\ u, v, w \in V.$

$(P2)$ $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle; \forall\ u, v \in V$ *e* $\forall\ \lambda \in \mathbb{K}$.

$(P3)$ $\langle u, v\rangle = \overline{\langle v, u\rangle}, \forall\ u, v \in V.$

$(P4)$ $\langle u, u\rangle \geq 0; \forall\ u \in V$ *com* $\langle u, u\rangle = 0 \Leftrightarrow u = 0_V.$

define um **produto interno** *no espaço vetorial complexo V.*

Um espaço vetorial real de dimensão finita dotado de um produto interno é frequentemente chamado um **espaço euclidiano**. Um espaço vetorial complexo dotado de um produto interno é frequentemente chamado de **espaço hermitiano** e o seu produto interno é as vezes chamado de **produto hermitiano**.

Podemos verificar que com as propriedades de simetria hermitiana, distributividade e homogeneidade temos que:

- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ para todos $u, v, w \in V$.
- $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ para todos $u, v \in V$ e $\lambda \in \mathbb{K}$.

A necessidade de tomar conjugados no caso complexo em (iii) justifica-se para assegurar consistência com (iv). De fato, se valesse simplesmente $\langle x, y\rangle = \langle y, x\rangle$ e (iv), teríamos

$$0 < \langle ix, ix\rangle = i^2\langle x, x\rangle = -\langle x, x\rangle < 0.$$

A condição (iv) tem prioridade sobre a comutatividade do produto interno, isto é, abdicamos da comutatividade $\langle x, y\rangle = \langle y, x\rangle$ a fim de que (iv) valha, porque é esta última propriedade que nos permite definir uma noção de norma de vetor a partir do produto interno.

Definição 1.13. *Seja V um $\mathbb{K}$ - espaço vetorial munido de um produto interno $\langle\ ,\ \rangle$. Para cada $v \in V$, chamamos de* **norma** *de v ao número real dado por*

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Exemplo 1.6. *Definimos o produto interno em $\mathbb{K}^n$ da seguinte forma.*
Se $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, então

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Este é o chamado produto interno canônico em $\mathbb{K}^n$.

Observação 1.3.

- *Identificando $\mathbb{R}^n$ com o espaço das matrizes reais $n \times 1$ (matrizes colunas reais), dada uma matriz real $n \times n$ invertível A, definimos um produto interno em $\mathbb{R}^n$ por*

$$\langle x, y\rangle = y^t (A^t A) x.$$

Note que $A^t A$ é uma matriz simétrica. Quando $A = I$, obtemos o produto interno canônico em $\mathbb{R}^n$.

- *Dada uma matriz complexa A $n \times n$, definimos a sua transposta conjugada A^* por*

$$(A^*)_{ij} = \overline{A}_{ji}.$$

Observe que se A possui apenas entradas reais, sua transposta conjugada coincide com sua transposta. Identificando $\mathbb{C}^n$ com o espaço das matrizes reais n (matrizes colunas complexas), dada uma matriz complexa $n \times n$ invertível A, definimos um produto interno em $\mathbb{C}^n$ por

$$\langle x, y\rangle = y^* A^* A x.$$

Note que $A^* A$ é uma matriz hermitiana, isto é, uma matriz B que satisfaz $B^* = B$ (matrizes hermitianas reais são matrizes simétricas). Quando $A = I$, obtemos o produto interno canônico em $\mathbb{C}^n$.

Definição 1.14. *Seja V um espaço vetorial sobre $\mathbb{K}$ com produto interno $\langle\ ,\ \rangle$ e sejam os vetores $u, v \in V$. Dizemos que u e v são* **ortogonais** *se $\langle u, v\rangle = 0$. Um subconjunto $\mathcal{A}$ de V é chamado de* **ortogonal** *se os seus elementos são ortogonais dois a dois e dizemos que $\mathcal{A}$ é um conjunto* **ortonormal** *se for um conjunto ortogonal e se $\|u\| = 1$, $\forall\ u \in \mathcal{A}$. Para indicar que dois vetores u e v são ortogonais, usa-se a seguinte notação: $u \perp v$.*

Observação 1.4. *O vetor nulo 0 é ortogonal a todos os vetores de V, pois $\langle 0, u\rangle = 0$, para todo $u \in V$.*

Teorema 1.3. *Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e seja $\{v_1, \ldots, v_n\}$ uma base ortonormal de V. Então para $v \in V$, temos*

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i.$$

Demonstração. Ver [8]. □

Teorema 1.4. *Todo espaço vetorial de dimensão finita $n \geq 1$ com produto interno possui uma base ortonormal.*

Demonstração. Ver [8]. □

Corolário 1.2. *Seja V um $\mathbb{K}$ - espaço vetorial munido de um produto interno. Sejam $\beta = \{u_1, \ldots, u_n\}$ e $\beta' = \{v_1, \ldots, v_n\}$ duas bases ortonormais de V. Se M é a matriz de mudança de bases β para β', então $M \cdot \overline{M}^t = \overline{M}^t \cdot M = Id_n$.*

Demonstração. Ver [8]. □

Definição 1.15. *Seja V um espaço vetorial com produto interno, e seja $S \subseteq V$ um subconjunto de V. Chamamos de* **ortogonal** *a S ao conjunto*

$$S^{\perp} = \{v \in V; \langle v, u\rangle = 0, \quad \forall\ \ u \in S\}.$$

No caso em que S é um subespaço vetorial de V, o conjunto $S^{\perp}$ é denominado **complemento ortogonal** de S em V.

Teorema 1.5. *O conjunto $S^{\perp}$ é um subespaço de V, mesmo que S não o seja. Além disso, tem-se que $S \cap S^{\perp} = \{0\}$ no caso em que S é um subespaço próprio de V.*

Demonstração. O conjunto $S^{\perp}$ é um subespaço vetorial de V. De fato,

- $0 \in S^{\perp}$ pois $\langle 0, v\rangle = 0$ para todo $v \in S$.
- Se $v_1, v_2 \in S^{\perp}$ então $\langle v_1, u\rangle = \langle v_2, u\rangle = 0\ \forall\ u \in S$.
  Portanto, temos $\langle v_1 + v_2, u\rangle = \langle v_1, u\rangle + \langle v_2, u\rangle = 0\ \forall\ u \in S$ e então $v_1, v_2 \in S^{\perp}$.
- De modo análogo, temos que $\lambda v \in S^{\perp}$ para todo $\lambda \in \mathbb{K}$.
  Considerando agora S um subconjunto de V, vamos mostrar que $S \cap S^{\perp} = \{0\}$. Seja $w \in S \cap S^{\perp}$, isto é, $w \in S^{\perp}$ e $w \in S$. Como $w \in S^{\perp}$, temos que $\langle w, v\rangle = 0$ para todo $v \in S$. Em particular para $v = w$, pois $w \in S$, obtemos $\langle w, w\rangle = 0$. Logo, $w = 0$, o que completa a demonstração.

$\square$

Teorema 1.6. *Seja V um $\mathbb{K}$ - espaço vetorial de dimensão $n \geq 1$ e com produto interno e seja $W \subsetneq V$ um subespaço de V. Então $V = W \oplus W^{\perp}$.*

Demonstração. Ver [8]. $\square$

Definição 1.16. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. Denotamos por $L(V, W)$ o conjunto de todas as transformações lineares de V em W, isto é,*

$$L(V, W) = \{T : V \longrightarrow W \mid T \text{ é uma transformação linear } \}.$$

Definição 1.17. *Dadas as transformações lineares $T, P \in L(V, W)$. Definimos a* **adição de transformações** $T + P : V \longrightarrow W$ *da seguinte forma:*

$$(T + P)(v) = T(v) + P(v); \quad \forall \quad v \in V.$$

A aplicação assim definida é também uma transformação linear.

Definição 1.18. *Dada uma transformação linear $T \in L(V, W)$ e o escalar $\lambda \in \mathbb{K}$. Definimos a* **multiplicação de uma transformação por um escalar** $\lambda T : V \longrightarrow W$ *da seguinte forma:*

$$(\lambda T)(v) = \lambda T(v); \quad \forall \quad v \in V.$$

A aplicação assim definida é também uma transformação linear.

Teorema 1.7. *$L(V, W)$ é um espaço vetorial sobre $\mathbb{K}$ com relação as operações de adição de transformações lineares e multiplicações por escalar definidas acima.*

Demonstração. Ver [7]. □

Teorema 1.8. *Sejam V e W dois espaços vetoriais sobre $\mathbb{K}$ com dimensões n e m, respectivamente. Então o espaço $L(V, W)$ tem dimensão $m \cdot n$.*

Demonstração. Ver [8]. □

Corolário 1.3. *Sejam V e W dois espaços vetoriais de dimensões n e m sobre $\mathbb{K}$, então o espaço $L(V, W)$ é isomorfo a $\mathbb{M}_{m\times n}(\mathbb{K})$.*

Demonstração. Ver [8]. □

Definição 1.19. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Um operador linear sobre V é uma transformação linear de V em V.*

$L(V, V)$ é um espaço vetorial sobre o corpo $\mathbb{K}$ com as operações de adição de operadores e multiplicações por escalar definidas para transformações lineares.

Definição 1.20. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Denotamos por $L(V, V)$ o conjunto de todos os operadores lineares sobre V, isto é,*

$$L(V, V) = \{T : V \longrightarrow V \mid T \text{ é um operador linear } \}.$$

1.4 Adjunto

Um funcional linear é um tipo especial de transformação cujos contradomínios sejam o corpo base $\mathbb{K}$.

Definição 1.21. *Seja V um $\mathbb{K}$ - espaço vetorial. Um* **funcional linear** *sobre V é uma transformação linear $f : V \longrightarrow \mathbb{K}$.*

Exemplo 1.7. *Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e $w \in V$. Consideremos a função*

$$\begin{aligned} f_w : \ V &\longrightarrow \mathbb{K} \\ u &\longmapsto f_w(u) = \langle u, w\rangle. \end{aligned}$$

f_w é uma transformação linear:

(i) $f_w(u_1 + u_2) = \langle u_1 + u_2, w\rangle = \langle u_1, w\rangle + \langle u_2, w\rangle = f_w(u_1) + f_w(u_2)$;

(ii) $f_w(\lambda u) = \langle \lambda u, w\rangle = \lambda\langle u, w\rangle = \lambda f_w(u)$.

O conjunto $L(V, \mathbb{K})$ dos funcionais lineares forma um espaço vetorial sobre $\mathbb{K}$. Vamos denotar $L(V, \mathbb{K})$ por V^* e chamá-lo de **espaço dual** de V.

Para $u \in V$, $f_w(u) = \{u, w\}$ determina um funcional linear. O resultado a seguir garante que todos os funcionais lineares $f : V \longrightarrow \mathbb{K}$ podem ser obtidos dessa forma.

Proposição 1.4. *(Teorema da Representação de Riesz) Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e dimensão finita e $f \in V^*$ um funcional linear. Então existe um único $w \in V$ tal que $f(u) = \langle u, w\rangle$ para todo $u \in V$.*

Demonstração. Seja $\{v_1, v_2, \ldots, v_n\}$ uma base ortonormal de V. Se $u \in V$, então

$$u = \sum_{j=1}^{n} \langle u, v_j\rangle v_j$$

Logo

$$f(u) = \sum_{j=1}^{n} \langle u, v_j\rangle f(v_j) = \sum_{j=1}^{n} \langle u, \overline{f(v_j)} v_j\rangle = \langle u, \sum_{j=1}^{n} \overline{f(v_j)} v_j\rangle.$$

Tome

$$w = \sum_{j=1}^{n} \overline{f(v_j)} v_j$$

Se $w' \in V$ é outro vetor tal que $f(u) = \langle u, w'\rangle$ para todo $u \in V$, então $\langle u, w\rangle = \langle u, w'\rangle$ para todo $u \in V$, donde $\langle u, w - w'\rangle = 0$. Tomando $u = w - w'$ concluímos que $w - w' = 0$. □

Exemplo 1.8. *Seja a transformação $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida pela matriz*

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}.$$

Considere o vetor $w = (1, -2) \in \mathbb{R}^2$ e forme a expressão $\langle T(u), w\rangle$, $\forall\ u \in \mathbb{R}^2$.

Se $u = (x, y)$, então

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + 4y \\ 3x + y \end{bmatrix}.$$

Assim, $T(u) = (2x + 4y, 3x + y)$. Portanto,

$$\langle T(u), w\rangle = \langle (2x + 4y, 3x + y), (1, -2)\rangle = 2x + 4y - 6x - 2y = -4x + 2y.$$

Observe que:

$$\langle T(u_1+u_2), w\rangle = \langle T(u_1), w\rangle + \langle T(u_2), w\rangle \quad e \quad \langle T(\lambda u), v\rangle = \lambda\langle T(u), v\rangle.$$

Representemos a expressão $\langle T(u), v\rangle$ por $f_w(u)$, para indicar que depende de w (obviamente dependerá também da transformação linear escolhida).

Como f_w é um funcional linear, existirá um vetor $w^* \in \mathbb{R}^2$ tal que $f_w(u) = \langle u, w^*\rangle$.

Segue-se imediatamente do exemplo anterior que $f_w(u) = -4x + 2y$, logo $w^* = (-4, 2)$. Assim $f_w(u) = \langle T(u), w\rangle = \langle u, w^*\rangle = \langle (x,y), (-4,2)\rangle$.

Observação 1.5. *Dada uma transformação linear $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ e fixado $w \in \mathbb{R}^2$, associamos a ele um funcional linear f_w definido por $f_w(u) = \langle T(u), w\rangle$. Mas então, este funcional linear é representado por um certo vetor $v^* \in \mathbb{R}^2$. Assim, partindo de um vetor $w \in \mathbb{R}^2$, chegamos a um vetor $w^* \in \mathbb{R}^2$; note que dado $w \in \mathbb{R}^2$, o w^* que satisfaz às condições acima é único. Temos uma função de domínio $\mathbb{R}^2$ e contradomínio $\mathbb{R}^2$. Esta função será chamada de* **adjunta** *de T e representada por T^*. Em verdade, T^* é uma transformação linear.*

Teorema 1.9. *Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e de dimensão finita. Se $T \in L(V,V)$, então existe um único operador $T^* \in L(V,V)$, tal que $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ para todos $u, v \in V$.*

Demonstração. Seja $v \in V$. Queremos definir $T^* \in V$. Para tanto, vamos considerar o seguinte funcional linear

$$\begin{array}{rccl} f: & V & \longrightarrow & \mathbb{K} \\ & u & \longmapsto & f(u) = \langle T(u), v\rangle. \end{array}$$

Observe que f é linear pois se u_1, $u_2 \in V$ e $\lambda \in \mathbb{K}$, então

$$\begin{aligned} f(u_1 + \lambda u_2) &= \langle T(u_1 + \lambda u_2), v\rangle = \langle T(u_1) + \lambda T(u_2), v\rangle \\ &= \langle T(u_1), v\rangle + \lambda\langle T(u_2), v\rangle = f(u_1) + \lambda f(u_2). \end{aligned}$$

Pela Proposição 1.4 sabemos que existe um único $w \in V$ tal que $f(u) = \langle u, w\rangle$, para todo $u \in V$, isto é, tal que $\langle T(u), v\rangle = \langle u, w\rangle$, $\forall\ u \in V$. Como w é determinado de modo único por v, definimos $T^*(v) = w$. Por construção, teremos então

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle, \forall\ u, v \in V.$$

Para mostrarmos que T^* definida acima é linear, considere vetores $u, v_1, v_2 \in V$ e um escalar $\lambda \in \mathbb{K}$. Então

$$\begin{aligned}\langle u, T^*(v_1 + \lambda v_2)\rangle &= \langle T(u), v_1 + \lambda v_2\rangle = \langle T(u), v_1\rangle + \overline{\lambda}\langle T(u), v_2\rangle \\ &= \langle u, T^*(v_1)\rangle + \overline{\lambda}\langle u, T^*(v_2)\rangle = \langle u, T^*(v_1) + \lambda T^*(v_2)\rangle.\end{aligned}$$

Portanto, $\langle u, T^*(v_1 + v_2) - T^*(v_1) + \lambda T^*(v_2)\rangle = 0,\ \forall\ u \in V$, o que implica que $T^*(v_1 + \lambda v_2) = T^*(v_1) + \lambda T^*(v_2)$. Portanto, T^* é linear. A unicidade decorre facilmente da construção feita. □

Definição 1.22. *Seja $T \in L(V, V)$, onde V é um $\mathbb{K}$-espaço vetorial com produto interno. Dizemos que T possui um adjunto se existir um operador linear $T^* \in L(V, V)$ tal que $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, para todos $u, v \in V$. Diremos, nesse caso, que T^* é o* **adjunto** *de T.*

A transposta de uma matriz $A = [a_{ij}] \in \mathbb{M}_{(m\times n)}$ é a matriz $A^t = [a_{ji}] \in \mathbb{M}_{(m\times n)}$ que tem como linhas as colunas de A e como colunas as linhas de A, na mesma ordem.

Teorema 1.10. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita e sejam $T \in L(V, V)$. Considere $\beta = \{v_1, \ldots, v_n\}$ uma base ortonormal de V e $[T]_\beta$ a matriz do operador linear $T : V \longrightarrow V$ na base β, então a matriz de T^* na base β é igual à transposta conjugada da matriz de T.*

Demonstração. Por definição de uma transformação linear, temos

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \text{ para cada } j = 1, \cdots, n.$$

e

$$T(v_j) = \sum_{i=1}^{n} c_{ij} v_i,$$

onde $[T^*]_\beta = (c_{ij})_{i,j}$ é a matriz do operador T^* com relação à base β a ser determinada.

Usando-se a definição de T^* e as propriedades do produto interno segue que como, β é uma base ortonormal, temos, para cada $i, j = 1, \ldots, n$

$$c_{ij} = \langle T^*(v_j), v_i\rangle = \overline{\langle v_i, T^*(v_j)\rangle} = \overline{\langle T(v_j), v_i\rangle} = \overline{a_{ij}}.$$

Portanto, $\overline{[T]}_\beta^{\,t} = [T^*]_\beta$, como queríamos. □

Proposição 1.5. *Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e de dimensão finita. Sejam $T, S \in L(V,V)$ operadores lineares que admitem adjuntos T^* e S^*, respectivamente e $\lambda \in \mathbb{K}$. Então*

(a) *$T + S$ admite ajunto e $(T+S)^* = T^* + S^*$.*

(b) *λT admite adjunto e $(\lambda T)^* = \overline{\lambda} T^*$.*

(c) *$T \circ S$ admite adjunto e $(T \circ S)^* = S^* \circ T^*$.*

(d) *T^* admite adjunto e $(T^*)^* = T$.*

Demonstração. Ver [8]. □

Dados um espaço vetorial V de dimensão finita e um operador linear $T : V \longrightarrow V$, queremos encontrar uma base β de V, na qual a representação $[T]_\beta$ desse operador seja a mais simples possível.

Definição 1.23. *Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$, com dimensão finita n e $T : V \longrightarrow V$. O polinômio $p(t) = det(tI - T)$ é o* **polinômio característico** *de T. As raízes $\lambda_i \in \mathbb{K}$ desse polinômio são os* **autovalores** *de T. A* **multiplicidade algébrica** *de um autovalor é a multiplicidade como raiz de $p(t)$. Os elementos não nulos do núcleo $ker(\lambda_i I - T)$ são os* **autovetores** *associado ao autovalor λ_i, ou simplesmente autovetores de T. O autoespaço V_λ associado ao autovalor λ é definido por*

$$V_\lambda = ker(\lambda I - T) = \{v \in V; (\lambda I - T)v = 0\}.$$

Se existir uma base β de V tal que $[T]_\beta$ seja uma matriz diagonal, dizemos que T é **diagonalizável**.

Teorema 1.11. *Se v_i for um autovetor de $T : V \longrightarrow V$ associado ao autovalor $\lambda_i \in \mathbb{K}$, e se $\lambda_i \neq \lambda_j$ para $i \neq j$, então o conjunto $\{v_1, \ldots, v_m\}$ é linearmente independente.*

Demonstração. Faremos indução no número m de elementos do conjunto $\{v_1, \ldots, v_m\}$. Se $m = 1$, o resultado é óbvio. Suponhamos verdadeiro para $m - 1$ vetores e consideremos o caso de m vetores. Dada a combinação linear nula

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m = 0, \tag{1.2}$$

aplicando o operador T a ambos os membros desta igualdade, levando em conta que $Tv_i = \lambda_i v_i$. Resulta então

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \alpha_m \lambda_m v_m = 0. \tag{1.3}$$

Multiplicando a igualdade (1.2) por λ_n e subtraindo de (1.3), vem:

$$\alpha_1(\lambda_1 - \lambda_m)v_1 = \alpha_2(\lambda_2 - \lambda_m)v_2 = \ldots = \alpha_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

Como os autovalores são todos diferentes, as $m-1$ diferenças nos parênteses são $\neq 0$, logo, $\alpha_1 = \alpha_2 = \ldots = \alpha_{m-1} = 0$. Isso reduz a igualdade (1.2) a $\alpha_m v_m = 0$. Como $v_m \neq 0$, segue-se que $\alpha_m = 0$. Assim, a igualdade (1.2) só pode ocorrer quando todos os coeficientes α_m são nulos, o que prova o teorema. □

Corolário 1.4. *Se V for um espaço vetorial de dimensão n e se o polinômio característico do operador linear $T : V \longrightarrow V$ possuir n raízes distintas, então V possui uma base β formada por autovetores de T. A aplicação T representada na base β é uma matriz diagonal, sendo os elementos da diagonal principal os autovetores de T.*

Teorema 1.12. *Se V for um espaço vetorial de dimensão finita. Uma aplicação linear $T : V \longrightarrow V$ é diagonalizável se, e somente se, existir uma base β de V formada por autovetores de T.*

Demonstração. Suponhamos que $\beta = \{v_1, \ldots, v_n\}$ seja uma base de V tal que $[T]_\beta$ seja a matriz diagonal (não estamos supondo que os λ_i sejam distintos!.):

$$[T]_\beta = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Claramente $De_i = \lambda_i e_i$. Seja $B : V \longrightarrow \mathbb{R}^n$ o isomorfismo dado por $B(v_i) = e_i$. Então $T = B^{-1}DB$ e

$$T(v_i) = B^{-1}D(Bv_i) = B^{-1}(\lambda_i e_i) = \lambda_i B^{-1} e_i = \lambda_i v_i$$

mostrando que cada v_i é autovalor de T.

Seja $\beta = \{v_1, \ldots, v_n\}$ uma base de V formadas por autovetores de T. então, como

$$\begin{aligned} T(v_1) &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + 0 v_n \\ T(v_2) &= 0 v_1 + \lambda_2 v_2 + \cdots + 0 v_n \\ &\vdots \\ T(v_n) &= 0 v_1 + 0 v_2 + \cdots + \lambda_n v_n, \end{aligned}$$

a matriz $[T]_\beta$ será uma matriz diagonal onde os elementos da diagonal principal são os autovalores λ_i, isto é,

$$[T]_\beta = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

$\square$

Definição 1.24. *Seja $T : V \longrightarrow V$ um operador linear onde V é um $\mathbb{K}$ - espaço vetorial e seja $W \subseteq V$ um subespaço vetorial de V. Dizemos que W é um* **subespaço T - invariante** *de V se $T(w) \in W$ para todo $w \in W$.*

Observação 1.6. *Seja $T : V \longrightarrow V$ um operador linear onde V é um $\mathbb{K}$- espaço vetorial.*

(a) *Os subespaços $ker(T)$ e $Im(T)$ são T - invariantes.*

(b) *Se λ for um autovalor de T, então V_λ é um subespaço T - invariante de V. De fato, se $v \in V_\lambda$, então $T(v) = \lambda v \in V_\lambda$.*

(c) *Se W é um subespaço T - invariante, então a restrição de T a W é um operador linear em $L(W, W)$.*

Capítulo 2

Teorema Espectral e Aplicação

Uma importante classe de operadores lineares é formada pelos operadores que coincidem com os respectivos adjuntos. Estudar tais operadores é o objetivo deste capítulo. As referências para os estudos realizados neste trabalho são [8], [10] e [11].

Definição 2.1. *Seja $T \in L(V,V)$, onde V é um $\mathbb{K}$ - espaço vetorial com produto interno. Dizemos que T é* **autoadjunta** *se T admite adjunta T^* e $T^* = T$.*

Observação 2.1. *No caso em que $\mathbb{K} = \mathbb{C}$, usamos também o termo* **hermitiano** *e no caso em que $\mathbb{K} = \mathbb{R}$, usamos também o termo* **simétrico**.

Proposição 2.1. *Sejam V um $\mathbb{K}$ - espaço vetorial com produto interno e de dimensão finita e $T \in L(V,V)$. As seguintes afirmações são equivalentes:*

(*a*) *T é autoadjunto.*

(*b*) *$\overline{[T]}^t_\beta = [T]_\beta$ para toda base ortonormal β de V.*

(*c*) *Existe uma base ortonormal β de V tal que $\overline{[T]}^t_\beta = [T]_\beta$.*

Demonstração. Seja β uma base ortonormal de V. Vimos no Teorema (1.10) temos que $[T^*]_\beta = \overline{[T]}^t_\beta$. Suponha T autoadjunto, segue que $\overline{[T]}^t_\beta = [T]_\beta$, o que prova a implicação $(a) \Rightarrow (b)$. Por outro lado, se supormos que $\overline{[T]}^t_\beta = [T]_\beta$ então $[T^*]_\beta = [T]_\beta$ e, portanto, T é autoadjunto, o que prova $(c) \Rightarrow (a)$. A implicação $(b) \Rightarrow (c)$ é trivial.

□

Corolário 2.1. *Seja V um $\mathbb{K}$ - espaço vetorial de dimensão finita e β uma base ortonormal de V. Se $T \in L(V,V)$ for um operador linear autoadjunto e se $[T]_\beta = (a_{ij})_{i,j}$ então $a_{ij} = \overline{a_{ji}}$. Em particular, os elementos da diagonal de $[T]_\beta$ são números reais.*

Definição 2.2. *Sejam V um $\mathbb{K}$ - espaço vetorial com produto interno e $T \in L(V,V)$. Dizemos que T é* **unitário** *se for um isomorfismo de espaços com produto interno.*

Proposição 2.2. *Seja $T \in L(V,V)$, onde V é um $\mathbb{K}$-espaço vetorial com produto interno. Então T é unitário se e somente se o adjunto T^* existir e $T \circ T^* = T^* \circ T = Id$.*

Demonstração. Suponha, que T seja unitário. Então T é inversível e, como preserva produto interno, temos

$$\langle T(u), v\rangle = \langle T(u), T \circ T^{-1}(v)\rangle = \langle u, T^{-1}(v)\rangle$$

para todos $u, v \in V$. Logo, $T^{-1} = T^*$. Donde $T^* \circ T = T \circ T^* = Id$.

Reciprocamente, suponha que T^* exista e que $T^* \circ T = T \circ T^* = Id$. Então T é invertível e $T^{-1} = T^*$. Resta mostrar que T preserva produtos internos. De fato,

$$\langle T(u), T(v)\rangle = \langle u, (T^* \circ T)(v)\rangle = \langle u, Id(v)\rangle = \langle u, v\rangle$$

para todos $u, v \in V$. Portanto, T é unitário. □

Exemplo 2.1. *Seja $V = \mathbb{M}_{n\times 1}(\mathbb{C})$ com produto interno $\langle M, N\rangle = \overline{N}^t M$ onde $M, N \in V$ e seja $A \in \mathbb{M}_{n\times n}(\mathbb{C})$. Defina $T : V \longrightarrow V$ por $T(X) = AX$. Calculando-se $\langle T(X), T(Y)\rangle$ teremos*

$$\langle T(X), T(Y)\rangle = \langle AX, AY\rangle = (\overline{AY})^t(AX) = \overline{Y}^t\overline{A}^t AX$$

para todos $X, Y \in V$. Conclui-se daí que T é unitário se e somente se $\overline{A}^t A = Id_n$.

Definição 2.3. *Seja $A \in \mathbb{M}_n(\mathbb{K})$. Dizemos que A é* **unitária** *se $\overline{A}^t A = A\overline{A}^t = Id_n$.*

Quando $\mathbb{K} = \mathbb{R}$, também dizemos que A é **ortogonal**.

O objetivo principal desta seção é resolver o seguinte problema: em que condições um operador linear possui uma base ortonormal em relação à qual a sua matriz é diagonal.

Definição 2.4. *Sejam V um espaço vetorial com produto interno e $T \in L(V,V)$. Dizemos que T é* **normal** *se existir T^* e $T \circ T^* = T^* \circ T$.*

Observação 2.2. *Todo operador autoadjunto é normal.*

Proposição 2.3. *Sejam V um $\mathbb{K}$ - espaço vetorial com produto interno e $T \in L(V,V)$ um operador normal. Então*

(a) $\|T(v)\| = \|T^*(v)\|, \quad \forall \quad v \in V.$

(b) *Se $T(v) = \alpha v$ para $\alpha \in \mathbb{K}$ e $v \in V$, então $T^*(v) = \overline{\alpha} \cdot v$.*

(c) *Se $T(v_1) = \alpha_1 v_1$ e $T(v_2) = \alpha_2 v_2$, para $v_1, v_2 \in V$ e $\alpha_1, \alpha_2 \in \mathbb{K}$, com $\alpha_1 \neq \alpha_2$, então $\langle v_1, v_2 \rangle = 0$.*

Demonstração. (a) Seja $v \in V$, então

$$\langle T(v), T(v) \rangle = \langle v, T^*(T(v)) \rangle = \langle v, T(T^*(v)) \rangle = \overline{\langle T(T^*(v)), v \rangle}.$$

Como $\langle T(v), T(v) \rangle$ é um número real, segue que

$$\overline{\langle T(T^*(v)), v \rangle} = \langle T(T^*(v)), v \rangle = \langle T^*(v), T^*(v) \rangle.$$

Portanto, $\|T(v)\| = \|T^*(v)\|$, como queríamos.

(b) Se $T(v) = \alpha v$, então $(T - \alpha Id)(v) = 0$. Logo, $\|(T - \alpha Id)(v)\| = 0$. Usando o item (a), concluímos que $\|(T - \alpha Id)^*(v)\| = 0$. Então $(T - \alpha Id)^*(v) = 0$ e, portanto, $T^*(v) = \overline{\alpha} v$.

(c) Temos que $T(v_1) = \alpha_1 v_1$ e $T(v_2) = \alpha_2 v_2$, assim,

$$\alpha_1 \langle v_1, v_2 \rangle = \langle \alpha_1 v_1, v_2 \rangle = \langle T(v_1), v_2 \rangle = \langle v_1, T^*(v_2) \rangle = \langle v_1, \overline{\alpha}_2 (v_2) \rangle = \alpha_2 \langle v_1, v_2 \rangle.$$

Daí, $\alpha_1 \langle v_1, v_2 \rangle = \alpha_2 \langle v_1, v_2 \rangle$ e, portanto, $(\alpha_1 - \alpha_2)\langle v_1, v_2 \rangle = 0$. Como $\alpha_1 \neq \alpha_2$, segue que $\langle v_1, v_2 \rangle = 0$. □

Notemos que um operador linear real pode não ter autovalores reais (por exemplo, uma rotação em $\mathbb{R}^2$). No entanto, é possível provar que todo operador autoadjunto tem pelo menos um autovalor que é real.

Proposição 2.4. *Sejam V um $\mathbb{K}$ - espaço vetorial de dimensão finita com produto interno. Se $T \in L(V,V)$ é autoadjunta, então todo autovalor de T é real.*

Demonstração. Se λ é um autovalor de T existe $v \in V$, $v \neq 0$ tal que $T(v) = \lambda v$. Por outro lado,

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle T(v), v \rangle = \langle v, T^*(v) \rangle = \langle v, T(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle,$$

como $v \neq 0$, $\lambda \langle v, v \rangle = \overline{\lambda} \langle v, v \rangle$ acarreta que $\lambda = \overline{\lambda}$, ou seja, λ é real. □

Teorema 2.1. *Seja V um $\mathbb{K}$ - espaço vetorial de dimensão finita com produto interno. Se $T \in L(V,V)$ é autoadjunta, então T possui um autovetor.*

Demonstração. Se $\mathbb{K} = \mathbb{C}$ então $p_T(x)$ tem raízes pela proposição 2.4 seus autovalores são reais.

Se $\mathbb{K} = \mathbb{R}$. Suponha $dim_{\mathbb{K}} V = n$ e seja $T \in L(V,V)$ um operador autoadjunto. Sejam β uma base ortonormal de V e $A = [T]_\beta$. Como T é um operador autoadjunto, isto é, que satisfaz $T = T^*$, a matriz A é uma matriz hermitiana (isto é, simétrica, pois é uma matriz real) temos pelo Teorema (1.10) que $A = \overline{A}^t$.

Se considerada como uma matriz de um espaço com produto interno complexo, seu polinômio característico possui uma raiz complexa λ, logo $det(\lambda I - A) = 0$ e a matriz $\lambda I - A$ é singular e possui núcleo não trivial; por exemplo, seja $W = \mathbb{M}_{n\times 1}(\mathbb{C})$ e considere o produto interno $\langle X, Y\rangle = \overline{Y}^t X$. Defina o operador linear

$$\begin{array}{rccc} F: & W & \longrightarrow & W \\ & X & \longmapsto & F(X) = AX \end{array}$$

para todo $X \in W$ e $n = dimV$. Mas como A é hermitiana, sabemos que $F^*(X) = \overline{A}^t X = AX$ e, portanto, F é autoadjunto e segue pela Proposição (2.4), que λ é real. Como $\lambda I - A$ é uma matriz real cujo o determinante é nulo, ela possui núcleo real não trivial, portanto λ é um autovalor de A. $\square$

Lema 2.1. *Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e de dimensão finita e $T \in L(V,V)$. Se W é um subespaço T - invariante de V, então $W^\perp$ é T^* - invariante.*

Demonstração. Sejam $v \in W$ e $w \in W^\perp$. Como W é T - invariante, então $T(v) \in W$ e, portanto, $\langle T(v), w\rangle = 0$. Logo,

$$\langle v, T^*(w)\rangle = \langle T(v), w\rangle = 0.$$

Portanto, $T^*(w) \in W^\perp$, para cada $w \in W^\perp$. $\square$

Teorema 2.2. *(***Teorema Espectral***) Seja V um $\mathbb{K}$ - espaço vetorial com produto interno e de dimensão finita. Se $T \in L(V,V)$ é autoadjunto, então existe uma base ortonormal de V cujos vetores são autovetores de T.*

Demonstração. Suponha que $dim_{\mathbb{K}} V = n \geq 1$. Pelo Teorema (2.1), T possui um autovetor v_1. Se $dim_{\mathbb{K}} V = 1$, então $\{\frac{v_1}{\|v_1\|}\}$ é uma base, como queríamos. Vamos supor agora que $n > 1$ e que o resultado vale para todo espaço vetorial de dimensão $n-1$. Seja $W = [v_1]$, onde v_1 é o autovetor acima. W é invariante por T. Pelo Lema (2.1) $W^\perp$ é T^* - invariante. Como $T = T^*$ segue que $W^\perp$ é T - invariante. Agora, como $W^\perp$ é um espaço de dimensão $n-1$ segue da hipótese de indução que $W^\perp$ possui uma base ortonormal $\{v_2, \ldots, v_n\}$ formada por autovetores da restrição $T : W^\perp \longrightarrow W^\perp$. Logo $\beta = \{\frac{v_1}{\|v_1\|}, v_2, \ldots, v_n\}$ é um conjunto ortonormal com $dim_{\mathbb{K}}V$ elementos e, portanto, uma base de V. Por construção, todos os elementos de β são autovetores e o resultado está provado. $\square$

Observação 2.3. *Vale a recíproca do Teorema Espectral: se existe uma base ortonormal $\{u_1, \ldots, u_n\} \subset V$ formada por autovetores do operador $T : V \longrightarrow V$ então este operador é autoadjunto. Com efeito, para quaisquer $i, j = 1, \ldots, n$ tem-se*

$$\langle T(u_i), u_j\rangle = \langle \lambda_i u_i, u_j\rangle = \lambda_i\delta_{ij} = \lambda_j\delta_{ij} = \langle u_i, \lambda_j u_j\rangle = \langle u_i, T(u_j)\rangle$$

e daí resulta que $\langle T(u), v\rangle = \langle u, T(v)\rangle$ para quaisquer $u, v \in V$.

Corolário 2.2. *Seja $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz simétrica. Então existe uma matriz invertível $M \in \mathbb{M}_n(\mathbb{R})$ tal que M^tAM é diagonal.*

Demonstração. Seja $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ um operador linear tal que $[T]_{can} = A$. Como A é simétrica segue que T é autoadjunta. Consequentemente, pelo Teorema 2.2, temos que existe uma base β ortonormal de $\mathbb{R}^n$ formada de autovetores de T. Considere agora M a matriz mudança de base canônica de $\mathbb{R}^n$ para a base β. Por (2.1) temos que $[T]_\beta = M^{-1}[T]_{can}M$. Como as bases em questão são ortonormais, segue do corolário (1.2) que $M^{-1} = M^t$. Portanto, $[T]_\beta = M^tAM$ satisfaz a afirmação desejada. $\square$

Exemplo 2.2. *Considere $\mathbb{R}^2$ com produto interno usual e $T(x, y) = (x + 3y, 3x - 7y)$. Seja β a base canônica do $\mathbb{R}^2$, chame*

$$A = [T]_\beta = \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix}$$

Logo,

$$det(A - \lambda I_2) = 0 \Leftrightarrow det\begin{bmatrix} 1-\lambda & 3 \\ 3 & -7-\lambda \end{bmatrix} = 0 \Leftrightarrow \lambda^2 + 6\lambda - 16 = (\lambda - 2)(\lambda + 8) = 0.$$

Como a matriz A é simétrica, segue que T é autoadjunto e, portanto, pelo Teorema Espectral este operador admite uma base de autovetores ortonormal.

Vamos determiná-la.

(i) *Para $\lambda = 2$ temos:*

$$(A - 2I_2)v = 0 \Leftrightarrow \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow -x + 3y = 0.$$

Portanto,

$$W_1 = \{(3y, y) \mid y \in \mathbb{R}\} = \{y(3,1) \mid y \in \mathbb{R}\}$$

(ii) *Ao invés de seguir o procedimento padrão, vamos usar a informação que o teorema Espectral nos fornece. Sabemos que os autovetores associado ao autovalor $\lambda = -8$ são ortogonais à $(3,1)$, e, como estamos em $\mathbb{R}^2$, um autovetor associado a $\lambda = -8$ deve ser $(-1,3)$. De fato, veja que*

$$\begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ -24 \end{bmatrix} = -8 \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Normalizando esses vetores, obtemos uma base ortonormal

$$\beta = \{(3/\sqrt{10}, 1/\sqrt{10}), (-1/\sqrt{10}, 3/\sqrt{10})\}.$$

Portanto a matriz

$$P = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}$$

diagonaliza A. Temos que P é ortogonal, isto é, $P^T = P^{-1}$, com base nisso, teremos

$$D = \begin{bmatrix} 2 & 0 \\ 0 & -8 \end{bmatrix} = P^T A P.$$

2.1 Formas Bilineares Simétricas e Formas Quadráticas

Agora consideremos formas bilineares simétricas e formas quadráticas, que são generalizações do produto escalar e do comprimento.

Definição 2.5. *Suponha V um espaço vetorial sobre $\mathbb{R}$. Dizemos que a aplicação $B : V \times V \longrightarrow \mathbb{R}$ será* **bilinear** *se ela satisfazer as seguintes propriedades para $u, v, w \in V$ e $\alpha \in \mathbb{R}$ qualquer*

$$B(u + \alpha w, v) = B(u, v) + \alpha B(w, v) \quad e \quad B(u, v + \alpha w) = B(u, v) + \alpha B(u, w)$$

Além disso, diremos que ela é simétrica se

$$B(u, v) = B(v, u)$$

para quaisquer $u, v \in V$.

Como exemplos de formas bilineares simétrica, podemos citar os produtos internos sobre $\mathbb{R}^n$.

Às formas bilineares simétricas, podemos associar, para cada base, uma matriz simétrica A da seguinte forma: fixe uma base $\alpha = \{v_1, v_2, \ldots, v_n\}$ de V, tome u e v vetores de V. Calculando as suas coordenadas, temos $u = \sum_{i=1}^{n} x_i v_i$, $v = \sum_{j=1}^{n} y_j v_j$ então,

$$\begin{aligned} B(u, v) &= B\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) \\ &= \sum_{i=1}^{n} x_i B\left(v_i, \sum_{j=1}^{n} y_j v_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j B(v_i, v_j). \end{aligned}$$

Portanto, para calcular $B(u, v)$, basta saber os valores $B(v_i, v_j)$, $1 \leq i, j \leq n$. Se criarmos a matriz (simétrica) $A = [B]_\alpha = [B(v_i, v_j)]$, isto é, a matriz que na posição ij tem o valor $B(v_i, v_j)$, então,

$$B(u, v) = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} B(v_1, v_1) & B(v_1, v_2) & \cdots & B(v_1, v_n) \\ B(v_2, v_1) & B(v_2, v_2) & \cdots & B(v_2, v_n) \\ \vdots & \vdots & \ddots & \vdots \\ B(v_n, v_1) & B(v_n, v_2) & \cdots & B(v_n, v_n) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Chamamos a matriz $A = [B]_\alpha = [B(v_i, v_j)]$ de **matriz associada à forma bilinear simétrica na base** α.

Exemplo 2.3. *Considere a forma bilinear B de $\mathbb{R}^3$ dada por*

$$B(v,w) = 2x_1y_1 - 4x_2y_1 - 8x_3y_1 - 4x_1y_2 + x_2y_2 + 7x_3y_2 - 8x_1y_3 + 7x_2y_3 + 5x_3y_3.$$

Seja $\beta = \{e_1, e_2, e_3\}$ a base canônica de $\mathbb{R}^3$. Calculando $B(e_1, e_1) = 2$, $B(e_1, e_2) = -4$ e todas as outras possibilidades, obteremos a matriz

$$\begin{bmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{bmatrix}$$

Definição 2.6. *Seja V um espaço vetorial real. Dizemos que uma aplicação $Q : V \longrightarrow \mathbb{R}$* **será uma forma quadrática** *se $Q(u) = B(u,u)$ para alguma forma bilinear simétrica B de V.*

Há uma maneira de, considerando uma forma quadrática Q, obtermos a forma bilinear B necessária para a definição 2.6 Considere a aplicação $B : V \times V \longrightarrow V$ dada por

$$B(u,v) = \tfrac{1}{2}[Q(u+v) - Q(u) - Q(v)] \text{ (Fórmula de Polarização).}$$

Portanto, dado uma forma quadrática $Q : V \longrightarrow \mathbb{R}$, podemos determinar uma forma bilinear simétrica B, através da fórmula de Polarização. Além disso, se fixarmos uma base α de V podemos associar uma matriz $A = [B]_\alpha$ à forma bilinear B. Por isso, também, chamamos a matriz A de matriz da forma quadrática Q com respeito à base α.

É possível fazer o caminho inverso. Se tivermos uma matriz $A = [a_{ij}]_{n\times n}$, simétrica, poderemos determinar uma forma quadrática fazendo $Q(u) = u^tAu$.

Exemplo 2.4. *Uma forma quadrática mais geral pode ser definida em $\mathbb{R}^n$ partindo de uma matriz A simétrica n por n. Então podemos definir uma forma bilinear simétrica $B_A : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ por*

$$B_A(x,y) = x^tAy$$

Nesta definição x e y são olhados como vetores coluna. A bilinearidade seguirá da linearidade da multiplicação de matrizes. A simetria resultará de A ser simétrica:

$$B_A(y,x) = y^tAx = y^tA^tx = (x^tAy)^t = x^tAy = B_A(x,y).$$

Observação 2.4. *Seja B uma forma bilinear representada pela matriz simétrica $A = [a_{ij}]$. Considere $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. Então a forma quadrática Q pode ser representada por*

$$\begin{aligned} Q(u) = B(u,u) &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \sum_{i,j} a_{ij}x_ix_j = \sum_i a_{ii}x_i^2 + \sum_{i<j} a_{ij}x_ix_j. \end{aligned}$$

Segue que Q será uma forma quadrática se, e somente se, $Q(u)$ for um polinômio homogêneo de grau 2 as entradas do vetor u.

Pela observação 2.4, dada a aplicação $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$, definida por

$$Q(x,y) = ax^2 + bxy + cy^2, \text{ onde } a, b, c \in \mathbb{R},$$

temos que esta é uma forma quadrática, pois consiste de um polinômio de grau 2. Além disso, é fácil verificar que

$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Essa observação se aplica a qualquer $n \geq 3$.

Para o caso $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$, onde

$$Q(x,y,z) = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz$$

a forma matricial é a seguinte

$$Q(x,y,z) = \begin{bmatrix} x & y & z \end{bmatrix} \cdot \begin{bmatrix} A & D/2 & E/2 \\ D/2 & B & F/2 \\ E/2 & F/2 & C \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Mudança de Coordenada e Formas Bilineares Simétricas

Suponha que x, y são as coordenadas de dois vetores do V em relação à base $\alpha = \{u_1, \ldots, u_n\}$. Dizemos que x' e y' representam os mesmos vetores em relação à outra base $\beta = \{v_1, \ldots, v_n\}$. Dessa forma, existe uma matriz invertível $P = [I]_\alpha^\beta$, que satisfaz

$$x = Px' \text{ e, de outra forma } y = Py'$$

Se $B : V \times V \longrightarrow \mathbb{R}$ é uma forma bilinear simétrica e $A = [f(u_i, u_j)]$, então:

$$B(x, y) = x^t Ay = (Px')^t APy' = x'^t P^t APy' = B(x', y').$$

Segue que a matriz de B, com respeito a base β, isto é, $B = [f(v_i, v_j)]$, deve satisfazer $B = P^t AP$.

Definição 2.7. *Diremos que uma matriz B será congruente a uma matriz A (e escrevemos $B \cong A$), se existir uma matriz invertível P tal que $B = P^t AP$.*

Nota 2.1. *Não confunda o conceito de congruência entre matrizes com o de semelhança entre matrizes. Apesar disso, se a matriz P for ortogonal, os dois conceitos irão coincidir.*

De acordo com a definição acima, podemos dizer que matrizes associadas a uma mesma forma quadrática, em relação a bases diferentes, serão sempre congruentes, e, reciprocamente, matrizes congruentes estarão associadas a uma mesma forma quadrática, só que relacionadas a outro sistema de coordenadas.

Teorema 2.3. *(Teorema dos Eixos Principais) Seja V um espaço vetorial real com produto interno $\langle , \rangle$ e Q uma forma quadrática sobre V. Então existe uma base ortonormal $v_1, \ldots, v_n$ de V (com relação a $\langle , \rangle$) tal que a matriz A para Q é diagonal.*

Teorema 2.4. *Se $Q(v) = B(v, v)$ uma forma quadrática em V. Existe uma base ortonormal β de V tal que se*

$$[v]_\beta = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

então $Q(v) = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$.

Demonstração. Seja α uma base ortonormal qualquer de V.
Então, $Q(v) = B(v,v) = [v]_\alpha^t\ [B]_\alpha^\alpha\ [v]_\alpha$. Logo, a matriz $[B]_\alpha^\alpha$ é uma matriz simétrica e, portanto, corresponde a um operador autoadjunto $T : V \to V$ que tem como matriz $[T]_\alpha^\alpha = [B]_\alpha^\alpha$. Como um operador autoadjunto pode ser diagonalizado mediante uma base β de autovetores ortonormais, então

$$
\begin{aligned}
[B]_\alpha^\alpha = [T]_\alpha^\alpha &= [I]_\alpha^\beta \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [I]_\beta^\alpha \\
&= ([I]_\beta^\alpha)^{-1} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [I]_\beta^\alpha \\
&= ([I]_\beta^\alpha)^t \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [I]_\beta^\alpha
\end{aligned}
$$

pois α e β são bases ortonormais e, portanto, $[I]_\beta^\alpha$ é uma matriz ortogonal. Então,

$$
\begin{aligned}
Q(v) &= [v]_\alpha^t \cdot ([I]_\beta^\alpha)^t \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [I]_\beta^\alpha \cdot [v]_\alpha \\
&= ([I]_\beta^\alpha [v]_\alpha)^t \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} ([I]_\beta^\alpha [v]_\alpha) \\
&= [v]_\beta^t \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [v]_\beta \\
&= [y_1 \cdots y_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\
&= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.
\end{aligned}
$$

□

2.2 Máximos e Mínimos de Funções de duas Variáveis

Seja $z = f(x, y)$ função diferenciável de duas variáveis e queremos achar os máximos e mínimos de f. Em particular estamos interessados nos máximos e mínimos relativos.

Definição 2.8. *Seja $f(x, y)$ uma função a valores reais e seja $(x_0, y_0) \in A$, com $A \subset D_f$.*

(i) Dizemos que $(x_0, y_0) \in D_f$ é ponto de **máximo relativo** *de f se existir uma bola aberta B de centro (x_0, y_0) tal que*

$$f(x, y) \leq f(x_0, y_0),$$

para todo $(x, y) \in B \cap D_f$.

(ii) Dizemos que $(x_0, y_0) \in D_f$ é ponto de **mínimo relativo** *de f se existir uma bola aberta B de centro (x_0, y_0) tal que*

$$f(x, y) \geq f(x_0, y_0),$$

para todo $(x, y) \in B \cap D_f$.

Definição 2.9. *Um número que é máximo relativo ou mínimo relativo de f chama-se em extremo de f.*

Definição 2.10. *Seja $z = f(x, y)$ uma função que admite derivadas parciais em (x_0, y_0). O vetor*

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

denomina-se **gradiente** *de f em (x_0, y_0).*

O gradiente $\nabla f(x_0, y_0)$ é um vetor aplicado no ponto (x_0, y_0).

Teorema 2.5. *Seja $z = f(x, y)$ uma função com domínio D, possuindo máximo e mínimo relativo num ponto (x_0, y_0) interior a D. se f for diferenciável nesse ponto, então as derivadas de primeira ordem se anulam.*

Demonstração. Ver [10]. □

Definição 2.11. *Se f é diferenciável em (x_0, y_0) e se $\nabla f(x_0, y_0) = 0$, o ponto (x_0, y_0) diz-se um* **ponto crítico** *ou* **estacionário** *de f.*

Observação 2.5. *Os pontos extremantes de f estão entre os seus pontos críticos. No entanto, um ponto crítico nem sempre é um ponto extremante.*

Um ponto crítico que não é um ponto extremante é chamado um ponto de sela.

Definição 2.12. *Uma função diferenciável $f(x, y)$ tem um* **ponto de sela** *em um ponto crítico (x_0, y_0) se toda bola aberta centrada em (x_0, y_0) contém pontos (x, y) onde $f(x, y) < f(x_0, y_0)$ e outros pontos para os quais $f(x, y) > f(x_0, y_0)$.*

Critério das derivadas de segunda ordem para extremos de funções de duas variáveis

No caso em que $n = 2$ a natureza de um ponto de crítico pode determinar-se pela a derivada de segunda ordem $f_{xx}(x_0, y_0)$ e o determinante da matriz Hessiana. O próximo teorema, que é provado em cálculo, torna possível analisar os pontos críticos usando somente derivadas.

Teorema 2.6. *Seja $z = f(x, y)$ uma função com derivadas parciais de segunda ordem contínuas em alguma região circular centrada num ponto crítico (x_0, y_0) de f. Então:*

(a) *f tem um mínimo relativo em (x_0, y_0) se $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0$ e $f_{xx}(x_0, y_0) > 0$.*

(b) *f tem um máximo relativo em (x_0, y_0) se $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0$ e $f_{xx}(x_0, y_0) < 0$.*

(c) *f tem um ponto de sela em (x_0, y_0) se $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) < 0$.*

(d) *O teste é inconclusivo se $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) = 0$.*

Exemplo 2.5. *Para a função $f(x, y) = 2x^2 + 6xy + y^2$, existe um valor crítico na origem. O determinante é dado por*

$$H(x, y) = \begin{vmatrix} 4 & 6 \\ 6 & 2 \end{vmatrix} = -28.$$

Logo há um ponto de sela na origem.

Exemplo 2.6. *A função* $f(x,y) = x$ *sen* $x-$ *cos* $y - xy$, *tem um ponto crítico na origem, pois* $f_1 = x$ *cos* $x+$ *sen* $x - y$ *e* $f_2 =$ *sen* $y - x$ *se anulam ambas aí. Quando calculamos a matriz obtemos* $f_{11} = -x$ *sen* $x + 2$ *cos* $x - y$, $f_{12} = -1$, $f_{22} =$ *cos* y. *Calculando na origem obtemos a matriz*

$$\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Como esta tem determinante positivo e $f_{11} > 0$ *há um mínimo local na origem.*

Nosso interesse aqui é mostrar como esse teorema pode ser expresso em termos de formas quadráticas.

Definição 2.13. *Seja* $f(x,y)$ *de classe* C^2. *A função* H *dada por*

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial y \partial x}(x,y) \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix}$$

denomina-se a matriz **hessiana**, *ou simplesmente, a* **hessiana** *de* f *no ponto em questão, em homenagem ao matemático e cientista alemão Ludwig Otto Hesse (1811 - 1874). A hessiana é de interesse porque*

$$det[H(x_0,y_0)] = \begin{vmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{xy}(x_0,y_0) & f_{yy}(x_0,y_0) \end{vmatrix} = f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - f_{xy}^2(x_0,y_0)$$

Considere a forma quadrática $Q(x,y)$ cuja a matriz é dada por $H(x,y)$. Com a hipótese de serem contínuas as derivadas segundas, sabe-se $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, assim essa matriz será simétrica. Naturalmente, ela varia com o ponto, mas variará continuamente. Disto pode-se mostrar que seus autovalores também variarão continuamente. A regra da cadeia pode ser combinada com o Teorema de Taylor para mostrar que para (x,y) próximo de (x_0,y_0), se pusermos $h = (x,y) - (x_0,y_0)$ então

$$f(x,y) = f(x_0,y_0) = \frac{1}{2}h^t H((x_0,y_0) + th)h$$

para algum t, $0 \leq t \leq 1$. Assim a questão de máximos locais, mínimos locais e pontos de sela se reduz à questão de saber se isto é sempre negativo ou positivo ou pode tomar valores positivos quanto negativos. Isto depende dos sinais dos autovalores de $H((x_0,y_0)+th)$. Mas se h é suficientemente pequeno e 0 não é autovalor de $H(x_0,y_0)$ então os autovalores de $H((x_0,y_0)+th)$ terão os mesmos sinais que os autovalores de $H(x_0,y_0)$. Assim, obtemos o critério.

Proposição 2.5. Critério para extremos locais *Suponha que as derivadas parciais primeira de f se anulem em (x_0, y_0) e que as derivadas parciais segundas de f são contínuas. Seja H a matriz hessiana das derivadas parciais de segunda ordem.*

(*i*) *Se $H(x_0, y_0)$ tem dois autovalores positivos então f tem um mínimo local em (x_0, y_0).*

(*ii*) *Se $H(x_0, y_0)$ tem dois autovalores negativos então f tem um máximo local em (x_0, y_0).*

(*iii*) *Se $H(x_0, y_0)$ tem um autovalor positivo e um negativo então f tem um ponto de sela em (x_0, y_0).*

Definição 2.14. *Seja A uma matriz simétrica real $n \times n$ e considere $Q(h) = h^T \cdot A \cdot h$ a forma Quadrática associada, tem-se então:*

(*a*) *$Q(h) > 0$ para todo $h \neq 0$ se e só se todos os autovalores de A são todos positivos. A forma quadrática diz-se* **positiva definida**.

(*b*) *$Q(h) < 0$ para todo $h \neq 0$ se e só se todos os autovalores de A são todos negativos. A forma quadrática diz-se* **negativa definida**.

(*c*) *A matriz A (ou a forma quadrática Q) é* **indefinida** *se existe* **h** *e* **k** *em $\mathbb{R}^n$ tais que*

$$Q(h) = h^T \cdot A \cdot h > 0 \text{ e } Q(k) = k^T \cdot A \cdot k < 0.$$

Nota 2.2. *Se nenhum dos autovalores é zero dizemos que a forma quadrática é* **não degenerada**.

Teorema 2.7. *Sejam $f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ uma função de classe C^2, $(x_0, y_0) \in U$ um ponto crítico de f e H a forma quadrática Hessiana de f no ponto (x_0, y_0). Então*

(*a*) *Se H é positiva definida, (x_0, y_0) é um ponto de mínimo local não degenerado;*

(*b*) *Se H é negativa definida, (x_0, y_0) é um ponto de máximo local não degenerado;*

(*c*) *Se H é indefinida, (x_0, y_0) é um ponto de sela f;*

(d) *O teste é inconclusivo nos demais casos.*

Demonstração. Ver [13]. □

Existe um critério simples para saber se a forma é não degenerada, para uma forma quadrática dada uma matriz, pois o determinante da matriz dá o produto dos autovalores. Logo a forma não é degenerada se o determinante da matriz associada não é zero.

No caso 2 por 2 terá autovalores tanto positivo quanto negativo se o determinante for negativo, mas os dois autovalores serão de mesmo sinal se o determinante for positivo.

Exemplo 2.7. *A matriz*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

é positiva definida, pois a forma quadrática associada,

$$Q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + x_2^2,$$

é maior que 0 *para todo* $(x_1, x_2) \neq (0, 0)$.

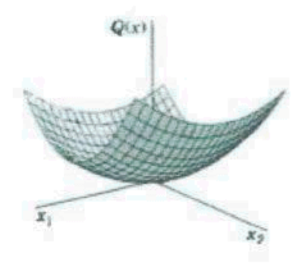

Figura 2.1: Gráfico da forma quadrática positiva definida.

Exemplo 2.8. *A matriz*

$$B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

é negativa definida, pois a forma quadrática associada,

$$Q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -x_1^2 - x_2^2,$$

é menor que 0 *para todo* $(x_1, x_2) \neq (0, 0)$.

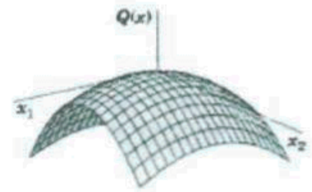

Figura 2.2: Gráfico da forma quadrática negativa definida.

Exemplo 2.9. *A matriz*

$$C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

é indefinida, pois a forma quadrática associada,

$$Q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - x_2^2,$$

é tal que $Q(1,0) = 1 > 0$ *e* $Q(0,1) = -1 < 0$.

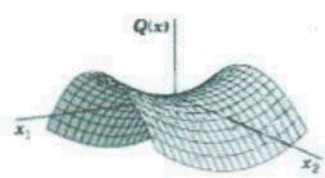

Figura 2.3: Gráfico da forma quadrática positiva indefinida.

Exemplo 2.10. *Classificar os pontos críticos da função*

$$f(x, y) = 3xy^2 + x^3 - 3x$$

O domínio de f é o $\mathbb{R}^2$. O únicos candidatos a extremantes locais são os pontos críticos, pois o domínio de f ($D_f = \mathbb{R}^2$).

- Pontos críticos

$$\frac{\partial f}{\partial x} = 3y^2 + 3x^2 - 3 \qquad e \qquad \frac{\partial f}{\partial y} = 6xy$$

resulta que os candidatos a extremantes locais são as soluções do sistema

$$\begin{cases} 3y^2 + 3x^2 - 3 = 0 \\ 6xy = 0 \end{cases}$$

As soluções do sistema são:

$$(0,1),\ \ (0,-1),\ \ (1,0)\ \ e\ \ (-1,0).$$

- A matriz hessiana é dada por

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial y \partial x}(x,y) \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix} = \begin{bmatrix} 6x & 6y \\ 6y & 6x \end{bmatrix}$$

- Calculo da matriz hessiana no ponto crítico $(0,1)$

$$H(0,1) = \begin{bmatrix} 0 & 6 \\ 6 & 0 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} -\lambda & 6 \\ 6 & -\lambda \end{bmatrix} = \lambda^2 - 36$$

Os autovalores

$$\lambda^2 - 36 = 0 \Leftrightarrow \lambda = \pm 6.$$

Assim, $\lambda_1 = -6$, e $\lambda_2 = 6$. Pelo o item (iii) da Proposição 2.5 $H(0,1)$ tem um autovalor positivo e um negativo, então f tem um ponto de sela em $(0,1)$.

- Cálculo da matriz hessiana no ponto crítico $(0,-1)$

$$H(0,-1) = \begin{bmatrix} 0 & -6 \\ -6 & 0 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} -\lambda & -6 \\ -6 & -\lambda \end{bmatrix} = \lambda^2 - 36$$

Os autovalores

$$\lambda^2 - 36 = 0 \Leftrightarrow \lambda = \pm 6.$$

Assim, $\lambda_1 = -6$, e $\lambda_2 = 6$. Pelo o item (iii) da Proposição 2.5 $H(0,-1)$ tem um autovalor positivo e um negativo, então f tem um ponto de sela em $(0,-1)$.

- Cálculo da matriz hessiana no ponto crítico $(1,0)$

$$H(1,0) = \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} 6-\lambda & 0 \\ 0 & 6-\lambda \end{bmatrix} = (6-\lambda)^2$$

Os autovalores

$$(6-\lambda)^2 = 0 \Leftrightarrow \lambda = 6.$$

Note que o autovalor tem multiplicidade 2. Assim, $\lambda_1 = \lambda_2 = 6$. Pelo o item (i) da Proposição 2.5 $H(1,0)$ tem dois autovalores positivos, então f tem um mínimo local em $(1,0)$.

- Cálculo da matriz hessiana no ponto crítico $(-1,0)$

$$H(1,0) = \begin{bmatrix} -6 & 0 \\ 0 & -6 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} -6-\lambda & 0 \\ 0 & -6-\lambda \end{bmatrix} = (-6-\lambda)^2$$

Os autovalores

$$(-6-\lambda)^2 = 0 \Leftrightarrow \lambda = -6.$$

Note que o autovalor tem multiplicidade 2. Assim, $\lambda_1 = \lambda_2 = -6$. Pelo o item (ii) da Proposição 2.5 $H(-1,0)$ tem dois autovalores negativos, então f tem um máximo local em $(-1,0)$.

Exemplo 2.11. *Considere a função*

$$f(x,y,z) = (1/2)(4x^2 + 4y^2 + 2z^2 + 2xy + 2xz - 2yz)$$

O domínio de f é $\mathbb{R}^3$. Os únicos candidatos a extremantes locais são os pontos críticos, pois o domínio de f $(D_f = \mathbb{R}^3)$.

Os pontos críticos são dados por

$$\frac{\partial f}{\partial x} = 4x + y + z, \quad \frac{\partial f}{\partial y} = 4y + x - z \ \ e \ \ \frac{\partial f}{\partial z} = 2z + x - y$$

resulta que os candidatos a extremantes locais são as soluções do sistema.

$$\begin{cases} 4x & - & y & + & z & = & 0 \\ 4y & + & x & - & z & = & 0 \\ 2z & + & x & - & y & = & 0 \end{cases}$$

A solução do sistema é: (0, 0, 0).

A matriz hessiana é dado por

$$H(x,y,x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

- *Cálculo da matriz hessiana no ponto crítico* $(0, 0, 0)$

$$H(0,0,0) = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} 4-\lambda & 1 & 1 \\ 1 & 4-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{bmatrix} = -\lambda^3 + 10\lambda^2 - 29\lambda + 20$$

Os autovalores são dados por 5, 4 e 1. Como H é positiva definida, pelo o Teorema 2.7 $(0,0,0)$ *é um ponto de mínimo local não degenerado.*

Exemplo 2.12. *Considere a função*

$$g(x,y,x) = \frac{11}{2}x^2 + \frac{11}{2}y^2 + z^2 + 4xy + 5xz - 5yz$$

O domínio de g é $\mathbb{R}^3$*. Os únicos candidatos a extremantes locais são os pontos críticos, pois o domínio de g* $(D_g = \mathbb{R}^3)$.

Os pontos críticos são dados por

$$\frac{\partial g}{\partial x} = 11x + 4y + 5z, \quad \frac{\partial g}{\partial y} = 11y + 4x - 5z \ \ e \ \ \frac{\partial g}{\partial z} = 2z + 5x - 5y$$

resulta que os candidatos a extremantes locais são as soluções do sistema.

$$\begin{cases} 11x & + & 4y & + & 5z & = & 0 \\ 11y & + & 4x & - & 5z & = & 0 \\ 2z & + & 5x & - & 5y & = & 0 \end{cases}$$

A solução do sistema é: (0, 0, 0).

A matriz hessiana é dado por

$$H(x,y,x) = \begin{bmatrix} \frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial x \partial z} \\ \frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial y^2} & \frac{\partial^2 g}{\partial y \partial z} \\ \frac{\partial^2 g}{\partial x \partial z} & \frac{\partial^2 g}{\partial y \partial z} & \frac{\partial^2 g}{\partial z^2} \end{bmatrix} = \begin{bmatrix} 11 & 4 & 5 \\ 4 & 11 & -5 \\ 5 & -5 & 2 \end{bmatrix}$$

- *Cálculo da matriz hessiana no ponto crítico* $(0,0,0)$

$$H(0,0,0) = \begin{bmatrix} 11 & 4 & 5 \\ 4 & 11 & -5 \\ 5 & -5 & 2 \end{bmatrix}$$

Polinômio característico

$$P(\lambda) = det(H - \lambda I) = det \begin{bmatrix} 11-\lambda & 4 & 5 \\ 4 & 11-\lambda & -5 \\ 5 & -5 & 2-\lambda \end{bmatrix} = -\lambda^3 + 24\lambda^2 - 99\lambda + 540$$

Os autovalores são dados por 15, 12 e -3 e portanto H indefinida. pelo o Teorema 2.7, g tem um ponto de sela na origem.

Capítulo 3

Apêndice

Neste capítulo encontram-se os resultados que foram utilizados no decorrer deste trabalho sem preocupação alguma em demonstrá-los, uma vez que não prejudica o desenvolvimento do mesmo. Indicaremos os textos [2], [10] e [13] para maiores detalhes.

Teorema 3.1. *(Teorema de Schwarz) Seja* $f : U \longrightarrow \mathbb{R}$ *duas vezes diferenciável no ponto* $c \in U \subset \mathbb{R}^n$. *para quaisquer* $1 \leq i, j \leq n$, *tem-se*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(c) = \frac{\partial^2 f}{\partial x_j \partial x_i}(c).$$

Demonstração. Ver [13]. □

A fórmula de Taylor, dando uma aproximação quadrática para $f(a+y) - f(a)$, toma a seguinte forma.

Teorema 3.2. *(Fórmula de Taylor de Segunda Ordem para Campos Escalares.) Se* f *é um campo escalar, admitindo derivadas parciais de segunda ordem,* $D_{ij}f$, *contínuas numa bola aberta* B *e centro* a, *então para todo* $y \in \mathbb{R}^n$ *tal que* $a + h \in B(a)$ *tem-se*

$$f(a+y) - f(a) = \nabla f(a) \cdot h + \frac{1}{2!} h H(a+th) h^t, com 0 < t < 1. \tag{3.1}$$

Esta também pode escrever-se na forma

$$f(a+y) - f(a) = \nabla f(a) \cdot h + \frac{1}{2!} h H(a+th) h^t + \|h\|^2 E_2(a,h), \tag{3.2}$$

onde $E_2(a,h) \longrightarrow 0$ *quando* $h \longrightarrow 0$.

Demonstração. Ver [2].

Teorema 3.3. *Seja* $A = [a_{ij}]$ *uma matriz simétrica real* $n \times n$ *e considere-se*

$$Q(h) = hAh^t = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} y_i y_j.$$

tem-se então:

(*a*) $Q(h) > 0$ *para todo* $y \neq 0$ *se e só se todos os valores próprios de* A *são positivos.*

(*b*) $Q(h) < 0$ *para todo* $y \neq 0$ *se e só se todos os valores próprios de* A *são negativos.*

Nota 3.1. *Na hipótese* (*a*), *a forma quadrática diz-se* **definida positiva**; *na hipótese* (*b*), *diz-se* **definida negativa**.

Demonstração. Ver [2]. □

Teorema 3.4. *Se* f *é um campo escalar com derivadas parciais de segunda ordem* $D_{ij}f$, *contínuas numa bola aberta* B *e centro* a, *e* $H(a)$ *representa a matriz hessiana num ponto crítico* a, *então tem-se:*

(*a*) *Se todos os valores próprios de* $H(a)$ *são positivos,* f *tem um mínimo relativo em* a.

(*b*) *Se todos os valores próprios de* $H(a)$ *são negativos,* f *tem um máximo relativo em* a.

(*c*) *Se* $H(a)$ *tem valores próprios positivos e negativos, então* f *tem um ponto de sela em* a.

Demonstração. Ver [2]. □

Nota 3.2. *Se todos os valores próprios de* $H(a)$ *são nulos, o teorema 3.4 não dá qualquer informação relativa ao ponto de crítico. Podem estabelecer-se critérios para estudar tais exemplos, os quais fazem intervir derivadas de ordem superior; estes porém não serão tratados aqui.*

Para ver isto, considere as funções $f(x, y) = x^4 + y^4$, $g(x, y) = -x^4 - y^4$ e $h(x, y) = x^4 - y^4$. O ponto p $= (0, 0)$ é um ponto crítico das três funções e a matriz hessiana dessas três funções em p $= (0, 0)$ é a matriz nula. Mas $\mathbf{p} = (0, 0)$ é um ponto de mínimo local de f, um ponto de máximo local de g e um ponto de sela de h.

Teorema 3.5. *Seja a um ponto de crítico de um campo escalar $f(x_1, x_2)$ admitindo derivadas parciais de segunda ordem contínuas numa bola aberta B centro a e sejam*

$$A = D_{1,1}f(a), \qquad B = D_{1,2}f(a), \qquad C = D_{2,2}f(a),$$

e

$$\Delta = det H(a) = det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2.$$

Então verifica-se:

(*a*) *Se $\Delta < 0$, f admite um ponto em a.*

(*b*) *Se $\Delta > 0$ e $A > 0$, f admite um mínimo relativo em a.*

(*c*) *Se $\Delta > 0$ e $A < 0$, f admite um máximo relativo em a.*

(*d*) *Se $\Delta = 0$ o critério não é conclusivo.*

Demonstração. Ver [2]. □

Qual a "melhor" quádrica (polinômio de grau dois em duas variáveis) que aproxima a função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ de classe C^2 em uma vizinhança de um ponto $\mathbf{p} = (a, b)$.

Teorema 3.6. *(O Polinômio de Taylor de Ordem 2) Considere uma função $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ de classe C^2 e $\mathbf{p} = (x_0, y_0)$ um ponto interior de D. Então existe um único polinômio p_2 de grau 2 que satisfaz as condições $p_2(\mathbf{p}) = f(\mathbf{p})$, $Df(\mathbf{p}) = Dp_2(\mathbf{p})$ e $D^2f(\mathbf{p}) = D^2p_2(\mathbf{p})$.*

$$p_2(x) = f(\mathbf{p}) + Df(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + \frac{1}{2}(\mathbf{x} - \mathbf{p})^T \cdot D^2f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}),$$

onde $\mathbf{x} = (x, y)$, o símbolo ^{T} indica a operação transposição de matrizes, $Df(\mathbf{p})$ é a matriz jacobiana de f no ponto $\mathbf{p} = (a, b)$ e $D^2f(\mathbf{p})$ é a matriz hessiana de f no ponto $\mathbf{p} = (a, b)$:

$$Df(\mathbf{p}) = \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) & \frac{\partial f}{\partial y}(\mathbf{p}) \end{bmatrix}, \qquad D^2f(\mathbf{p}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{p}) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{p}) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{p}) & \frac{\partial^2 f}{\partial y^2}(\mathbf{p}) \end{bmatrix}.$$

Mais ainda, vale que

$$f(\mathbf{x}) = p_2(\mathbf{x}) + R_2(\mathbf{p}, \mathbf{x}),$$

onde o erro $R_2(\mathbf{p}, \mathbf{x})$ satisfaz a condição

$$\lim_{\mathbf{x} \to \mathbf{p}} \frac{R_2(\mathbf{p}, \mathbf{x})}{\|\mathbf{x} - \mathbf{p}\|^2} = 0,$$

isto é, a diferença entre as funções $f(\mathbf{x})$ e $p_2(\mathbf{x})$ vai para zero mais rapidamente do que o quadrado da distância entre os pontos $\mathbf{p}$ *e* $\mathbf{x}$.

Demonstração. Ver [2]. □

A aproximação de Taylor indica que o fato de $\mathbf{p}$ ser um extremo local de f está intimamente relacionado com o sinal do polinômio de grau 2 em várias variáveis $Q(\mathbf{h}) = \mathbf{h}^T \cdot D^2 f(\mathbf{p}) \cdot \mathbf{h}$ onde $D^2 f(\mathbf{p})$ é a matriz hessiana de f no ponto crítico $\mathbf{p}$.

Se $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ é uma função de classe C^2 e $\mathbf{p}$ é um ponto interior a D, então

$$f(\mathbf{x}) = f(\mathbf{p}) + Df(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + \frac{1}{2}(\mathbf{x} - \mathbf{p})^T \cdot D^2 f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p})n + R_2(\mathbf{p}, \mathbf{x}).$$

Se $\mathbf{p}$ é um ponto crítico de f, então $Df(\mathbf{p}) = \mathbf{0}$ e, portanto,

$$f(\mathbf{x}) = f(\mathbf{p}) + \frac{1}{2}(\mathbf{x} - \mathbf{p})^T \cdot D^2 f(\mathbf{p}) \cdot (\mathbf{x} - \mathbf{p})n + R_2(\mathbf{p}, \mathbf{x}).$$

Usando a variável $\mathbf{h} = \mathbf{x} - \mathbf{p}$ (que representa o deslocamento com relação ao ponto $\mathbf{p}$) e negligenciando o erro $R_2(\mathbf{p}, \mathbf{x})$ (que vai para zero rapidamente quando $\mathbf{x} \longrightarrow \mathbf{p}$), concluímos que

$$f(\mathbf{p} + \mathbf{h}) \approx f(\mathbf{p}) + \frac{1}{2} \cdot \mathbf{h}^T \cdot D^2 f(\mathbf{p}) \cdot \mathbf{h},$$

isto é,

$$f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p}) \approx \frac{1}{2} \cdot \mathbf{h}^T \cdot D^2 f(\mathbf{p}) \cdot \mathbf{h}.$$

Se ao lado direito da expressão acima é maior do que zero para todo $\mathbf{h} \neq \mathbf{0}$ suficientemente pequeno, então o lado esquerdo também deve ser maior do que zero:

$$f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p}) > 0 \qquad \text{ou} \qquad f(\mathbf{p} + \mathbf{h}) > f(\mathbf{p}),$$

para todo $\mathbf{h} \neq \mathbf{0}$ suficientemente pequeno. Desta maneira $\mathbf{p}$ é um **mínimo local** de f. Analogamente, se $\mathbf{h}^T \cdot D^2 f(\mathbf{p}) \cdot \mathbf{h}$ é menor do que zero para $\mathbf{h} \neq \mathbf{0}$ suficientemente pequeno, então o lado esquerdo também deve ser menor do que zero:

$$f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p}) < 0 \qquad \text{ou} \qquad f(\mathbf{p} + \mathbf{h}) < f(\mathbf{p}),$$

para todo $\mathbf{h} \neq \mathbf{0}$ suficientemente pequeno. Desta maneira $\mathbf{p}$ é um **máximo local** de f.

Proposição 3.1. *Seja $f(x,y,z)$ de classe C^2 e seja (x_0,y_0,z_0) um ponto interior de Df. Suponha que (x_0,y_0,z_0) é ponto crítico de f. Sejam $H(x,y,z)$ e $H_1(x_0,y_0,z_0)$ dadas por*

$$H = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{vmatrix} \quad e \quad H_1 = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}$$

Mostre que:

(*i*) *Se $\frac{\partial^2 f}{\partial x^2}(x_0,y_0,z_0) > 0$, $H_1(x_0,y_0,z_0) > 0$ e $H(x,y,z) > 0$ então (x_0,y_0,z_0) será ponto de mínimo local.*

(*ii*) *Se $\frac{\partial^2 f}{\partial x^2}(x_0,y_0,z_0) < 0$, $H_1(x_0,y_0,z_0) > 0$ e $H(x,y,z) < 0$ então (x_0,y_0,z_0) será ponto de máximo local.*

Demonstração. ver [13]. □

Referências Bibliográficas

[1] ANTON, Howard, *Álgebra Linear com Aplicações*/ Anton Howard e Chris Rorres; trad Claus Ivo Doering. - 8ª ed. Porto Alegre, 2001.

[2] APOSTOL, T.M. Apostol, *Calculus,* Volume II, 2ª ed. Jonh Wiley & Sons, New York, 1969.

[3] BORTOLOSSI, Humberto José. *Cálculo Diferencial a Várias Variáveis: uma introdução à teoria de otimização.* 5ª ed Rio de Janeiro: PUC-RIO, 2011.

[4] BOYER, Carl B. *História da Matemática.* Traduçao de Elza F. Gomide. 2ª ed. São Paulo: Edgard Blücher, 1996.

[5] BUENO, Hamilton Prado. *Álgebra Linear: Um Segundo Curso.* Rio de Janeiro, SBM, 2006.

[6] CAJORI, Florian. *Uma História da Matemática.* Rio de Janeiro, Editora Ciência Moderna Ltda., 2007.

[7] CARVALHO, João Pitombeira.; *Introdução à Álgebra Linear.* Rio de Janeiro: Livros Técnicos e Científicos Editora S.A. 1974.

[8] COELHO, Flávio U. LOURENÇO, Mary L. *Um Curso de Álgebra Linear.* São Paulo, Editora da Universidade de São Paulo, 2001.

[9] EVES, Howard. *Introdução à história da matemática.* tradução: Hygino H. Domingues. - Campinas, SP, Editora da UNICAMP, 2004.

[10] GUIDORIZZI, Hamilton L. *Cálculo.* Volume 2, 5ª ed. - Rio de Janeiro, LTC editora, 2001.

[11] LAWSON, Terry. *Álgebra Linear.* Tradução: Elza F. Gomide. - São Paulo: Blucher, 1997.

[12] LIMA, Elon Lages. *Álgebra Linear.* 7ª ed. Rio de Janeiro, IMPA, 2008.

[13] LIMA, Elon Lages. *Curso de Análise*, vol.2. 9ª ed. Rio de Janeiro, IMPA, 2006.

[14] POOLE, David. *Álgebra Linear.* Tradutoras técnicas Martha Salerno Monteiro (coord.) ... [et al.]. - São Paulo: Thomson Learning, 2006.

[15] STEINBRUCH, Alfredo. *Álgebra Linear I.* 2ª ed. São Paulo, Pearson Makron Books, 1987.

[16] *Breve Historia del Álgebra Matricial.* Disponível em `http://http://personales.upv.es/jbenitez/cajon_sastre/histam.pdf` (Acessado em 05/09/2013.)

[17] *El Teorema Espectral. Orígenes y Evolucion.* Disponível em `http://www.ehu.es/~mtpalezp/libros/anafunap2.pdf` (Acessado em 05/09/2013.)

[18] *Sobre el Surgimiento del Concepto de Valor Propio: Una Historia Selecta sobre los Orígenes de la Teoría Espectral.* Disponível em `http://www.miscelaneamatematica.org/Misc43/CarMtnez_a.pdf` (Acessado em 05/09/2013.)