UNIVERSIDADE FEDERAL DE ALFENAS

ESTELA COSTA FERREIRA

ESPAÇOS DE HILBERT DE REEPRODUÇÃO E APROXIMAÇÃO DE SOLUÇÕES DE EQUAÇÕES INTEGRAIS DE VOLTERRA

ALFENAS/MG

2016

ESTELA COSTA FERREIRA

ESPAÇOS DE HILBERT DE REEPRODUÇÃO E APROXIMAÇÃO DE SOLUÇÕES DE EQUAÇÕES INTEGRAIS DE VOLTERRA

Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Área de concentração: matemática aplicada e modelagem matemática.
Orientador: Prof. Dr. José Claudinei Ferreira
Coorientador: Prof. Dr. Evandro Monteiro.

ALFENAS/MG

2016

Dados Internacionais de Catalogação-na-Publicação (CIP)
Biblioteca Central da Universidade Federal de Alfenas

Ferreira, Estela Costa.
Espaços de Hilbert de reprodução e aproximação de soluçoes e equações integrais de volterra / Estela Costa Ferreira. -- Alfenas/MG, 2016.
102 f.

Orientador: José Claudinei Ferreira.
Dissertação (mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Alfenas, 2016.
Bibliografia.

1. Matrizes não-negativas. 2. Volterra, Equações de. 3. Hilbert, Espaço de. I. Ferreira, José Claudinei. II. Título.

CDD-519

MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Alfenas / UNIFAL-MG
Programa de Pós-graduação em Estatística Aplicada e Biometria
Rua Gabriel Monteiro da Silva, 700. Alfenas - MG CEP 37130-000
Fone: (35) 3299-1392 (Secretaria) (35) 3299-1121(Coordenação)
https://www.unifal-mg.edu.br/ppgeab/

1914 100 anos 2014
Universidade Federal de Alfenas

ESTELA COSTA FERREIRA

"ESPAÇOS DE HILBERT DE REPRODUÇÃO E APROXIMAÇÃO DE SOLUÇÕES DE EQUAÇÕES INTEGRAIS DE VOLTERRA"

A Banca Examinadora, abaixo assinada, aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Linha de Pesquisa: Matemática Aplicada e Modelagem Matemática.

Aprovado em: 29 de fevereiro de 2016.

Prof. Dr. José Claudinei Ferreira
Instituição: UNIFAL-MG

Assinatura: ____________________

Profa. Dra. Thaís Jordão
Instituição: ICMC-USP

Assinatura: ____________________

Prof. Dr. Mário Henrique de Castro
Instituição: FAMAT-UFU

Assinatura: ____________________

Dedico este trabalho a Deus por guiar meus caminhos.
À minha família pelo incentivo, amor e carinho.
Aos meus amigos pela convivência, apoio e atenção.

AGRADECIMENTOS

Inicio meus agradecimentos a DEUS, já que Ele colocou pessoas tão especiais a meu lado, sem as quais certamente não teria dado conta! À minha família meu infinito agradecimento. Sempre acreditaram em minha capacidade. Isso só me fortaleceu e me fez tentar fazer o melhor de mim. Obrigada pelo amor incondicional!

Aos professores que expandiram o meu campo de visão. Mostrando-me tantas coisas novas e impulsionando o meu crescimento de tantas formas diferentes. Em especial ao Prof. Dr. José Claudinei, meu orientador, pela paciência e dedicação em me explicar inúmeras vezes a mesma dúvida.

À minha amiga de sempre, Maria Cecília, por só querer o meu bem e me valorizar tanto como pessoa. Obrigada pela amizade e tantos fins de semana me ouvindo e apoiando! Sua amizade foi imprescindível para o meu crescimento e sua família se tornou a minha família. Agradeço ainda a outra amiga que sempre torceu por mim. Obrigada Laís por me acolher tantas vezes em sua casa com tanta hospitalidade. Aos colegas de estudo por tantas horas quebrando a cabeça juntos e em especial a Bruna que tanto me ajudou com minhas dúvidas.

Todos vocês fizeram-me enxergar que existe mais que pesquisadores e resultados por trás de uma dissertação. Existe um caminho a ser escolhido, uma opção de futuro, uma lição de vida. Vocês foram e são referências profissionais e pessoais para meu crescimento.

RESUMO

O objetivo deste trabalho é encontrar uma solução exata para um sistema de equações integrais de Volterra. Para isso, usaremos a teoria de espaços de reprodução e núcleos positivos definidos, visto que as técnicas usuais de resoluções de equações diferenciais e integrais possuem restrições. Grande parte do estudo voltado a solução de equações se baseia em analisar o comportamento das soluções, o chamado estudo qualitativo. Este não é o nosso interesse, queremos aproximar a solução do problema usando a representação dessa solução em uma base ortonormal especial de um espaço de Hilbert de reprodução gerado por um núcleo positivo definido adequado. Dessa forma, truncando a série encontrada para a solução do sistema de Volterra podemos exibir uma boa aproximação para a solução do sistema. As equações integrais de Volterra, foco deste trabalho, são importantes para a modelagem de fenômenos físicos, demográficos ou epidemiológicos. Para a resolução de tais equações, faremos um estudo introdutório sobre conceitos de álgebra linear, análise e teoria da medida com o intuito de abranger temas como: existência de base de um espaço vetorial, o processo de ortogonalização de Gram-Schmidt, os espaços $\mathcal{L}^p$, entre outros. Faremos uma breve análise sobre a transformada de Laplace, assim como resolveremos uma equação diferencial e integral usando este método. Também resolveremos um sistema de equações integrais através da transformada de Laplace para exemplificar o método. Cabe lembrar que a maioria das equações não pode ser resolvida por meio da transformada de Laplace. Faremos um estudo de resolução de equações lineares de Volterra e então abrangeremos esse estudo para equações não lineares.

Palavras-chave: Núcleos positivos definidos. Equações de Volterra. Espaços de Reprodução.

ABSTRACT

The aim of this study is to give the exact solution to a system of linear Volterra integral equations. So do it, we will use the theory of reproduction Kernel method and positive definite kernels, since the usual method to solve differential and integral equations have restrictions. Much of the study about solving equations is based on analyzing the behavior of solutions, called qualitative study. This is not our interest, we want to approach the solution of the problem using the representation of the solution in a special orthonormal basis of the reproduction kernel Hilbert space generated by an appropriate positive definite kernel. Thus, truncating the series found for the solution of the Volterra system, we can give a good approximation to the system solution. The Volterra integral equations, focus of this work, are important to modeling physical, demographic or epidemiological phenomena. For solving such equations, we make an introductory study of linear algebra, analysis and measure theory in order to comprehend topics such as: existence of a base in a vector space, the Gram-Schmidt orthogonalization process, the spaces $\mathcal{L}^p$, and others. We make a brief analysis of the Laplace transform, as well as solve a differential and integral equation using this method. We also solve a system of integral equations by Laplace transform to illustrate the method. It should be noted that most of the equations can not be solved by means of the Laplace transform. We will study how to solve linear Volterra equations and then extend the study to nonlinear equations.

Keywords: Positive definite kernel. Volterra integral equation. Reproducing kernel spaces.

SUMÁRIO

1 INTRODUÇÃO

Modelos matemáticos são importantes para entendermos o comportamento de determinados fenômenos. Na modelagem de diversos fenômenos físicos e biológicos usamos as equações integrais de Volterra. Estas foram introduzidas por Vito Volterra, um famoso matemático italiano que publicou, em 1896, um trabalho sobre equações integrais, que conhecemos hoje como equações integrais de Volterra. Os trabalhos de Volterra juntamente com os de Ivar Fredholm, matemático sueco, marcaram o começo do estudo da análise funcional [1].

O intuito deste trabalho é encontrar uma solução para um sistema de equações integrais de Volterra. Para tanto, precisaremos de dois espaços de Hilbert de reprodução. Usaremos os núcleos positivos definidos desses espaços para encontrar uma base ortonormal conveniente de tal forma que consigamos calcular os coeficientes da solução do sistema de equações de Volterra.

Usaremos a teoria de espaços de Hilbert de reprodução e núcleos positivos definidos para a resolução desse sistema, pois os métodos usuais para resolução de equações diferenciais e integrais possuem restrições quanto a sua utilização. O uso da transformada de Fourier na resolução de equações diferenciais e integrais está limitado por se aplicar a funções integráveis em $\mathbb{R}$, o que exige que as funções decresçam para zero suficientemente rápido em $\pm\infty$. Contudo, é frequentemente necessário considerar equações diferenciais com soluções em que isso não acontece. Esta limitação pode ser contornada, até certo ponto, pela consideração de uma outra transformada, a transformada de Laplace. Pelo fato de considerar integrais em $\mathbb{R}^+$ a transformada de Laplace é particularmente apropriada em problemas de valor inicial para equações diferenciais onde o objetivo é determinar as soluções no intervalo de tempo que se segue ao instante inicial, enquanto a transformada de Fourier é mais adequada para problemas de valores na fronteira. Contudo, a transformada de Laplace não é suficiente para a resolução de todas as equações diferencias e integrais. Por isso, são necessários outros métodos de resolução de equações como a utilização da teoria de espaços de reprodução e núcleos positivos definidos, justificando o estudo deste trabalho.

No começo do século XIX, a teoria de núcleos positivos definidos começou a se desenvolver com maior intensidade. Desde então, podemos dividi-la em duas tendências principais: estudar os núcleos positivos definidos e suas propriedades ou estudar as funções pertencentes ao espaços gerados por núcleos positivos definidos [2].

A primeira dessas tendências tem origem na teoria de equações integrais, onde os operadores integrais são gerados por núcleos positivos definidos. Antes de receber o nome de núcleo positivo definido, Carl Frederich Gauss, por volta de 1809, mencionou um núcleo positivo definido em seu livro sobre movimento de corpos celestes, referindo-se ao cálculo de mínimos quadrados em máxima verossimilhança. Um dos núcleos positivos definidos mais importantes é o núcleo Gaussiano, também conhecido como distribuição normal. Posteriormente, Mercer usou-o para muitas outras equações integrais.

Na segunda tendência, Zaremba foi o primeiro a introduzir em seu trabalho sobre problemas de valores iniciais, em um caso particular, um núcleo correspondendo a uma classe de funções. Moore, provou o teorema que liga as duas tendências, mostrando que cada núcleo positivo definido gera um espaço que pode ser completado formando um espaço de Hilbert, que conhecemos hoje como espaço de Hilbert de reprodução. Ao longo dos anos, os matemáticos que trabalharam em cada uma das tendências perceberam as ideias gerais e indispensáveis que estavam sendo usadas. Ainda segundo [2], as duas tendências, como matriz hermitiana positiva ou núcleos de reprodução, são equivalentes e os métodos elaborados em uma tendência se tornam fundamentais na outra.

A teoria de espaços de reprodução e núcleos positivos definidos tem sido aplicada com sucesso em muitos problemas lineares e não lineares, tais como equações diferenciais, modelos populacionais e diversas equações que aparecem na física e na engenharia. Este avanço na resolução de equações se deve à construção dos espaços de Hilbert de reprodução, $W[0,1]$ e $W_1[0,1]$. Nosso estudo será baseado nos artigos [3] e [18] que tratam da resolução de sistemas lineares e não lineares, respectivamente. Os artigos [4, 5, 6] serão usados como auxiliares.

No primeiro capítulo, faremos uma revisão de conceitos de análise, álgebra linear e teoria da medida. Apresentaremos o teorema do ponto fixo de Banach, que será essencial para garantir que a solução do sistema de equações de Volterra é única, e daremos uma breve introdução sobre transformada de Laplace na resolução de equações. Daremos ênfase a Identidade de Parseval pois esta será a estrutura para a construção da solução do sistema de equações Volterra.

No segundo capítulo, abrangemos a teoria de núcleos positivos definidos com algumas definições e teoremas pertinentes ao interesse desse trabalho. Em seguida, abordamos a teoria de espaços de reprodução, relacionando-a com núcleos positivos definidos. Também apresentaremos os espaços $W[0,1]$ e $W_1[0,1]$ que serão usados, no capítulo seguinte, para a resolução das equações integrais de Volterra.

No capítulo quatro, apresentaremos a definição das equações de Volterra e a construção das bases para os espaços $W[0,1]$ e $W_1[0,1]$. Assim como a solução algébrica do sistema de equações integrais de Volterra dada em função dos respectivos núcleos positivos definidos. Por fim, faremos a generalização do método de resolução de sistemas lineares de Volterra para resolvermos sistemas não lineares. Este capítulo é baseado no artigo [18] e é estruturado pelos mesmos conceitos preliminares estudados nos capítulos anteriores.

2 CONCEITOS PRELIMINARES

Determinar a área ou o volume de uma região do plano ou espaço é um problema antigo. O comprimento, a área e o volume são conceitos importantes de medida. Outros conceitos, como massa e a carga elétrica, são essenciais na física. Na estatística, medimos a probabilidade de x pertencer a S. Segundo [7], foi Kolmogorov, em 1933, quem reformulou a teoria das probabilidades em termos da teoria de Lebergue e de medidas σ-aditivas. Faremos um estudo introdutório de alguns conceitos de teoria da medida que serão importantes neste trabalho.

Neste capítulo, apresentaremos alguns conceitos preliminares de análise e teoria da medida. Também introduziremos os conceitos de transformada de Laplace, obtendo uma forma de resolver algumas equações integrais de Volterra, porém este método possui limitações para um conjunto maior de equações. Dessa forma, há motivação para o estudo de métodos de resolução através da teoria de núcleos positivos definidos.

2.1 BASE E BASE ORTONORMAL

Estamos interessados em encontrar um conjunto B contido em um espaço vetorial V de tal forma que todos os elementos de V possam ser escritos como combinação linear finita dos elementos de B e queremos que B seja um conjunto linearmente independente. Se pudermos encontrar tais vetores, encontraremos uma base do espaço vetorial V.

Primeiramente, apresentaremos algumas definições pertinentes ao assunto. Como o conceito de combinação linear, que é uma expressão construída a partir da soma dos elementos de um conjunto, onde cada elemento é multiplicado por uma constante.

Definição 2.1.1 *Seja V um espaço vetorial sobre $\mathbb{K}$. Um vetor $v \in V$ é uma combinação linear finita dor vetores $v_1, v_2, \cdots, v_n \in V$ se existirem escalares $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$ tais que:*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i.$$

Definição 2.1.2 *Sejam V um espaço vetorial sobre $\mathbb{K}$ e B um subconjunto de V. Dizemos que B é um conjunto gerador de V se todo elemento de V for uma combinação linear finita* *de elementos de B.*

Denotaremos por $V = [B]$.

Definição 2.1.3 *Sejam V um $\mathbb{K}$-espaço vetorial e B um subconjunto de V. Então,*

a) B é linearmente independente (LI) se a única forma de se escrever o vetor nulo $\overline{0}$ de V como combinação linear dos elementos de B é com todos os escalares iguais a zero;

b) O conjunto B é chamado linearmente dependente (LD) se não for linearmente independente.

Ao juntarmos os conceitos de conjunto gerador e conjunto linearmente independente obtemos a ideia de base de um espaço vetorial.

Definição 2.1.4 *Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um subconjunto B de V é uma base de V se:*

a) B for um conjunto gerador de V, e

b) B for linearmente independente.

Se existe uma base finita de V com n elementos, diz-se que a dimensão de V, denotada por dim V, é finita e igual a n. Caso contrário, a dimensão de V é infinita.

O próximo teorema depende do lema de Zorn. Mas antes, faremos uma série de definições necessárias para o seu entendimento (Veja [1] para mais detalhes).

Definição 2.1.5 *Seja X um conjunto qualquer diferente de vazio. Uma relação de ordem parcial sobre X, que denotaremos por $\prec$, é uma relação que satisfaz:*

a) Reflexividade: $x \prec x$ para qualquer $x \in X$;

b) Transitividade: $x \prec y$ e $y \prec z$ então $x \prec z$, com $x,y,z \in X$;

c) Antissimetria: $x \prec y$ e $y \prec x$ então $x = y$, com $x,y \in X$.

Um conjunto parcialmente ordenado é um par $(X, \prec)$ consistindo de um conjunto X e uma ordem parcial sobre o mesmo.

Definição 2.1.6 *Uma relação de ordem total sobre X é uma relação de ordem parcial $\prec$ sobre X que para quaisquer $x,y \in X$,*

$$x \prec y \quad ou \quad y \prec x$$

Um conjunto totalmente ordenado é um par $(X, \prec)$ consistindo de um conjunto X e uma ordem total sobre o mesmo.

Definição 2.1.7 *Seja $A \subseteq X$. Um elemento m é denominado cota superior para A se $x \prec m$ para todo $x \in A$.*

Definição 2.1.8 *Considere $(X, \prec)$ um conjunto parcialmente ordenado. Um elemento $m \in X$ é denominado maximal se $m \prec x$ implicar $x = m$, $x \in X$.*

Lema 2.1.9 (Lema de Zorn) *Um conjunto não vazio, parcialmente ordenado, no qual todo subconjunto totalmente ordenado possui um cota superior, possui um elemento maximal.*

Teorema 2.1.10 *Todo espaço vetorial não-trivial possui uma base.*

Demonstração: Sejam V um espaço vetorial sobre $\mathbb{K}$ e C um subconjunto LI de V. Observe que o conjunto C existe, pois um conjunto com um único elemento não nulo é sempre LI. Seja P a classe de todos os subconjuntos LI que contenham C. Note que $P \neq \varnothing$, pois $C \in P$, e P é parcialmente ordenado, pela relação de ordem obtida pela inclusão de conjuntos, já que nem sempre podemos dizer que um conjunto contém o outro ou o contrário.

Para usar o lema de Zorn, precisamos mostrar que todo subconjunto totalmente ordenado possui cota superior. Assim, seja $D = \{A_\alpha\}_{\alpha \in I}$ um subconjunto totalmente ordenado de P. Um candidato a cota superior de D é $\cup_{\alpha \in I} A_\alpha$. Note que todo $A_\alpha \in D$ está na união e ainda, $C \in \cup_{\alpha \in I} A_\alpha$. Vamos mostrar que $\cup_{\alpha \in I} A_\alpha$ é LI.

Seja $L = \{v_1,v_2,\cdots,v_n\}$ um subconjunto de $\cup_{\alpha \in I} A_\alpha$. Então, para cada $i = 1,\cdots,n$, $\exists\, \alpha_i \in I$ tal que $v_i \in A_{\alpha_i}$. Como D é um conjunto totalmente ordenado, podemos reorganizar os elementos de L afim de obter:

$$A_{\alpha_1} \subset A_{\alpha_2} \subset \cdots \subset A_{\alpha_n}.$$

Assim, podemos afirmar que $v_i \in A_{\alpha_n}$ $\forall i = 1, \cdots ,n$. Como D é subconjunto de P e P é a classe dos subconjuntos LI que contém C, então os conjuntos em D são LI. Logo, A_α é LI e então L é LI. Como L é um subconjunto qualquer de $\cup_{\alpha \in I} A_\alpha$, segue que $\cup_{\alpha \in I} A_\alpha$ é LI. Logo $\cup_{\alpha \in I} A_\alpha$ é cota superior de D.

Pelo lema de Zorn, P possui um elemento maximal. Denotemos por B o elemento maximal de P e vamos mostrar que B gera V.

Suponhamos que exista $v \in V$ tal que v não seja gerado pela combinação linear finita dos elementos de B. Então, $B \cup \{v\}$ é LI. Mas B é maximal, então $B \cup \{v\} \subset B$. Contradição. Logo, não existe $v \in V$ que não possa ser escrito como combinação linear finita dos elementos de B. Portanto, $V = [B]$. Como B é LI segue que B é base de V. $\square$

Sempre é possível exibir uma base de um espaço vetorial não nulo de dimensão finita. Porém, isso nem sempre é verdade para espaços vetoriais de dimensão infinita.

Se o espaço vetorial em questão tiver um produto interno e o produto interno entre dois elementos de um espaço vetorial é igual a zero, dizemos que estes elementos são ortogonais. Quando os elementos são ortogonais, podemos garantir que formam um conjunto linearmente independente.

Definição 2.1.11 *Dois elementos x,y em um espaço com produto interno V são ditos ortogonais se $\langle x, y \rangle = 0$.*

Denotaremos por $E^\perp$ o conjunto de todos os vetores de V ortogonais ao subespaço E de V. Ou seja,

$$E^\perp = \{x \in V; \langle x, z \rangle = 0, \forall z \in E\}.$$

A partir do conceito de ortogonalidade podemos definir uma base ortogonal, onde os vetores são ortogonais entre si. E ainda, a ideia de base ortonormal, abrangendo os conceitos de ortogonalidade e vetores unitários.

Definição 2.1.12 *Um conjunto B em V é uma base ortonormal se satisfizer:*

a) O produto interno entre dois elementos distintos de B é zero, ou seja, os elementos são ortogonais;

b) Cada vetor possui norma unitária; e

c) B é um conjunto ortonormal total, ou seja, $\overline{[B]} = V$, onde $\overline{[B]}$ denota o menos conjunto fechado que contém $[B]$, ou seja, o fecho de $[B]$.

Definição 2.1.13 *Um espaço vetorial V é a soma direta de dois de seus subespaços U e W se todo $v \in V$ possui representação única dada por*

$$v = u + w, \quad u \in U \quad e \quad v \in V.$$

Denotamos a soma direta por:

$$V = U \oplus W.$$

Os espaços de Hilbert são convenientes pois apresentam conjuntos ortonormais, que podem ser usados para decompor vetores.

Teorema 2.1.14 *Se E é um subespaço vetorial fechado de um espaço de Hilbert H, então*

$$H = E \oplus E^{\perp}.$$

O subespaço $E^{\perp}$ é chamado de complemento ortogonal do subespaço E em H (Veja [1, p. 129]).

Teorema 2.1.15 *Seja H um espaço de Hilbert.*

a) Se E é um subespaço fechado de H, então $(E^{\perp})^{\perp} = E$. Assim, neste caso, E é o complemento ortogonal de $E^{\perp}$.

b) Se M é um subconjunto de H, então

$$\overline{[M]} = H \Leftrightarrow M^{\perp} = \{0\}.$$

Demonstração: (a). Sabendo que $E^{\perp} = \{x \in H; \langle x, y\rangle = 0, \forall y \in E\}$, podemos escrever que $(E^{\perp})^{\perp} = \{x \in H; \langle x, z\rangle = 0, \forall z \in E^{\perp}\}$. Dessa forma, notamos que $E \subseteq (E^{\perp})^{\perp}$. Como E é subespaço fechado de H, temos que

$$E \oplus E^{\perp} = H. \tag{2.1.1}$$

Note ainda que $E^{\perp\perp} = (E^{\perp})^{\perp}$, logo

$$(E^{\perp})^{\perp} \oplus E^{\perp} = H. \tag{2.1.2}$$

Assim, por (2.1.1) e (2.1.2), obtemos

$$(E^{\perp})^{\perp} \oplus E^{\perp} = E \oplus E^{\perp}.$$

Da unicidade da soma direta, concluímos que

$$E = E^{\perp\perp}.$$

(b) Sejam $N = [M]$ e $\overline{N}$ o seu fecho. Como $\overline{N}$ é fechado, segue que

$$\overline{N} \oplus \overline{N}^{\perp} = H.$$

Observe que $N^{\perp} = M^{\perp}$, pois N é obtido através de combinações lineares finitas dos elementos de M. Observe ainda que $N^{\perp} = \overline{N}^{\perp}$, pois, se $\langle x_n, y\rangle = 0$ e $x_n \to u$, então $\langle u, y\rangle = 0$.

Assim,

$$\overline{N} \oplus M^{\perp} = H.$$

Dessa forma, $\overline{N} = H \Leftrightarrow M^{\perp} = \{0\}$. □

Teorema 2.1.16 (Representação de Riez) *Sejam H um espaço de Hilbert e $f : H \to K$ um funcional linear contínuo. Existe um único $v \in H$ tal que $f(u) = \langle u,v\rangle$, $\forall u \in H$.*

Demonstração: Se f for identicamente nula, basta tomar $v = 0$. Caso contrário, considere o $Nuc(f) = \{u \in H; f(u) = 0\}$. O conjunto $Nuc(f)$ é um subespaço fechado de H. Logo,

$$Nuc(f) \oplus (Nuc(f))^{\perp} = H.$$

Seja $z \in (Nuc(f))^{\perp}$ tal que $\|z\| = 1$. Definindo $w = f(u)z - f(z)u$, então $w \in Nuc(f)$, pois

$$\begin{aligned} f(w) &= f(f(u)z - f(z)u) \\ &= f(u)f(z) - f(z)f(u) \\ &= 0 \end{aligned}$$

Portanto, $w \in Nuc(f)$. Assim,

$$\begin{aligned}
0 &= \langle w, z\rangle \\
&= \langle f(u)z - f(z)u, z\rangle \\
&= \langle f(u)z, z\rangle - \langle f(z)u, z\rangle \\
&= f(u)\langle z,z\rangle - f(z)\langle u,z\rangle \\
&= f(u)\|z\| - f(z)\langle u,z\rangle \\
&= f(u) - \langle u,\overline{f(z)}z\rangle.
\end{aligned}$$

Dessa forma, $f(u) = \langle u,\overline{f(z)}z\rangle$. Seja $v = \overline{f(z)}z$, então $f(u) = \langle u, v\rangle$.

Para provarmos a unicidade, suponhamos $v_1, v_2 \in H$ tais que $f(u) = \langle u, v_1\rangle = \langle u, v_2\rangle$, $\forall u \in H$. Assim,

$$\begin{aligned}
f(v_1 - v_2) &= \langle v_1 - v_2, v_1\rangle = \langle v_1 - v_2, v_2\rangle \\
&\Leftrightarrow \langle v_1 - v_2, v_1 - v_2\rangle = 0 \\
&\Leftrightarrow v_1 - v_2 = 0 \\
&\Leftrightarrow v_1 = v_2.
\end{aligned}$$

$\square$

Teorema 2.1.17 *Todo espaço de Hilbert não trivial possui uma base ortonormal.*

Demonstração: Seja $H \neq \{0\}$ um espaço de Hilbert. Seja ϑ a classe de todos os subconjuntos ortonormais em H. Note que $\vartheta \neq \varnothing$, pois dado um elemento $\xi \neq 0$ pertencente a H, o conjunto que contém apenas $\dfrac{\xi}{\|\xi\|}$ é ortonormal, ou seja, $\left\{\dfrac{\xi}{\|\xi\|}\right\}$ é um subconjunto ortonormal de H. Observe que ϑ é um subconjunto parcialmente ordenado de H, pela relação de inclusão de conjuntos. Semelhante a demonstração do Teorema 2.1.10 considere um subconjunto $D = \{A_\alpha\}_{\alpha\in I}$ parcialmente ordenado de ϑ. Uma cota superior de D é $\cup_{\alpha\in I} A_\alpha$, a qual pertence a ϑ, já que H é um espaço de Hilbert. Dessa forma, pelo lema de Zorn, existe um elemento maximal de ϑ. Seja M tal elemento e o candidato a base ortonormal de H.

Suponhamos que M não seja base ortonormal de H. Assim, pelo Teorema 2.1.15, existe um vetor v não nulo tal que $v \in \overline{[M]}^{\perp}$. Assim, $M \cup \left\{\dfrac{v}{\|v\|}\right\}$ é um conjunto ortonormal. O que contraria o fato de M ser um elemento maximal em ϑ. Logo, M é uma base ortonormal de ϑ. $\square$

Teorema 2.1.18 (Desigualdade de Bessel) *Se $B = \{u_i\}_{i \in I}$ é um conjunto ortonormal em H, então para cada $u \in H$*

$$\sum_{\alpha \in I} |\langle u_\alpha, u \rangle|^2 \leq \|u\|^2.$$

Demonstração:

Seja B_1 um subconjunto finito de B e considere $W = \overline{[B_1]}$, assim:

$$H = W \oplus W^\perp.$$

Assim, dado $u \in H$, podemos escrevê-lo como $u = p + q$, com $p \in W$ e $q \in W^\perp$. Como $p \in W$, existem escalares $\alpha_1, \cdots, \alpha_n$, tais que

$$p = \sum_{i=1}^{n} \alpha_i u_i.$$

Observe que $u - p = q \in W^\perp$, logo $u - p \perp u_i$, $i = 1, \cdots, n$. Dessa forma,

$$\begin{aligned}
0 &= \langle u_i, u - p \rangle \\
&= \langle u_i, u \rangle - \langle u_i, p \rangle \\
&= \langle u_i, u \rangle - \left\langle u_i, \sum_{i=j}^{n} \alpha_j u_j \right\rangle \\
&= \langle u_i, u \rangle - \sum_{i=j}^{n} \alpha_j \langle u_i, u_j \rangle \\
&= \langle u_i, u \rangle - \alpha_i.
\end{aligned}$$

Portanto,

$$\alpha_i = \langle u_i, u \rangle$$

Ou seja,

$$p = \sum_{i=1}^{n} \langle u_i, u \rangle u_i.$$

Note que,

$$
\begin{aligned}
0 &\leq \langle u-p, u-p\rangle \\
&= \left\langle u - \sum_{i=1}^{n}\langle u_i, u\rangle u_i, u - \sum_{i=1}^{n}\langle u_i, u\rangle u_i \right\rangle \\
&= \langle u, u\rangle - \left\langle u, \sum_{i=1}^{n}\langle u_i, u\rangle u_i \right\rangle - \left\langle \sum_{i=1}^{n}\langle u_i, u\rangle u_i, u \right\rangle + \left\langle \sum_{i=1}^{n}\langle u_i, u\rangle u_i, \sum_{i=1}^{n}\langle u_i, u\rangle u_i \right\rangle \\
&= \|u\|^2 - \overline{\sum_{i=1}^{n}\langle u_i, u\rangle u_i \langle u, u_i\rangle} - \sum_{i=1}^{n}\langle u_i, u\rangle u_i \langle u_i, u\rangle + \sum_{i=1}^{n}\langle u_i, u\rangle u_i \overline{\langle u_i, u\rangle}\langle u_i, u_i\rangle \\
&= \|u\|^2 - 2\sum_{i=1}^{n}\langle u, u_i\rangle\langle u, u_i\rangle + \sum_{i=1}^{n}|\langle u, u_i\rangle|^2 \\
&= \|u\|^2 - \sum_{i=1}^{n}|\langle u, u_i\rangle|^2.
\end{aligned}
$$

Portanto,

$$\sum_{i=1}^{n}|\langle u, u_i\rangle|^2 \leq \|u\|^2.$$

Para o caso em que os elementos u_i, $i = 1, \cdots, n$ formam um conjunto ortonormal finito, a demonstração está concluída. Se B não for enumerável, considere o conjunto J dado por

$$J = \{i \in I; \langle u, u_i\rangle \neq 0\}.$$

Afirmação: J é finito ou enumerável.

De fato, para cada $K \in \mathbb{N}$, definimos

$$J_k = \left\{ i \in J; |\langle u, u_i\rangle| > \frac{1}{k} \right\}.$$

Para obter $J = \bigcup_{k=1}^{\infty} J_k$, basta mostrar que J_k é finito. Sabemos que para todo subconjunto finito J_0 de J, vale a desigualdade

$$\sum_{i\in J_0}|\langle u, u_i\rangle|^2 \leq \|u\|^2.$$

Em particular, dado um número finito de elementos $i_1, \cdots, i_n$ de J_k, podemos escrever

$$\begin{aligned} \frac{n}{k^2} &= \frac{1}{k^2} + \cdots + \frac{1}{k^2} \\ &< |\langle u, u_{i_1} \rangle|^2 + \cdots + |\langle u, u_{i_n} \rangle|^2 \\ &\leq \|u\|^2. \end{aligned}$$

Dessa forma, $n < k^2\|u\|^2$. Ou seja, J_k é finito, pois o número de elementos de qualquer subconjunto de J_k é inferior a $k^2\|u\|^2$. Consequentemente, J é finito ou enumerável. Assim, seja $i_1, i_2 \ldots,$ uma enumeração qualquer dos elementos de J. Para cada $n \in \mathbb{N}$, temos

$$\sum_{k=1}^{n} |\langle u, u_{i_k} \rangle|^2 \leq \|u\|^2.$$

Dessa forma, fazendo n tender a infinito, obtemos

$$\sum_{i \in J} |\langle u, u_i \rangle|^2 \leq \|u\|^2.$$

□

Dada uma base ortonormal em um espaço de Hilbert, é possível definir coordenadas de vetores desse espaço através dessa base. Essas coordenadas são conhecidas como coeficientes de Fourier [1] (Veja [1]).

Teorema 2.1.19 *Sejam H um espaço de Hilbert e B um conjunto ortonormal em H. Então as seguintes afirmações são equivalentes:*

a) B é uma base ortonormal de H;

b) Se $u \in H$ satisfizer $\langle u, v \rangle = 0$, para todo $v \in B$, então $u = 0$.

Demonstração: $(a) \Rightarrow (b)$

Sejam $B = \{v_i\}_{i \in I}$ uma base ortonormal de H e $u \in H$. Quaisquer que sejam

[1] O termo "coeficientes de Fourier"é dado pela similaridade as séries de Fourier. O matemático e físico Jean-Baptiste Joseph Fourier desenvolveu a sua teoria enquanto estudava uma solução para a equação da onda, publicou os seus resultados por volta de 1800 em um trabalho intitulado "A teoria analítica do calor". Este trabalho recebeu muitas críticas de Laplace e Lagrange por excesso de simplificação e falta de rigor matemático.

v_i,$v_j \in B$, temos que

$$\langle v_i, v_j\rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

Além disso, para cada $n \in \mathbb{N}$, existe $p(n) \in \mathbb{N}$ tal que

$$u_n = \sum_{j=1}^{p(n)} \beta_{jn} v_{\alpha_{jn}}, \quad v_{\alpha_{jn}} \in B,$$

com β_{jn} escalares, tais que $\|u - u_n\| \to 0$. Seque que, se u satisfizer $\langle u, v\rangle = 0$, para todo $v \in B$, então

$$\|u - u_n\|^2 = \|u\|^2 + \sum_{j=1}^{p(n)} |\beta_{jn}|^2 \to 0.$$

Assim, $\|u\| = 0$ e $u = 0$.

$(b) \Rightarrow (a)$

Seja $M = [B]$, então $\overline{M}^{\perp} = B^{\perp}$. Pelo teorema 2.1.15, obtemos que

$$H = \overline{M} \oplus B^{\perp}.$$

Por hipótese, temos $B^{\perp} = \{0\}$, e assim, podemos concluir que B é um conjunto ortonormal maximal, pois $\overline{[B]} = H$. Ou seja, B é uma base ortonormal de H. □

Teorema 2.1.20 (Identidade de Parseval) *Seja $B = \{u_i\}_{i\in I}$ uma base ortonormal de H, então para todo $u \in H$, temos que*

$$u = \sum_{i\in I} \langle u_i, u\rangle u_i.$$

Além disso,

$$\|u\|^2 = \sum_{i\in I} |\langle u_i, u\rangle|^2.$$

Demonstração: Seja $u \in H$. Pelo Teorema 2.1.18 temos que $\sum_{i\in J} |\langle u_i, u\rangle|^2$ é convergente. Assim, os termos da série $\sum_{i\in J} \langle u_i, u\rangle u_i$ formam uma sequência de Cauchy. Defina $x = u - \sum_{i\in J} \langle u_i, u\rangle u_i$. Note que $\langle u_i, x\rangle = 0$, pois

$$\begin{aligned}
\langle u_i, x\rangle &= \left\langle u_i, u - \sum_{i\in J}\langle u_i, u\rangle \right\rangle \\
&= \langle u_i, u\rangle - \sum_{i\in J}\langle u_i, u\rangle\langle u_i, u_i\rangle \\
&= 0.
\end{aligned}$$

Pelo teorema (2.1.19), concluímos que $x = 0$, ou seja:

$$u = \sum_{i\in J}\langle u_i, u\rangle u_i$$

Para a segunda igualdade, note que

$$\|u\|^2 = \langle u, u\rangle$$

Dessa forma,

$$\begin{aligned}
\|u\|^2 &= \left\langle \sum_{i\in I}\langle u, u_i\rangle u_i, \sum_{i\in I}\langle u, u_i\rangle u_i \right\rangle \\
&= \sum_{i\in I} |\langle u, u_i\rangle\langle u, u_i\rangle| \langle u_i, u_i\rangle \\
&= \sum_{i\in I} |\langle u, u_i\rangle|^2 \langle u_i, u_i\rangle \\
&= \sum_{i\in I} |\langle u, u_i\rangle|^2.
\end{aligned}$$

□

2.1.1 O processo de ortogonalização de Gram-Schmidt

A partir de um subconjunto B linearmente independente qualquer de um espaço vetorial V sobre $\mathbb{K}$, com produto interno $\langle\cdot,\cdot\rangle$, podemos criar outro conjunto linearmente independente e ortogonal C para o qual $[B] = [C]$. Para isso, usaremos um processo construtivo. Considere um conjunto linearmente independente $A = \{v_1, v_2, \cdots, v_n\} \subseteq V$.

Construiremos um outro conjunto $A' = \{w_1, w_2, \cdots, w_n\} \subseteq V$ que seja ortogonal e de tal forma que o espaço gerado por A e A' seja o mesmo, ou seja, $[A] = [A']$. A construção é intuitiva, como segue:

$w_1 = v_1$.

$w_2 = v_2 - \dfrac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1$.

Definidos w_1,w_2,$\cdots$,w_k, para $1 < k < n$, podemos definir w_{k+1} da seguinte maneira:

$$\begin{aligned} w_{k+1} &= v_{k+1} - \frac{\langle v_{k+1}, w_1\rangle}{\|w_1\|^2} w_1 - \cdots - \frac{\langle v_{k+1}, w_k\rangle}{\|w_k\|^2} w_k \\ &= v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j\rangle}{\|w_j\|^2} w_j. \end{aligned}$$

Como $\{v_1, v_2\}$ é um conjunto LI, segue que $w_2 \neq 0$, pois v_2 não pode ser escrito como combinação linear de v_1. Note ainda que w_2 é ortogonal a w_1. De fato,

$$\langle w_2, w_1\rangle = \left\langle v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1, w_1 \right\rangle.$$

Assim,

$$\begin{aligned} \langle w_2, w_1\rangle &= \langle v_2, w_1\rangle - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} \langle w_1, w_1\rangle \\ &= \langle v_2, w_1\rangle - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} \|w_1\|^2 \\ &= 0. \end{aligned}$$

Logo, $w_1 \perp w_2$. Observe que o conjunto $\{w_1$,w_2,$\cdots$,$w_n\}$ pela maneira como foi definido é ortogonal. Mostraremos que $A' = \{w_1$,w_2,$\cdots$,$w_n\}$ é um conjunto LI. Sejam $w \in [A']$ e α_1,$\cdots$,$\alpha_n \in \mathbb{K}$. Assim, podemos escrever:

$$w = \sum_{i=1}^{n} \alpha_i w_i.$$

Como A' é um conjunto ortogonal $\langle w_i$,$w_j\rangle = 0$ sempre que $i \neq j$, concluímos então que:

$$\begin{aligned}\langle w, w_j\rangle &= \left\langle \sum_{i=1}^{n} \alpha_i w_i, w_j \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle w_i, w_j\rangle \\ &= \alpha_j \langle w_j, w_j\rangle.\end{aligned}$$

Para $j = 1, \cdots, n$.

Portanto, $\alpha_j = \dfrac{\langle w, w_j\rangle}{\|w_j\|^2}$, $\forall j = 1, \cdots, n$, e $w = \displaystyle\sum_{i=1}^{n} \frac{\langle w, w_i\rangle}{\|w_i\|^2} w_i$.

Agora, suponhamos que existam escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$ tais que:

$$\alpha_1 w_1 + \cdots + \alpha_n w_n = 0.$$

Acabamos de concluir que $\alpha_j = \dfrac{\langle w, w_j\rangle}{\|w_j\|^2}$, $\forall j = 1, \cdots, n$. Em particular, para $w = 0$, temos:

$$\alpha_i = \frac{\langle 0, w_i\rangle}{\|w_i\|^2} = 0 \quad i = 1, \cdots, n.$$

Portanto, A' é um conjunto LI. Observe ainda que, $dim[A] = dim[A'] = n$ e que para cada $i = 1, \cdots, n$, w_i é dado por uma combinação linear dos elementos de A. Logo, os subespaços gerados por A e A' são iguais. Segue que A' é uma base ortogonal de V.

O processo de ortogonalização de Gram-Schmidt [2] também pode ser usado em uma base enumerável. Formalizaremos esse resultado no teorema a seguir.

Teorema 2.1.21 *Sejam H um espaço de Hilbert e $\{v_n\}$ uma coleção enumerável linearmente independente em H. Assim, existe uma coleção ortonormal $\{w_n\}$ em H com a mesma cardinalidade de $\{v_n\}$. Além disso, para todo $m \in \mathbb{N}$, temos que $\{v_1, \cdots, v_m\}$ e $\{w_1, \cdots, w_m\}$ geram o mesmo subespaço vetorial.*

Demonstração: Seja $\{v_n\}$ uma coleção enumerável linearmente independente em H. Vamos construir a sequência $\{w_n\}$ por recorrência. Defina $w_1 = \dfrac{v_1}{\|v_1\|}$. Note que $\{v_1\}$ e $\{w_1\}$ geram o mesmo subespaço vetorial. Similar ao que foi feito anteriormente, temos

[2] J.P. Gram introduziu, em 1883, uma generalização no processo de mínimos quadrados. Ele usou o processo de ortogonalização que é atribuído a Erhard Schmidt, um aluno de Hilbert. De onde vem a nomenclatura "ortogonalização de Gram-Schmidt".

que, dado um conjunto ortonormal $\{\eta_1, \cdots ,\eta_k\}$, o vetor $\eta - \sum_{j=1}^{k} \langle \eta_j,\eta \rangle \eta_j$ é ortogonal a cada η_j. Assim, seja $v'_{k+1} = v_k - \sum_{j=1}^{k} \langle v_j,v_k \rangle v_j$ e $w_{k+1} = \dfrac{v'_{k+1}}{\|v'_{k+1}\|}$. Pela construção o conjunto $\{w_1, \ldots ,w_{k+1}\}$ é ortonormal e gera o mesmo subespaço que $\{v_1, \ldots ,v_{k+1}\}$. □

2.2 O ESPAÇO $\mathcal{L}^p$

Os espaços $\mathcal{L}^p$ são espaços de Banach muito estudados em análise funcional. Algumas definições e teoremas úteis para a definição desses espaços serão feitas a seguir. Para podermos estudar propriedades dos espaços de Hilbert de reprodução que definiremos posteriormente no trabalho precisamos de alguns resultados desta seção.

Definição 2.2.1 *Uma classe $\mathcal{A}$ de subconjuntos de um conjunto X é denominada álgebra quando são satisfeitas as seguintes propriedades:*

a) $X \in \mathcal{A}$;

b) Se $A \in \mathcal{A}$ então $A^C \in \mathcal{A}$;

c) Se $A, B \in \mathcal{A}$ então $A \cup B \in \mathcal{A}$.

Quando para todo $A_n \in \mathcal{A}$, com $n \in \mathbb{N}$, temos $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$ e, além disso, essa classe satisfaz todas as propriedades de uma álgebra, denominamos $\mathcal{A}$ como σ-álgebra. Ou seja, uma σ-álgebra é uma álgebra que é fechada para a união enumerável de conjuntos.

As álgebras são muito usadas para definir medidas. O conceito é importante em análise e probabilidade. Note que o espaço amostral $S(\Omega)$ das probabilidades sempre é uma σ-álgebra. A obervação a seguir generaliza este fato.

Observação 2.2.2 *Dado um conjunto $\mathcal{A}$, o conjunto de suas partes $\mathbb{P}(\mathcal{A})$, assim como o conjunto $\{\varnothing,\mathcal{A}\}$, são σ-álgebras de $\mathcal{A}$.*

Definição 2.2.3 *Sejam M um espaço qualquer e $\mathcal{A}$ uma σ-álgebra associada de subconjuntos de M. O par $(M,\mathcal{A})$ denomina-se espaço mensurável.*

Definiremos a medida em $(M,\mathcal{A})$ como a função $\mu : \mathcal{A} \to [0,\infty]$ tal que:

a) $\mu(\emptyset) = 0$;

b) Se A_n é uma sequência de elementos disjuntos de $\mathcal{A}$ então $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

A tripla $(M,\mathcal{A},\mu)$ é denominada espaço de medida.

Teorema 2.2.4 *Seja $(M,\mathcal{A},\mu)$ um espaço de medida.*

a) (Monotonicidade) Se $E, F \in \mathcal{A}$ e $E \subset F$, então $\mu(E) \leq \mu(F)$;

b) (Sub-aditividade) Se $\{E_j\}_{j=1}^{\infty} \subset \mathcal{A}$, então $\mu(\cup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \mu(E_j)$;

c) (Semi-continuidade Inferior) Se $\{E_j\}_{j=1}^{\infty}$ é uma sequência de conjuntos em $\mathcal{A}$ tal que $E_1 \subset E_2 \subset \cdots$, então $\mu(\cup_{j=1}^{\infty} E_j) = \lim_{j\to\infty} \mu(E_j)$;

d) (Semi-continuidade Superior) Se $\{E_j\}_{j=1}^{\infty}$ é uma sequência de conjuntos em $\mathcal{A}$ tal que $E_1 \supset E_2 \supset ...$ e $\mu(E_1) < \infty$, então $\mu(\cap_{j=1}^{\infty} E_j) = \lim_{j\to\infty} \mu(E_j)$.

Demonstração: (a) Como $E \subset F$, então

$$\mu(F) = \mu(E) + \mu(F \setminus E).$$

Assim,

$$\mu(F) \geqslant \mu(E).$$

(b) Sejam $F_1 = E_1$, $F_2 = E_2 \cap [E_1]^c, \cdots, F_k = E_k \cap [\cup_{j=1}^{\infty} E_j]^c$. Logo, os $F_k's$ são disjuntos e $\cup_{j=1}^{\infty} F_j = \cup_{j=1}^{\infty} E_j$, $\forall n$. Assim

$$\mu(\cup_{j=1}^{\infty} E_j) = \mu(\cup_{j=1}^{\infty} F_j).$$

Por definição, temos

$$\mu(\cup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(F_j).$$

Como $F_k \subset E_k$ $\forall k$, então pelo item (1)

$$\mu(\cup_{j=1}^{\infty} E_j) \leqslant \sum_{j=1}^{\infty} \mu(E_j).$$

(c) Seja $E_0 = \emptyset$, então

$$\mu(\cup_{j=1}^{\infty} E_j) = \mu(\cup_{j=1}^{\infty} E_j \backslash E_{j-1}).$$

Por definição, temos

$$\mu(\cup_{j=1}^{\infty} E_j \backslash E_{j-1}) = \sum_{j=1}^{\infty} \mu(E_j \backslash E_{j-1}) = \lim_{n\to\infty} \sum_{j=1}^{n} \mu(E_j \backslash E_{j-1}).$$

Note que $\cup_{j=1}^{\infty} E_j \setminus E_{j-1}$ é uma união disjunta e por hipótese $E_1 \subset E_2 \subset \cdots$, logo

$$\mu(\cup_{j=1}^{\infty} E_j) = \lim_{n\to\infty} \mu(E_n).$$

(d) Sejam $F_j = E_1 \backslash E_j$. Como $E_1 \supset E_2 \supset \ldots$, então $F_1 \subset F_2 \subset \ldots$. Note que

$$\mu(E_1) = \mu(F_j) + \mu(E_j) \tag{2.2.1}$$

e

$$\cup_{j=1}^{\infty} F_j = E_1 \backslash \cap_{j=1}^{\infty} E_j. \tag{2.2.2}$$

Pelo demonstrado em (c), temos

$$\mu(\cup_{j=1}^{\infty} F_j) = \lim_{j\to\infty} \mu(F_j).$$

Assim, substituindo em (2.2.1) e aplicando o limite, temos

$$\begin{aligned} \lim_{j\to\infty} \mu(E_1) &= \lim_{j\to\infty} \mu(F_j) + \lim_{j\to\infty} \mu(E_j) \\ &= \mu(\cup_{j=1}^{\infty} F_j) + \mu(\cap_{j=1}^{\infty} E_j). \end{aligned}$$

Pela igualdade (2.2.2), obtemos

$$\begin{aligned} \lim_{j\to\infty} \mu(E_1) &= \mu(E_1 \backslash \cap_{j=1}^{\infty} E_j) + \mu(\cap_{j=1}^{\infty} E_j) \\ &= \mu(E_1) - \lim_{j\to\infty} \mu(E_j) + \mu(\cap_{j=1}^{\infty} E_j). \end{aligned}$$

Como $\mu(E_1) < \infty$, podemos subtraí-lo em ambos os lados da equação, obtemos

$$\mu(\cap_{j=1}^{\infty} E_j) = \lim_{j\to\infty} \mu(E_j),$$

conforme o desejado. □

A ideia de continuidade corresponde à ideia de preservação da família de subconjuntos abertos. Já a ideia de mensurabilidade diz que a família de conjuntos mensuráveis é preservada. Dessa forma, as funções mensuráveis em teoria da medida possuem um papel semelhante ao das funções contínuas em topologia (veja [10] para mais detalhes). A definição a seguir formaliza essa ideia.

Definição 2.2.5 *Seja $\mathcal{A}$ uma σ-álgebra em M. Uma função $f : M \to \mathbb{R}$ é $\mathcal{A}$-mensurável quando para todo $r \in \mathbb{R}$ vale $[f > r] \in \mathcal{A}$, onde*

$$[f > r] = \{x \in M | f(x) > r\} = f^{-1}(r, +\infty).$$

Definição 2.2.6 *Sejam (X,M,μ) um espaço da medida e $p \in [1,\infty)$. Define-se o espaço $\mathcal{L}^p$ como o espaço de funções dado por*

$$\mathcal{L}^p = \{f : X \to \mathbb{C}; f \text{ é } \mu\text{-mensurável e } \|f\|_p < \infty\}$$

Onde,

$$\|f\|_p = \left[\int_X |f(x)|^p d\mu(x)\right]^{\frac{1}{p}}$$

Segundo [11], o conjunto $\mathcal{L}^p(X)$ torna-se um espaço vetorial quando identificamos quaisquer duas funções f e g de $\mathcal{L}^p(X)$ que são idênticas a menos de um conjunto de medida nula. Dizemos que f e g são iguais quase sempre ou, simplificadamente, $f = g$ q.s. Antes de mostrarmos que $\mathcal{L}^p(X)$ é um espaço de Banach, faremos algumas considerações.

Lema 2.2.7 *(Desigualdade de Young) Sejam $A,B \in \mathbb{R}$ positivos. Se $p,q \in [1,\infty)$ são tais que $\frac{1}{p} + \frac{1}{q} = 1$, então*

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}$$

Demonstração: Considere a função f definida em $(0,\infty)$, dada por $f(t) = \alpha t - t^{\alpha}$, $0 < \alpha < 1$. Temos que

$$f'(t) = \alpha - \alpha t^{\alpha - 1} = \alpha \left(1 - \frac{1}{t^{1-\alpha}}\right).$$

Se $0 < t < 1$, então $f'(t) < 0$. Se $t > 1$, então $f'(t) > 0$. Dessa forma, $t = 1$ é ponto de mínimo local, ou seja, $f(1) \leq f(t)$, $\forall t \in (0,\infty)$. Logo,

$$\begin{aligned} \alpha - 1 &\leq \alpha t - t^{\alpha} \\ t^{\alpha} &\leq \alpha t - \alpha + 1, \qquad \forall t \in (0,\infty). \end{aligned}$$

Fazendo $t = \dfrac{a}{b}$, com $a,b \in (0,\infty)$, obtemos

$$\begin{aligned} \left(\frac{a}{b}\right)^{\alpha} &\leq \alpha \left(\frac{a}{b}\right) - \alpha + 1 \\ \Leftrightarrow a^{\alpha} b^{1-\alpha} &\leq \alpha a - \alpha b + b \\ \Leftrightarrow a^{\alpha} b^{1-\alpha} &\leq \alpha a + b(1-\alpha). \end{aligned}$$

Tomando $\alpha = \dfrac{1}{p}$, $A = a^{\frac{1}{p}}$ e $B = b^{1-\frac{1}{p}}$, concluímos que

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

$\square$

Lema 2.2.8 *(Desigualdade de Holder) Sejam $f \in \mathcal{L}^p$ e $g \in \mathcal{L}^q$, onde $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Então $fg \in \mathcal{L}^1$ e vale que:*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Demonstração: Se $f = 0$ ou $g = 0$ o resultado segue. Suponhamos que $\|f\|_p \neq 0$ $\|g\|_q \neq 0$. Temos que fg é mensurável pois o produto de funções mensuráveis é mensurável ([7],p. 93). Note que podemos supor que $\|f\|_p = \|g\|_q = 1$, já que podemos multiplicar constantes positivas na desigualdade do enunciado. Pela desigualdade de Young, temos

$$|fg(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}.$$

Aplicando a integral, temos

$$\int_M |fg|d\mu \leq \int_M \frac{|f^p|}{p}d\mu + \frac{|g^q|}{q}d\mu.$$

Logo,

$$\|fg\|_1 \leq \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} = 1 = \|f\|_p\|g\|_q,$$

e segue que $fg \in \mathcal{L}^1$. $\square$

Lema 2.2.9 *(Desigualdade de Minkowsky) Sejam* $1 < p < \infty$ *e* $f,g \in \mathcal{L}^p$*. Então*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Demonstração: Note que

$$\begin{aligned} |f+g|^p &\leq (max(|f|,|g|))^p \\ &\leq 2^p(|f|^p + |g|^p). \end{aligned}$$

Logo,

$$\begin{aligned} \int_M |f+g|^p &\leq \int_M 2^p(|f|^p + |g|^p)d\mu \\ &= 2^p\left(\int_M |f|^p d\mu + \int_M |g|^p d\mu\right) \\ &< \infty. \end{aligned}$$

Logo, $f+g \in \mathcal{L}^p$. Vamos estimar o valor de $\|f+g\|_p$. Note que

$$\begin{aligned} |f+g|^p &= |f+g||f+g|^{p-1} \\ &\leq (|f|+|g|)|f+g|^{p-1} \end{aligned}$$

Observe ainda que $p = (p-1)q$, pois p é o conjugado de q, assim

$$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}.$$

Então

$$p = (p-1)q.$$

A partir desse fato e aplicando a desigualdade de Holder, obtemos

$$\begin{aligned}\int |f+g|^p &\leq \int \|f\|_p \||f+g|^{p-1}\|_q + \int \|g\|_p \||f+g|^{p-1}\|_q \\ &= (\|f\|_p + \|g\|_p) \left(\int |f+g|^p\right)^{\frac{1}{q}}.\end{aligned}$$

Isolando o termo desejado, temos

$$\left(\int |f+g|^p\right)^{1-\frac{1}{q}} \leq \|f\|_p + \|g\|_p.$$

Portanto, $\|f+g\|_p = \left(\int |f+g|^p\right)^{1-\frac{1}{q}} \leq \|f\|_p + \|g\|_p.$ □

Teorema 2.2.10 *(Convergência Dominada em $\mathcal{L}^p$) Seja $0 \leq p < \infty$. Sejam f_n e $f : M \to \mathbb{K}^m$ funções μ-mensuráveis tais que $f_n \to f$ quase sempre. Se existe alguma função μ-mensurável $g : M \to [0,\infty]$ tal que $\int_M g^p d\mu < \infty$ e $|f_n| \leq g$ quase sempre, $\forall n$, então f_n e $f \in \mathcal{L}^p$ e $\|f_n - f\|_p \to 0$.*

Teorema 2.2.11 *(Riesz-Fisher) Seja $1 < p < \infty$. Então, o espaço $\mathcal{L}^p$ munido com a norma*

$$\|f\|_p = \left(\int_M |f|^p d\mu\right)^{\frac{1}{p}}$$

é um espaço de Banach.

Teorema 2.2.12 (Teorema de Fubini) *Sejam (X, M, μ) e (Y, N, ν) espaços de medida σ finito, $\sigma = \mu \times \nu$ a medida produto e $f : X \times Y \to \mathbb{R}$ uma função σ-mensurável. Se $f \in \mathcal{L}^1(\sigma)$, então*

$$\int_{X\times Y} f d\sigma = \int_X \left(\int_Y f(x,y) d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y) d\mu(x)\right) d\nu(y).$$

Mais detalhes sobre o teorema de Fubini podem ser encontrados em [11].

2.3 TEOREMA DO PONTO FIXO DE BANACH

Muitos problemas de aplicações envolvem resolver uma equação não linear do tipo

$$A(x) = x, \quad x \in X.$$

Esse tipo de equação é conhecido como problema de pontos fixos pela aplicação $A : X \to X$ e será de grande valia para garantirmos a existência de soluções das equações integrais de Volterra que trataremos à frente. O teorema do ponto fixo de Banach, além de garantir a existência e unicidade de um ponto fixo, em um contexto bastante geral, também fornece um processo iterativo para encontrá-lo.

Definição 2.3.1 *Sejam* (X,d) *e* (Y,D) *espaços métricos. Uma aplicação* $A : X \to Y$ *é uma contração se existe uma constante* $0 \leq c < 1$ *de modo que, para todos* $x, y \in X$,

$$D(A(x),A(y)) \leq c\ d(x,y).$$

Um ponto fixo é definido como um ponto que não é alterado por uma aplicação. A próxima definição formaliza essa ideia.

Definição 2.3.2 *Um ponto fixo de uma aplicação* $T : X \to X$ *é um ponto* $\overline{x}$ *tal que,*

$$T(\overline{x}) = \overline{x}.$$

Exemplo 2.3.3 *Considere a função* $T : \mathbb{R} \to \mathbb{R}$ *dada por* $T(x) = x^3$.

Para encontrarmos os pontos fixos de T devemos descobrir os pontos $\overline{x}$ tais que

$$T(\overline{x}) = \overline{x}.$$

Assim, basta resolver a equação

$$\overline{x}^3 = \overline{x}$$

Dessa forma,

$$\overline{x}^3 - \overline{x} = 0$$

$$\overline{x}(\overline{x}^2 - 1) = 0$$

Assim, os pontos -1, 0 e 1 são os pontos fixos de T.

Teorema 2.3.4 (Teorema do Ponto Fixo de Banach) *Seja* (X,d) *um espaço métrico completo. Seja* T *uma contração definida em* X*. Então*

a) Existe um único $\overline{x}$ *pertencente a* X *tal que* $T(\overline{x}) = \overline{x}$*;*

b) Qualquer que seja x_0 *pertencente a* X*, a sequência definida por* $a_0 = x_0$ *e* $a_n = T(a_{n-1})$ *converge para* $\overline{x}$*;*

c) Qualquer sequência da forma anterior satisfaz:

$$d(a_n,\overline{x}) \leq \frac{c^{n-1}}{1-c} d(a_0,a_1).$$

Demonstração: Vamos mostrar que se existe um ponto fixo de T, ele é único.

Suponhamos x_1 e x_2 pontos fixos distintos, então,

$$d(T(x_1),T(x_2)) = d(x_1,x_2). \tag{2.3.1}$$

Como T é contração, então:

$$d(T(x_1),T(x_2)) \leq cd(x_1,x_2). \tag{2.3.2}$$

Por (2.3.1) e (2.3.2), podemos afirmar que:

$$d(x_1,x_2) \leq cd(x_1,x_2).$$

Pela lei do cancelamento, concluímos que:

$$c \geq 1$$

Dessa forma, chegamos a um absurdo, pois $0 < c < 1$. Logo, $x_1 = x_2$. Portanto existe um único ponto fixo. Para provar a existência de um ponto fixo, precisaremos da afirmação a seguir.

Afirmação: Nas condições do teorema, vale a desigualdade:

$$d(a_n,a_{n+1}) \leq c^{n-1}d(a_0,a_1).$$

De fato, usaremos indução para mostrar este resultado. Primeiramente, para $n = 0$, temos:

$$d(a_0,a_1) \leq c^{-1}d(a_0,a_1).$$

Como $0 < c < 1$ então $c^{-1} > 1$, logo a desigualdade é valida. Suponhamos por indução, que a hipótese é válida para $n = k$, ou seja:

$$d(a_k,a_{k+1}) \leq c^{k-1}d(a_0,a_1).$$

Vamos mostrar que vale para $n = k + 1$. Usando o fato de T ser contração e nossa hipótese de indução, obtemos:

$$\begin{aligned}
d(a_{k+1},a_{k+2}) &= d(T(a_k),T(a_{k+1})) \\
&\leq cd(a_k,a_{k+1}) \\
&\leq c\, c^{k-1}d(a_0,a_1) \\
&= c^k d(a_0,a_1)
\end{aligned}$$

Portanto, a relação vale para todo $n \in \mathbb{N}$.

Agora, para concluir a existência de ponto fixo, precisamos provar que a sequência definida no teorema é uma sequência de Cauchy. Usando a desigualdade triangular podemos escrever

$$\begin{aligned}
d(a_n,a_{n+p}) &\leq d(a_n,a_{n+1}) + d(a_{n+1},a_{n+2}) + \cdots + d(a_{n+p-1},a_{n+p}) \\
&\leq c^{n-1}d(a_0,a_1) + c^n d(a_0,a_1) + \cdots + c^{n+p-2}d(a_0,a_1) \\
&= (c^{n-1} + c^n + \cdots + c^{n+p-2})d(a_0,a_1) \\
&= c^{n-1}\left(\sum_{i=0}^{p-1} c^i\right) d(a_0,a_1)
\end{aligned}$$

Note que $\sum_{i=0}^{p-1} c^i$ é uma soma parcial de uma série geométrica que converge para $\dfrac{1}{1-c}$. Assim,

$$\begin{aligned} d(a_n,a_{n+p}) &\leq c^{n-1}\frac{1}{1-c}d(a_0,a_1) \\ &= \frac{c^{n-1}}{1-c}d(a_0,a_1) \end{aligned}$$

Como $\dfrac{c^{n-1}}{1-c}$ torna-se arbitrariamente pequeno, então a sequência é de Cauchy. Como X é completo, a sequência converge. Ou seja, existe $\overline{x}$ tal que $lim_{n\to\infty}a_n = \overline{x}$.

Por fim, basta mostrar que $\overline{x}$ é ponto fixo de T. Para isso, observe que

$$\begin{aligned} d(T(\overline{x}),a_{n+1}) &= d(T(\overline{x}),T(a_n)) \\ &\leq d(\overline{x},a_n). \end{aligned}$$

Como $a_n \to \overline{x}$, então $d(\overline{x},a_n) \to 0$. Dessa forma, $d(T(\overline{x}), a_{n+1}) \to 0$. Pela unicidade do limite, podemos concluir que:

$$T(\overline{x}) = \overline{x}.$$

□

Corolário 2.3.5 *Sejam R um conjunto fechado do espaço métrico completo (X,d) e $T : R \to R$. Se existe $m \in \mathbb{N}$ tal que $T^m : R \to R$ é uma contração, então T possui um, e somente um, ponto fixo em R.*

Demonstração: Seja $T^m : R \to R$ uma contração. Logo, pelo teorema do ponto fixo de Banach, existe um único $\overline{x} \in R$ tal que $T^m(\overline{x}) = \overline{x}$. Note que

$$T^m(T(\overline{x})) = T(T^m(\overline{x})) = T(\overline{x}).$$

Logo, $T(\overline{x})$ é ponto fixo de T^m. Mais ainda, todo ponto fixo de T é ponto fixo de T^m. Assim, pela unicidade do ponto fixo, temos que

$$T(\overline{x}) = \overline{x}.$$

Portanto, $\overline{x}$ é o único ponto fixo de T em R. □

2.4 TRANSFORMADA DE LAPLACE

A transformada de Laplace [3] é muito importante para a resolução de equações diferenciais e integrais. Esse método permite resolver os cálculos para algumas equações integrais lineares de Volterra e com isso podemos mostrar uma forma direta de resolução desse tipo de equação.

2.4.1 Transformada de Laplace

Os primórdios da transformada de Laplace se encontram nos trabalhos de Euler, seguido por outros grandes matemáticos como Lagrange, Abel, Liouville, Riemann, Poincaré, entre outros. J.L. Lagrange considerou, em 1773, as integrais referentes à transformada de Laplace com o propósito de analisar as funções densidade de probabilidade. P.S. Laplace começou, em 1782, a considerar fórmulas integrais para expressar soluções de equações diferenciais no espírito introduzido por Euler e em 1810 tinha avançado substancialmente a aplicação da transformada para equações integrais e diferenciais.

A transformada de Laplace é definida da seguinte forma:

Definição 2.4.1 *Seja $f(t)$ uma função definida no intervalo $0 \leq t < \infty$. Então a integral*

$$\mathfrak{L}[f] \;=\; \int_0^\infty e^{-st} f(t)dt, \tag{2.4.1}$$

é chamada Transformada de Laplace, desde que a integral exista, e s é chamado parâmetro da transformada.

Por exemplo, a transformada de Laplace da função $f(t) = t$, é dada por $\frac{1}{s^2}$. De fato,

$$\mathfrak{L}[t] = \int_0^\infty e^{-st} t dt$$

Assim,

[3] Poincaré, em uma publicação em 1884, foi o primeiro a utilizar o termo “transformada de Laplace”na acepção atual; as referências anteriores eram a “método de Laplace”.

$$\begin{aligned}
\mathfrak{L}[t] &= \lim_{b\to\infty}\int_0^b e^{-st}t dt \\
&= \lim_{b\to\infty}\left(-\frac{te^{-st}}{s}\ |_0^b - \int_0^b -\frac{e^{-st}}{s}dt\right) \\
&= \lim_{b\to\infty}\left(\frac{1}{s}\right)\left(\frac{e^{-sb}}{-s}+\frac{e^{-s0}}{s}\right) \\
&= \frac{1}{s^2}
\end{aligned}$$

O resultado foi encontrado utilizando a substituição $u = t$ e $dv = e^{-st}dt$.

Agora note que, quando $\mathfrak{L}[t]$ converge, o resultado é uma função na variável s, ou seja $\mathfrak{L}[f] = F(s)$.

Definição 2.4.2 *Uma função $f : [0{,}\infty) \to \mathbb{R}$ é chamada admissível ou de ordem exponencial se são satisfeitas as seguintes condições:*
a) *A função f for contínua por partes em $[0,\infty)$.*
b) *Se existirem duas constantes M e μ tais que para todo $t \in [0,\infty)$ vale a desigualdade*

$$|f(t)| \leq Me^{\mu t}.$$

Teorema 2.4.3 *Seja $f : [0,\infty) \to \mathbb{R}$ contínua por partes e de ordem exponencial para $t > T$. Então $\mathfrak{L}[f]$ existe para $s > \mu$.*

Demonstração: Pela definição de transformada de Laplace, temos:

$$\begin{aligned}
\mathfrak{L}[f] &= \int_0^\infty e^{-st}f(t)dt \\
&= \int_0^T e^{-st}f(t)dt + \int_T^\infty e^{-st}f(t)dt.
\end{aligned}$$

Note que a integral $\displaystyle\int_0^T e^{-st}f(t)dt$, do lado esquerdo da soma sempre é um valor real, pois T é um número, assim a integral se tornará uma constante.

Estudando a integral do lado direito da soma,

$$\int_T^\infty |e^{-st}f(t)|dt \leq \int_T^\infty e^{-st}|f(t)|dt \leq \int_T^\infty e^{-st}Me^{\mu t}dt.$$

Agora,

$$\begin{aligned}
\int_T^{\infty} e^{-st} M e^{\mu t} dt &= M\left(\lim_{b\to\infty}\int_T^{b} e^{-(s-\mu)t} dt\right)\\
&= M\left(\lim_{b\to\infty}\frac{e^{-(s-\mu)t}}{-(s-\mu)}|_T^b\right)\\
&= M\left(\lim_{b\to\infty}\frac{e^{-(s-\mu)b}}{-(s-\mu)}-\frac{e^{-(s-\mu)T}}{-(s-\mu)}\right)\\
&= \frac{Me^{-(s-\mu)T}}{s-\mu}\\
&< \infty
\end{aligned}$$

Portanto essa integral converge e a transformada de Laplace existe. □

A seguir estão listadas propriedades importantes da transformada de Laplace.

Propriedade 2.4.4 *Sejam f e g funções admissíveis e $a,b \in \mathbb{R}$. Sabemos que $\mathfrak{L}[f] = F(s)$ e $\mathfrak{L}[g] = G(s)$, então $\mathfrak{L}[af + bg] = aF(s) + bG(s)$.* **Demonstração:**

$$\begin{aligned}
\mathfrak{L}[af(t)+bg(t)] &= \int_0^{\infty} e^{-st}(af(t)+bg(t))dt\\
&= \int_0^{\infty} e^{-st}af(t)dt + \int_0^{\infty} e^{-st}bg(t)dt\\
&= a\int_0^{\infty} e^{-st}f(t)dt + b\int_0^{\infty} e^{-st}g(t)dt\\
&= aF(s)+bG(s).
\end{aligned}$$

□

Propriedade 2.4.5 *Se $\mathfrak{L}[f] = F(s)$, então $\mathfrak{L}[e^{at}f(t)] = F(s-a)$.* **Demonstração:**

$$\begin{aligned}
\mathfrak{L}[e^{at}f(t)] &= \int_0^{\infty} e^{-st}(e^{at}f(t))dt\\
&= \int_0^{\infty} e^{(-s+a)t}f(t)dt\\
&= \int_0^{\infty} e^{-(s-a)t}f(t)dt\\
&= F(s-a).
\end{aligned}$$

□

No que segue, algumas demonstrações serão omitidas por fugirem do interesse desse trabalho, mas podem ser encontradas em [15].

Propriedade 2.4.6 *Se f for uma função admissível, então* $\mathfrak{L}[t^n f(t)] = (-1)^n \dfrac{d^n}{ds^n}\mathfrak{L}[f(t)]$.

Propriedade 2.4.7 *Sejam f e f' funções integráveis em $[0,b]$, para todo $b > 0$. Se f for de ordem exponencial, então f' é admissível e existe $\mathfrak{L}[f'] = s\mathfrak{L}[f] - f(0)$.*

Propriedade 2.4.8 *Sejam $f, f', ..., f^{(n-1)}$ contínuas em $[0,\infty)$ e de ordem exponencial. Se $f^{(n)}$ for contínua por partes em $[0,\infty)$, então $f^{(n)}$ é admissível e*

$$\mathfrak{L}[f^{(n)}] = s^n F(s) - s^{(n-1)}f(0) - s^{(n-2)}f'(0) - ... - f^{(n-1)}(0),$$

onde onde $F(s) = \mathfrak{L}[f]$.

Em geral, a transformada de Laplace do produto de duas funções não é o produto das transformadas. Porém a seguir introduziremos o conceito de produto de convolução, que é um produto conveniente para que esta propriedade seja válida, ou seja, a transformada do produto de convolução é o produto das transformadas.

Definição 2.4.9 *Sejam f e g duas funções de ordem exponencial α e β e com transformadas de Laplace $F(s)$ e $G(s)$, respectivamente. Define-se a convolução de f e g, denotada por $(f * g)(t)$, como sendo*

$$(f * g)(t) = \int_0^t f(t-y)g(y)dy.$$

Propriedade 2.4.10 *Se $f * g$ representa a convolução das funções $f, g : [0,\infty) \to \mathbb{R}$, então $\mathfrak{L}[(f * g)] = \mathfrak{L}[f]\mathfrak{L}[g]$.*

Note que vale a igualdade,

$$\int_0^t f(t-y)g(y)dy = \int_0^t f(y)g(t-y)dy.$$

Agora, para calcularmos a transformada de Laplace de um produto de convolução, usaremos o teorema de Fubini (Veja 2.2.12).

$$\begin{aligned}
\mathfrak{L}[(f * g)(t)] &= \int_0^\infty e^{-st} \int_0^t f(t-y)g(y)dydt \\
&= \int_0^\infty \int_y^\infty e^{-st} f(t-y)g(y)dtdy
\end{aligned}$$

Para $x = t - y$ então $dt = dx$, tem-se:

$$\begin{aligned}
\mathfrak{L}[(f * g)(t)] &= \int_0^\infty \int_0^\infty e^{-sx} e^{-sy} f(x) g(y) dx dy \\
&= \int_0^\infty e^{-sx} f(x) dx \int_0^\infty e^{-sy} g(y) dy \\
&= \mathfrak{L}[f(t)] \mathfrak{L}[g(t)].
\end{aligned}$$

Portanto a transformada de Laplace do produto de convolução é o produto das transformadas.

Propriedade 2.4.11 $\mathfrak{L}[\cos(at)] = \dfrac{s}{s^2 + a^2}$.

$$\begin{aligned}
\mathfrak{L}[\cos(at)] &= \int_0^\infty e^{-st} \cos(at) dt \\
&= \lim_{b \to \infty} \int_0^b e^{-st} \cos(at) dt
\end{aligned}$$

Tomando $u = \cos(at), dv = e^{-st} dt$ para calcular a integral $\int e^{-st} \cos(at) dt$, tem-se:

$$\begin{aligned}
\int e^{-st} \cos(at) dt &= cos(at) \frac{e^{-st}}{-s} - \int \frac{e^{-st}}{-s} (-asen(at)) dt \\
&= -cos(at) \frac{e^{-st}}{s} - \frac{a}{s} \int e^{-st} sen(at) dt
\end{aligned}$$

Novamente, usando a substituição $u = sen(at), dv = e^{-st} dt$, tem-se:

$$\begin{aligned}
\int e^{-st} \cos(at) dt &= -cos(at) \frac{e^{-st}}{s} - \left(\frac{a}{s}\right) \left(-\frac{e^{-st}}{s} sen(at) - \int \frac{e^{-st}}{-s} a \cos(at) dt\right) \\
&= -cos(at) \frac{e^{-st}}{s} + \frac{a}{s^2} e^{-st} sen(at) - \frac{a^2}{s^2} \int e^{-st} \cos(at) dt
\end{aligned}$$

Logo,

$$\begin{aligned}
\left(1 + \frac{a^2}{s^2}\right) \int e^{-st} \cos(at) dt &= -cos(at) \frac{e^{-st}}{s} + \frac{a}{s^2} e^{-st} sen(at) \\
\int e^{-st} \cos(at) dt &= \left(\frac{s^2}{s^2 + a^2}\right) \left(-cos(at) \frac{e^{-st}}{s} + \frac{a}{s^2} e^{-st} sen(at)\right)
\end{aligned}$$

Voltando ao início,

$$
\begin{aligned}
\mathfrak{L}[\cos(at)] &= \lim_{b\to\infty}\int_0^b e^{-st}\cos(at)\\
&= \lim_{b\to\infty}\left(\frac{s^2}{s^2+a^2}\right)\left(-cos(at)\frac{e^{-st}}{s}+\frac{a}{s^2}e^{-st}sen(at)\right)|_0^b\\
&= \lim_{b\to\infty}\left(\frac{s^2}{s^2+a^2}\right)\left(-cos(ab)\frac{e^{-sb}}{s}+\frac{a}{s^2}e^{-sb}sen(ab)+\frac{1}{s}-0\right)\\
&= \left(\frac{s^2}{s^2+a^2}\right)\left(\frac{1}{s}\right)\\
&= \frac{s}{s^2+a^2}
\end{aligned}
$$

Propriedade 2.4.12 $\mathfrak{L}[sen(at)] = \dfrac{a}{s^2+a^2}$

A demonstração é análoga à anterior.

Em um trabalho de 1911, Poincaré discute a teoria dos *quantum* onde é incluída a fórmula da inversão da transformada de Laplace na forma usual de hoje em dia, embora B. Riemann a tivesse obtido anteriormente de uma forma um pouco diferente da atual, que é dada pela integral de Merllin,

$$\frac{1}{2\pi i}\lim_{b\to\infty}\int_{\lambda-ib}^{\lambda+ib} e^{st}F(s)ds.$$

Onde a integração é feito ao longo da linha vertical $Re(s)=\gamma$ no plano complexo em que γ é menor que a parte real de F(s).

Definição 2.4.13 *Seja $F(s)=\mathfrak{L}[f(t)]$ a transformada de Laplace da função $f(t)$. Então a transformada inversa é da forma $\mathfrak{L}^{-1}[F(s)]=f(t)$.*

Como M. Lerch estabeleceu, em 1892, a unicidade de transformadas de Laplace de funções contínuas, com base no Teorema de Aproximação de Weierstrass, que diz que qualquer função continua $f:[a,b]\to\mathbb{R}$ pode ser aproximada por funções polinomiais, o uso de tabelas para a inversão de transformadas de Laplace foi legitimado [4]. Entretanto, nem toda função possui transformada de Laplace inversa tabelada e existem, por exemplo, métodos numéricos e diversos trabalhos a respeito da inversão da transformada de Laplace,

[4] O trabalho de Lerch foi publicado em checo, ficando praticamente desconhecido até 1903 quando foi incluído num artigo publicado em francês. Este resultado, apareceu frequentemente citado como teorema de Lerch em textos de aplicações da transformada de Laplace para sustentar a utilização de tabelas de transformadas de Laplace.

uma vez que a fórmula geral de inversão é difícil e envolve a teoria de funçoes de variáveis complexas. Um exemplo disso é o trabalho [16], de 2010, que usa a teoria de espaços de Hilbert de reprodução para a inversão númerica.

Dessa forma, podemos escrever:

$$\begin{aligned}
\mathfrak{L}[t] &= \frac{1}{s^2} \quad \Rightarrow \quad \mathfrak{L}^{-1}\left[\frac{1}{s^2}\right] = t \\
\mathfrak{L}[cos(at)] &= \frac{s}{s^2+a^2} \quad \Rightarrow \quad \mathfrak{L}^{-1}\left[\frac{s}{s^2+a^2}\right] = cos(at) \\
\mathfrak{L}[sen(at)] &= \frac{a}{s^2+a^2} \quad \Rightarrow \quad \mathfrak{L}^{-1}\left[\frac{a}{s^2+a^2}\right] = sen(at) \\
\mathfrak{L}[tcos(at)] &= \frac{s^2-a^2}{(s^2+a^2)^2} \quad \Rightarrow \quad \mathfrak{L}^{-1}\left[\frac{s^2-a^2}{(s^2+a^2)^2}\right] = tcos(at) \\
\mathfrak{L}[tsen(at)] &= \frac{2as}{(s^2+a^2)^2} \quad \Rightarrow \quad \mathfrak{L}^{-1}\left[\frac{2as}{(s^2+a^2)^2}\right] = tsen(at)
\end{aligned}$$

2.4.2 Problema de Valor Inicial

Em 1833, J. Petzval desenvolveu consideravelmente a transformada de Laplace para resolução de equações diferenciais, publicando seu primeiro artigo sobre o assunto em 1847.

Vamos usar a transformada de Laplace para resolver a equação diferencial

$$f''(t) + 16f(t) = cos(4t),$$

onde $f(0) = 0$ e $f'(0) = 1$.

Aplicando a transformada de Laplace, obtemos:

$$\mathfrak{L}[f''(t) + 16f(t)] = \mathfrak{L}[cos(4t)]$$

$$\Rightarrow \mathfrak{L}[f''(t)] + \mathfrak{L}[16f(t)] = \mathfrak{L}[cos(4t)]$$

$$\Rightarrow s^2\mathfrak{L}[f(t)] - sf(0) - 1f'(0) + 16\mathfrak{L}[f(t)] = \frac{s}{s^2+16}$$

$$\Rightarrow (s^2+16)F(s) = \frac{s}{s^2+16} + 1$$

$$\Rightarrow F(s) = \frac{s}{(s^2+16)^2} + \frac{1}{s^2+16}$$

Aplicando a transformada de Laplace inversa, temos

$$\begin{aligned} f(t) &= \mathfrak{L}^{-1}\left[\frac{s}{(s^2+16)^2}\right] + \mathfrak{L}^{-1}\left[\frac{1}{s^2+16}\right] \\ &= \frac{1}{8}\mathfrak{L}^{-1}\left[\frac{8s}{(s^2+16)^2}\right] + \frac{1}{4}\mathfrak{L}^{-1}\left[\frac{4}{s^2+16}\right] \\ &= \frac{1}{8}tsen(4t) + \frac{1}{4}sen(4t) \end{aligned}$$

Agora, vamos usar a transformada de Laplace na resolução de um sistema linear de equações diferenciais. Considere o problema de valor inicial dado por:

$$\begin{cases} f_1'(t) = & -3f_1(t) + 2f_2(t) \\ f_2'(t) = & -4f_1(t) + f_2(t) + 2sent(t) \end{cases}$$

Com,

$$\begin{cases} f_1(0) = & 0 \\ f_2(0) = & 0. \end{cases}$$

Aplicando a transformada de Laplace às equações, obtemos

$$\begin{cases} \mathfrak{L}[f_1'(t)] = & \mathfrak{L}[-3f_1(t) + 2f_2(t)] \\ \mathfrak{L}[f_2'(t)] = & \mathfrak{L}[-4f_1(t) + f_2(t) + 2sent(t)]. \end{cases}$$

Como a transformada de Laplace é um operador linear e usando a propriedade 2.4.7 podemos escrever:

$$\begin{cases} s\mathfrak{L}[f_1(t)] - f_1(0) = & -3\mathfrak{L}[f_1(t)] + 2\mathfrak{L}[f_2(t)] \\ s\mathfrak{L}[f_2(t)] - f_2(0) = & -4\mathfrak{L}[f_1(t)] + \mathfrak{L}[f_2(t)] + 2\mathfrak{L}[sent(t)]. \end{cases}$$

Como o sistema está sujeito as condições iniciais, $f_1(0) =$ e $f_2(0) = 0$, temos,

$$\begin{cases} s\mathfrak{L}[f_1(t)] = & -3\mathfrak{L}[f_1(t)] + 2\mathfrak{L}[f_2(t)] \\ s\mathfrak{L}[f_2(t)] = & -4\mathfrak{L}[f_1(t)] + \mathfrak{L}[f_2(t)] + \dfrac{2}{s^2+1}. \end{cases}$$

Com algumas manipulações algébricas, obtemos:

$$\begin{cases} \mathfrak{L}[f_1(t)] = & \dfrac{4}{(s^2+1)(2s+5)} \\ \\ \mathfrak{L}[f_2(t)] = & \dfrac{2(s+3)}{(s^2+1)(s^2+2s+5)}. \end{cases}$$

Usando frações parciais do lado direito das igualdades, chagamos a:

$$\begin{cases} \mathfrak{L}[f_1(t)] = & \dfrac{1}{5}\left(\dfrac{-2s+4}{s^2+1}\right) + \dfrac{2}{5}\left(\dfrac{s}{s^2+2s+5}\right) \\ \\ \mathfrak{L}[f_2(t)] = & \dfrac{-1}{5}\left(\dfrac{s-7}{s^2+1}\right) + \dfrac{1}{5}\left(\dfrac{s-5}{s^2+2s+5}\right). \end{cases}$$

Aplicando a transformada de Laplace inversa,

$$\begin{cases} f_1(t) = & \dfrac{-2}{5}\mathfrak{L}^{-1}\left[\dfrac{s}{s^2+1}\right] + \dfrac{4}{5}\mathfrak{L}^{-1}\left[\dfrac{1}{s^2+1}\right] + \dfrac{2}{5}\mathfrak{L}^{-1}\left[\dfrac{s+1-1}{(s+1)^2+4}\right] \\ \\ f_2(t) = & \dfrac{-1}{5}\mathfrak{L}^{-1}\left[\dfrac{s}{s^2+1}\right] + \dfrac{7}{5}\mathfrak{L}^{-1}\left[\dfrac{1}{s^2+1}\right] + \dfrac{1}{5}\mathfrak{L}^{-1}\left[\dfrac{s-5+1-1}{(s+1)^2+4}\right]. \end{cases}$$

Dessa forma,

$$\begin{cases} f_1(t) = & \dfrac{-2}{5}\mathfrak{L}^{-1}\left[\dfrac{s}{s^2+1}\right] + \dfrac{4}{5}\mathfrak{L}^{-1}\left[\dfrac{1}{s^2+1}\right] + \dfrac{2}{5}\mathfrak{L}^{-1}\left[\dfrac{s+1}{(s+1)^2+4}\right] - \dfrac{1}{5}\mathfrak{L}^{-1}\left[\dfrac{2}{(s+1)^2+4}\right] \\ \\ f_2(t) = & \dfrac{-1}{5}\mathfrak{L}^{-1}\left[\dfrac{s}{s^2+1}\right] + \dfrac{7}{5}\mathfrak{L}^{-1}\left[\dfrac{1}{s^2+1}\right] + \dfrac{1}{5}\mathfrak{L}^{-1}\left[\dfrac{s+1}{(s+1)^2+4}\right] - \dfrac{3}{5}\mathfrak{L}^{-1}\left[\dfrac{2}{(s+1)^2+4}\right]. \end{cases}$$

Portanto, a solução do problema de valor inicial é:

$$\begin{cases} f_1(t) = & \dfrac{-2}{5}cos(t) + \dfrac{4}{5}sen(t) + \dfrac{2}{5}e^{-t}cos(2t) - \dfrac{1}{5}e^{-t}sen(2t) \\ \\ f_2(t) = & \dfrac{-1}{5}cos(t) + \dfrac{7}{5}sen(t) + \dfrac{1}{5}e^{-t}cos(2t) - \dfrac{3}{5}e^{-t}sen(2t). \end{cases}$$

2.4.3 Equação Integral

Agora, vamos resolver uma equação integral de Volterra usando a transformada de Laplace. No capítulo (5) as equações integrais de Volterra serão abordadas mais detalhadamente.

Seja $f(t) = 3t^2 - e^{-t} - \int_0^t f(x)e^{t-x}dx$.

Note que podemos escrever a integral acima como um produto de convolução, onde $g(t) = e^t$. Dessa forma, reescrevendo $f(t)$, obtemos

$$f(t) = 3t^2 - e^{-t} - (f * g)(t).$$

Aplicando a transformada de Laplace, temos

$$\begin{aligned}
F(s) &= 3\mathfrak{L}[t^2] - \mathfrak{L}[e^{-t}] - \mathfrak{L}[f(t)]\mathfrak{L}[g(t)] \\
\Rightarrow F(s) &= 3\mathfrak{L}[t^2] - \mathfrak{L}[e^{-t}] - \mathfrak{L}[f(t)]\mathfrak{L}[e^t] \\
\Rightarrow F(s) &= 3\left(\frac{2}{s^3}\right) - \frac{1}{s+1} - F(s)\left(\frac{1}{s-1}\right).
\end{aligned}$$

Assim,

$$\begin{aligned}
F(s) + F(s)\left(\frac{1}{s-1}\right) &= \left(\frac{6}{s^3}\right) - \frac{1}{s+1}. \\
\Rightarrow F(s)\left(1 + \frac{1}{s-1}\right) &= \left(\frac{6}{s^3}\right) - \frac{1}{s+1}. \\
\Rightarrow F(s)\left(\frac{s}{s-1}\right) &= \left(\frac{6}{s^3}\right) - \frac{1}{s+1}. \\
\Rightarrow F(s) &= \frac{6(s-1)}{s^4} - \frac{s-1}{s(s+1)}.
\end{aligned}$$

Usando frações parciais, obtemos

$$F(s) = \frac{6}{s^3} - \frac{6}{s^4} + \frac{1}{s} - \frac{2}{s+1}.$$

Dessa forma, aplicando transformada de Laplace inversa

$$\begin{aligned} f(t) &= \mathfrak{L}^{-1}\left[\frac{6}{s^3}\right] - \mathfrak{L}^{-1}\left[\frac{6}{s^4}\right] + \mathfrak{L}^{-1}\left[\frac{1}{s}\right] - \mathfrak{L}^{-1}\left[\frac{2}{s+1}\right] \\ &= 3\mathfrak{L}^{-1}\left[\frac{2!}{s^3}\right] - \mathfrak{L}^{-1}\left[\frac{3!}{s^4}\right] + \mathfrak{L}^{-1}\left[\frac{1}{s}\right] - 2\mathfrak{L}^{-1}\left[\frac{1}{s+1}\right] \\ &= 3t^2 - t^3 + 1 - 2e^{-t}. \end{aligned}$$

A transformada de Laplace também pode ser usada na resolução de sistemas de equações integrais. Algumas equações integrais de Volterra, que são o objeto de estudo desse trabalho, podem ser resolvidas com essa técnica. Porém é necessário que as integrais de Volterra possam ser escritas como um produto de convolução e, é claro, que existam as transformadas de Laplace de cada função e suas respectivas transformadas de Laplace inversa. Como exemplo, considere o sistema de equações integrais dado por

$$\begin{cases} f_1(t) = & 1 - 2\displaystyle\int_0^t e^{2(t-x)} f_1(x)dx + \int_0^t f_2(x)dx \\ \\ f_2(t) = & 3t - \displaystyle\int_0^t f_1(t)dt + 4\int_0^t (t-x) f_2(x)dx. \end{cases}$$

Podemos escrever as integrais acima usando o produto de convolução. Considere $g(t) = e^{2t}$, $h(t) = 1$ e $y(t) = t$, assim

$$\begin{cases} f_1(t) = & 1 - 2(f_1 * g)(t) + (f_2 * h)(t) \\ f_2(t) = & 3t - (f_1 * h)(t) + 4(f_2 * y)(t). \end{cases}$$

Aplicando a transformada de Laplace, obtemos

$$\begin{cases} \mathfrak{L}[f_1(t)] = & \mathfrak{L}[1] - 2\mathfrak{L}[f_1(t)]\mathfrak{L}[g(t)] + \mathfrak{L}[f_2(t)]\mathfrak{L}[h(t)] \\ \mathfrak{L}[f_2(t)] = & 3\mathfrak{L}[t] - \mathfrak{L}[f_1(t)]\mathfrak{L}[h(t)] + 4\mathfrak{L}[f_2(t)]\mathfrak{L}[y(t)]. \end{cases}$$

Dessa forma, temos

$$\begin{cases} \mathfrak{L}[f_1(t)] = & \dfrac{1}{s} - \dfrac{2}{s-2}\mathfrak{L}[f_1(t)] + \dfrac{1}{s}\mathfrak{L}[f_2(t)] \\ \\ \mathfrak{L}[f_2(t)] = & \dfrac{3}{s} - \dfrac{1}{s}\mathfrak{L}[f_1(t)] + \dfrac{4}{s^2}\mathfrak{L}[f_2(t)]. \end{cases}$$

Com algumas manipulações algébricas, chegamos a

$$\begin{cases} \mathfrak{L}[f_1(t)] = & \dfrac{s-4}{s(s+1)} \\ \\ \mathfrak{L}[f_2(t)] = & \dfrac{3s^2+2s-s+4}{s^2(s+1)}. \end{cases}$$

Agora, usando frações parciais obtemos

$$\begin{cases} \mathfrak{L}[f_1(t)] = & -\dfrac{4}{s} + \dfrac{5}{s+1} \\ \\ \mathfrak{L}[f_2(t)] = & -\dfrac{2}{s} + \dfrac{4}{s^2} + \dfrac{5}{s+1}. \end{cases}$$

Aplicando a transformada de Laplace inversa, temos

$$\begin{cases} f_1(t) = & -4\mathfrak{L}^{-1}\left[\dfrac{1}{s}\right] + 5\mathfrak{L}^{-1}\left[\dfrac{1}{s-(-1)}\right] \\ \\ f_2(t) = & -2\mathfrak{L}^{-1}\left[\dfrac{1}{s}\right] + 4\mathfrak{L}^{-1}\left[\dfrac{1}{s^2}\right] + 5\mathfrak{L}^{-1}\left[\dfrac{1}{s+1}\right]. \end{cases}$$

Portando, a solução do sistema de equações dado

$$\begin{cases} f_1(t) = & -4+5e^{-t} \\ f_2(t) = & -2+4t+5e^{-t}. \end{cases}$$

O estudo de resolução de equações usando a transformada de Laplace exige que exista a transformada de Laplace inversa. Dessa forma, para simulações numéricas, são necessários métodos que aproximem tais funções. O uso da teoria de núcleos positivos definidos para esse fim tem se mostrado satisfatório e pode ser visto com mais detalhes em [16].

3 NÚCLEOS POSITIVOS DEFINIDOS

Os núcleos positivos definidos têm um papel crescente em diversas aplicações, tais como soluções numéricas de equações diferenciais e gráficos e experimentos computacionais, também são úteis em problemas de interpolação. Em trabalhos recentes, Michael Scheuerer, Robert Schaback e Martin Schlather questionam se a interpolação de dados espaciais é um processo estocástico ou um problema determinístico que pode ser resolvido com o uso da teoria de núcleos positivos definidos e espaços de reprodução [12].

Ainda segundo [12], o início do estudo de funções positivas definidas foi creditado a James Mercer e a Maximilian Mathias, que era aluno de Erhard Schmidt. Como estavam interessados em analisar estas funções no contexto de transformada de Fourier, eles não perceberam que Mercer, cerca de uma década antes, havia considerado os núcleos positivos definidos no seu trabalho sobre equações integrais. A contribuição mais significativa de funções positivas definidas em termos de transformada de Fourier foi feita por Salomon Bochner e Iso Schoenberg. O teorema conhecido como teorema de Bochner foi usado por Aleksander Khinchin, por volta de 1930, para estabelecer as bases para o estudo de processos estocásticos estacionários em teoria da probabilidade.

O conceito de espaços de reprodução foi introduzido por Nachman Aronszajn e Stefan Bergman, por volta de 1950. Moore mostrou que cada núcleo positivo definido gera um espaço que pode ser completado formando um espaço de Hilbert de reprodução, ligando as duas teorias (Veja [2]).

3.1 NÚCLEOS POSITIVOS DEFINIDOS

Muitos problemas são descritos por equações integrais, que muitas vezes não podem ser resolvidos com as técnicas usuais. A teoria de núcleos positivos definidos nos proporciona uma técnica para a resolução de tais problemas, logo, vemos a importância dessa teoria.

Neste capítulo, discutiremos algumas definições e propriedades referentes aos núcleos positivos definidos, assim como faremos alguns exemplos.

Um núcleo é uma função de duas variáveis que possui contradomínio pertencente ao conjunto dos números complexos. Iniciaremos com uma definição formal de núcleo.

Definição 3.1.1 *Seja X um conjunto não vazio. Um núcleo K é uma função $K : X \times X \to \mathbb{C}$.*

Exemplo 3.1.2 *Seja X um conjunto de pontos de $\mathbb{R}^2$, então $K : X \times X \to \mathbb{C}$ dada por $K((x_1,y_1),(x_2,y_2)) = 3x_1y_2 + ix_2y_1$ é um núcleo.*

Um conceito importante para o nosso estudo é o de núcleo positivo definido. Para isso, precisamos conhecer a definição de matriz não negativa definida.

Definição 3.1.3 *Uma matriz $A_{n\times n}(\mathbb{C})$ é não negativa definida quando a forma quadrática*

$$B(y) = \overline{y}Ay^t, \qquad y \in \mathbb{C}^n,$$

for não negativa, ou seja,

$$\overline{y}Ay^t \geq 0, \qquad y \in \mathbb{C}^n.$$

A definição anterior é usada, por alguns autores, para definir matrizes positivas definidas. No nosso caso, usaremos essa nomeação quando a desigualdade for estritamente positiva.

Exemplo 3.1.4 *A matriz A dada por:*

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix},$$

é não negativa definida.

De fato,

$$\overline{y}Ay^t \;=\; \overline{(y_1,y_2,y_3)} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Dessa forma,

$$
\begin{aligned}
\overline{y}Ay^t &= \overline{(y_1,y_2,y_3)} \begin{pmatrix} 2y_1 - y_2 \\ -y_1 + 2y_2 - y_3 \\ -y_2 + 2y_3 \end{pmatrix} \\
&= 2y_1\overline{y_1} - (\overline{y_1}y_2 + y_1\overline{y_2}) + 2y_2\overline{y_2} - (\overline{y_2}y_3 + y_2\overline{y_3}) + 2y_3\overline{y_3} \\
&= 2y_1\overline{y_1} - 2Re(y_1\overline{y_2}) + 2y_2\overline{y_2} - 2Re(\overline{y_2}y_3) + 2y_3\overline{y_3}
\end{aligned}
$$

Fazendo $y_j = a_j + ib_j$, com $a_j,b_j \in \mathbb{R}$ e $j = 1,2,3$, temos:

$$
\begin{aligned}
\overline{y}Ay^t &= 2(a_1 + ib_1)(a_1 - ib_1) - 2Re((a_1 + ib_1)(a_2 - ib_2)) + 2(a_2 + ib + 2)(a_2 - ib_2) - \\
&- 2Re((a_2 - ib_2)(a_3 + ib_3)) + 2(a_3 + ib_3)(a_3 - ib_3) \\
&= 2a_1^2 + 2b_1^2 - 2a_1a_2 - 2b_1b_2 + 2a_2^2 + 2b_2^2 - 2a_2a_3 - 2b_2b_3 + 2a_3^2 + 2b_3^2 \\
&= (a_1 - a_2)^2 + (a_2 - a_3)^2 + (b_1 - b_2)^2 + (b_2 - b_3)^2 + a_1^2 + a_3^2 + b_1^2 + b_3^2 \\
&\geq 0
\end{aligned}
$$

Logo, a matriz A é não negativa definida.

A partir do conceito de matriz não negativa definida, podemos introduzir a ideia de núcleo positivo definido, que será o objeto de nossos estudos.

Definição 3.1.5 *Seja X um conjunto não vazio. Dizemos que um núcleo $K : X \times X \to \mathbb{C}$ é positivo definido quando a matriz $A = (K(x_i,x_j))$, de ordem n, é não negativa definida para qualquer $n \geq 1$ e qualquer n-úpla $(x_1,x_2, ..., x_n) \in X^n$.*

Denotaremos por $PD(X)$ o conjunto dos núcleos positivos definidos com domínio $X \times X$.

Observação 3.1.6 *Note que a definição (3.1.5) é equivalente a seguinte desigualdade ser satisfeita:*

$$\sum_{i,j=1}^{n} \overline{c_i}c_jK(x_i,x_j) \geq 0 \tag{3.1.1}$$

Dizemos então que K é estritamente positivo definido caso a desigualdade acima seja positiva sempre que algum c_i for não nulo.

De fato,

$$\begin{aligned}
\overline{c}K(x_i,x_j)c^t &= \overline{(c_1,c_2,...,c_n)} \begin{pmatrix} K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n) \\ K(x_2,x_1) & K(x_2,x_2) & \cdots & K(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\
&= \sum_{j=1}^{n} \overline{c_1} c_j K(x_1,x_j) + ... + \sum_{j=1}^{n} \overline{c_n} c_j K(x_n,x_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{c_i} c_j K(x_i,x_j) \\
&= \sum_{i,j=1}^{n} \overline{c_i} c_j K(x_i,x_j)
\end{aligned}$$

Observação 3.1.7 *Seja X um conjunto não vazio, $f : X \to \mathbb{C}$ e $K : X \times X \to \mathbb{C}$ um núcleo. Dados $x_i \in X$, $i = 1, \ldots, n$, e tendo valores tabelados $y_i = f(x_i)$. Um problema de interpolação pode ser escrito então como,*

$$S_f(x_j) = f(x_j) = \sum_{i=1}^{n} c_i K(x_j,x_i), \quad j = 1, \ldots, n,$$

quando tentamos aproximar f por uma função

$$S_f(x) = \sum_{i=1}^{n} c_i K(x,x_i), \quad x \in X,$$

e por isso queremos determinar os valores de c_i. Caso o núcleo K seja estritamente positivo definido, podemos garantir que o problema tem solução única e podemos usar diversos métodos numéricos para a sua resolução. Como caso particular podemos resolver o sistema por eliminação de Gauss sem a necessidade de usar pivotamento ou o método iterativo de Gauss-Seidel (Veja [13, p.119]).

A seguir, faremos alguns exemplos de núcleos positivos definidos.

Exemplo 3.1.8 *Seja $K : X \times X \to \mathbb{C}$ dada por:*

$$K(x,y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases}$$

Se $n \in \mathbb{N}$, $\{x_1,x_2,\ldots,x_n\}$ é um subconjunto de X e $\{c_1,c_2,\ldots,c_n\}$ são números complexos, então:

$$\sum_{i,j=1}^{n} c_i\overline{c_j}K(x_i,x_j) = \sum_{i=1}^{n} c_i\overline{c_i} = \sum_{i=1}^{n} |c_i|^2 \geq 0.$$

Logo, K é um núcleo positivo definido.

Exemplo 3.1.9 *Se $f : X \to \mathbb{C}$ é uma função qualquer, o núcleo dado pela expressão $K(x,y) = f(x)\overline{f(y)}$, $\forall x,y \in X$, é positivo definido.*

De fato, seja $n \in \mathbb{N}$, $\{x_1,x_2,\ldots,x_n\}$ um subconjunto de X e $\{c_1,c_2,\ldots,c_n\}$ números complexos, então:

$$\begin{aligned}
\sum_{i,j=1}^{n} \overline{c_i}c_jK(x_i,x_j) &= \sum_{i,j=1}^{n} \overline{c_i}c_jf(x_i)\overline{f(x_j)} \\
&= \sum_{i,j=1}^{n} \overline{c_i}f(x_i)\overline{\overline{c_j}f(x_j)} \\
&= \left|\sum_{j=1}^{n} \overline{c_j}f(x_j)\right|^2 \\
&\geq 0
\end{aligned}$$

Logo, $K(x,y) = f(x)\overline{f(y)}$ é um núcleo positivo definido.

Exemplo 3.1.10 *Sejam (X,μ) um espaço da medida e $f : X \times X \to \mathbb{C}$ tal que $f(\cdot,x) \in \mathcal{L}^2$, para todo $x \in X$. Nestas condições, o núcleo K dado por:*

$$K(x,y) = \int_X f(z,y)\overline{f(z,x)}d\mu(z)$$

é positivo definido.

De fato,

$$\begin{aligned}
\sum_{i,j=1}^{n} \overline{c_i}c_jK(x_i,x_j) &= \sum_{i,j=1}^{n} \overline{c_i}c_j \int_X f(z,x_i)\overline{f(z,x_j)}d\mu(z) \\
&= \int_X \sum_{i,j=1}^{n} \overline{c_i}f(z,x_i)\overline{\overline{c_j}f(z,x_j)}d\mu(z) \\
&= \int_X \left|\sum_{j=1}^{n} c_jf(z,x_j)\right|^2 d\mu(z) \\
&\geq 0
\end{aligned}$$

Garantindo que esta soma é sempre não negativa podemos concluir que o núcleo apresentado é positivo definido.

O exemplo a seguir será usado como motivação para o lema (3.1.12).

Exemplo 3.1.11 *Se X é um espaço vetorial complexo com produto interno $\langle \cdot,\cdot \rangle_X$, então $K(x,y) = \langle x,y \rangle_X$, com $x,y \in X$ é positivo definido.*

De fato,

$$\begin{aligned}
\sum_{i,j=1}^{n} \overline{c_i} c_j K(x_i,x_j) &= \sum_{i,j=1}^{n} \overline{c_i} c_j \langle x_i,x_j \rangle_X \\
&= \sum_{i,j=1}^{n} \overline{c_i} c_j \overline{\langle x_j,x_i \rangle_X} \\
&= \overline{\left\langle \sum_{j=1}^{n} c_j x_j, \sum_{i=1}^{n} c_i x_i \right\rangle_X} \\
&= \left\langle \sum_{i=1}^{n} c_i x_i, \sum_{j=1}^{n} c_j x_j \right\rangle_X \\
&= \left\| \sum_{i=1}^{n} c_i x_i \right\|^2 \\
&\geq 0
\end{aligned}$$

Dessa forma, podemos generalizar esta ideia com o seguinte lema.

Lema 3.1.12 *Sejam $\mathcal{H}$ um espaço de Hilbert, X um conjunto não vazio e $g : X \to \mathcal{H}$ uma função qualquer.*

a) O núcleo $K : X \times X \to \mathcal{H}$ dado por $K(x,y) = \langle g(y),g(x) \rangle_{\mathcal{H}}$, $\forall x,y \in X$ é positivo definido;

b) O núcleo $K_1 : X \times X \to \mathcal{H}$ dado por $K_1(x,y) = \langle g(x),g(y) \rangle_{\mathcal{H}}$, $\forall x,y \in X$ é positivo definido.

Demonstração:

a) Basta observar que:

$$\begin{aligned}
\sum_{i,j=1}^{n} \overline{c_j} c_i K(x_j,x_i) &= \sum_{i,j=1}^{n} \overline{c_j} c_i \langle g(x_i), g(x_j)\rangle_{\mathcal{H}} \\
&= \left\langle \sum_{i=1}^{n} c_i g(x_i), \sum_{j=1}^{n} c_j g(x_j) \right\rangle_{\mathcal{H}} \\
&= \left\| \sum_{i=1}^{n} c_i g(x_i) \right\|^2 \\
&\geq 0
\end{aligned}$$

b) Note que:

$$\begin{aligned}
K_1(x,y) &= \langle g(x), g(y)\rangle \\
&= \overline{\langle g(y), g(x)\rangle} \\
&= \overline{K(x,y)}
\end{aligned}$$

Logo, K_1 também é um núcleo positivo definido. $\square$

Para a próxima proposição são necessários os conceitos de matriz autoadjunta e autovalores. Com esses dois conceitos garantimos que o determinante de uma matriz não negativa definida é sempre positivo, mais ainda, o determinante da matriz autoadjunta é o produto de seus autovalores. Para fixar esses conceitos, primeiramente, vamos definir matriz adjunta e autoadjunta.

Definição 3.1.13 *Sejam H_1, H_2 espaços de Hilbert. Dado um operador linear contínuo $A : H_1 \to H_2$, existe um único operador linear contínuo $A^* : H_2 \to H_1$ tal que:*

$$\langle Ax, y\rangle_{H_2} = \langle x, A^* y\rangle_{H_1}, \quad \forall x, y \in H.$$

O operador A^* é o adjunto de A (Veja [1, p.138]).

No caso $H_1 = H_2$, podemos definir um autoadjunto, ou seja, quando o operador coincide com seu adjunto. Para o caso de espaços euclidianos temos a seguinte definição:

Definição 3.1.14 *Uma matriz $A = (a_{ij}) \in \mathbb{M}_n(\mathbb{K})$ é autoadjunta quando:*

$$\langle Ax^t, y\rangle = \langle x, Ay^t\rangle, \quad \forall x,y \in \mathbb{K}^n.$$

Onde $\langle \cdot\, , \cdot \rangle$ *é o produto interno canônico em* $\mathbb{K}^n$

Autovalores e autovetores são bastante utilizados em problemas que envolvem sistemas dinâmicos, pois são necessários no estudo da estabilidade de soluções. A determinação de autovalores e autovetores de uma matriz são conceitos que merecem uma maior atenção por haver inúmeras aplicações práticas em várias áreas, como por exemplo na mecânica quântica, no processamento de imagem, na Estatística, etc. A definição a seguir traz a ideia de autovalores e autovetores.

Definição 3.1.15 *Dizemos que* $\lambda \in \mathbb{K}$ *e um vetor não nulo* $v \in \mathbb{K}^n$ *são respectivamente autovalor e autovetor de uma matriz* A *se* $Av^t = \lambda v^t$.

Observação 3.1.16 *Vale ressaltar que se* A *for uma matriz não negativa definida todos os autovalores de* A *são não negativos. Note que:*

$$0 \leq \overline{v}Av^t = \overline{v}\lambda v^t = \lambda \overline{v}v^t = \lambda \|v\|^2.$$

Ou seja, $\lambda \geq 0$.

Proposição 3.1.17 *Seja* $K \in PD(X)$, *então para quaisquer* $x, y \in X$, *as seguintes afirmações são verdadeiras.*

a) $K(x,x) \geq 0$;

b) K *é hermitiano, ou seja,* $K(x,y) = \overline{K(y,x)}$;

c) $|K(x,y)|^2 \leq K(x,x)K(y,y)$

Demonstração:

a) Por definição a matriz $A = [K(x_i,x_j)]$ é não negativa definida para todo $n \geq 1$. Em particular, para $n = 1$, concluímos que $K(x,x) \geq 0$.

b) A matriz $A = [K(x_i,x_j)]$ é não negativa definida $\forall n \geq 1$ e $\forall$ n-úpla $(x_1,x_2,...,x_n) \in X^n$, pois $K \in PD(X)$. Assim,

$$\overline{z}Az^t \geq 0.$$

Fazendo $n = 2$, $z = (1,1)$ e $K(x_i,x_j) = a_{ij}$, temos:

$$\begin{aligned} 0 &\leq \overline{(1,1)} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= a_{11} + a_{12} + a_{21} + a_{22}. \end{aligned}$$

Pelo item i., $K(x,x) \geq 0$, logo $a_{11},a_{22} \in \mathbb{R}$. Daí, concluímos que :

$$Im(a_{12}) = -Im(a_{21}). \tag{3.1.2}$$

Por outro lado, fazendo $z = (1,i)$, obtemos:

$$\begin{aligned} 0 &\leq \overline{(1,i)} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= a_{11} + ia_{12} - ia_{21} + a_{22}. \end{aligned}$$

Do mesmo modo, $a_{11},a_{22} \in \mathbb{R}$, então devemos ter:

$$Re(a_{12}) = Re(a_{21}) \tag{3.1.3}$$

A partir de (3.1.2) e (3.1.3) obtemos que $a_{12} = \overline{a_{21}}$. Ou seja, $K(x,y) = \overline{K(y,x)}$, $\forall x,y \in X$.

c) Note que A é autoadjunta, pois $A = \overline{A}^t$. Assim, A é diagonalizável (Veja [9, p.228]) e, pela Observação 3.1.16, possui todos os autovalores não negativos. Segue que seu determinante, que é o produto de seus autovalores, é não negativo (veja a Observação 3.1.18 a seguir). Com isso,

$$detA = det \begin{pmatrix} a_{11} & \overline{a_{21}} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - |a_{21}|^2 \geq 0.$$

Dessa forma,

$$|a_{21}|^2 \leq a_{11}a_{22}.$$

Isto é,

$$|k(y,x)|^2 \leq K(x,x)K(y,y), \quad \forall x,y \in X.$$

$\square$

Observação 3.1.18 *Sendo A uma matriz diagonalizável sabemos que A possui n autovetores. Logo, existe uma matriz P que é a matriz mudança de base da base canônica $\mathbb{K}^n$ para a base formada pelos autovetores de A. Também podemos encontrar a matriz diagonal D composta pelos n autovalores de A dispostos na diagonal principal. Como a matriz A é semelhante a matriz D, temos que:*

$$A = P^{-1}DP.$$

Dessa forma, podemos concluir que $det(A) = det(D)$, pois:

$$\begin{aligned} det(A) &= det(P^{-1}DP) \\ &= det(P^{-1})det(D)det(P) \\ &= \frac{1}{det(P)}det(D)det(P) \\ &= det(D). \end{aligned}$$

Como D é uma matriz diagonal, seu determinante é o produto dos valores da diagonal principal. Ou seja, quando A é uma matriz diagonalizável seu determinante é o produto de seus autovalores. Sendo assim, sabendo que toda matriz autoadjunta é diagonalizável, podemos concluir que o determinante de uma matriz autoadjunta é dado pelo produto de seus autovalores.

Teorema 3.1.19 (Unicidade da Raiz Quadrada) *Seja A uma matriz autoadjunta e não negativa definida. Então, existe uma única matriz P autoadjunta e não negativa definida tal que $A = P^2$. (Veja [1, p. 226])*

Teorema 3.1.20 *Uma matriz A é autoadjunta se, e somente se, $A = \overline{A}^T$.*

Demonstração: Sejam $A = (a_{ij})$ e $x,y \in \mathbb{K}^n$. Note que,

$$\begin{aligned} \langle Ax^T, y\rangle &= \left\langle \left(\sum_{i=1}^{n} a_{1i}x_i, \cdots, \sum_{i=1}^{n} a_{ni}x_i\right), (y_1, \cdots, y_n)\right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ji}x_i\overline{y_i} \\ &= \left\langle (x_1, \cdots, x_n), \left(\sum_{j=1}^{n} a_{j1}y_j, \cdots, \sum_{j=1}^{n} a_{jn}y_j\right)\right\rangle \\ &= \langle x, \overline{A}^T y^T\rangle \end{aligned}$$

Logo, A é autoadjunta se, e somente se, $A = \overline{A}^T$. □

Teorema 3.1.21 *Sejam $K : X \times X \to \mathbb{C}$ um núcleo positivo definido e a matriz $A = (K(x_i,x_j))$, $i,j = 1, \cdots, n$, então existe uma única matriz G tal que $A = G\overline{G}^T$*

Demonstração: Como K é um núcleo positivo definido, segue da proposição (3.1.5) que $A = \overline{A}^T$. Logo, pelo teorema (3.1.20), A é uma matriz autoadjunta e portanto existe uma única matriz G autoadjunta não negativa definida tal que $A = G^2$, conforme enunciado no teorema (3.1.19).

Como G é autoadjunta, temos que $G = \overline{G}^T$. Logo $A = G\overline{G}^T$. □

Teorema 3.1.22 *Sejam $K_1, \cdots, K_p \in PD(X)$ e $d_1, \cdots, d_p \geq 0$:*

i. A soma $\sum_{i=1}^{p} d_iK_i$ está em $PD(X)$;

a) O produto K_1K_2 está em $PD(x)$;

b) Se K_n converge para K então $K \in PD(X)$.

Demonstração: a) Seja $K = \sum_{i=1}^{p} d_iK_i$, então:

$$\begin{aligned} \sum_{i,j=1}^{n} c_i\overline{c_j}K(x_i,x_j) &= \sum_{i,j=1}^{n} c_i\overline{c_j}\sum_{l=1}^{p} d_lK_l(x_i,x_j) \\ &= \sum_{l=1}^{p} d_l \sum_{i,j=1}^{n} c_i\overline{c_j}K_l(x_i,x_j) \\ &\geq 0 \end{aligned}$$

b) Sejam $A = (a_{ij}) = (K_1(x_i,x_j))$ e $B = (b_{ij}) = (K_2(x_i,x_j))$. Como A é uma matriz não negativa definida, pelo teorema (3.1.20), existe uma única matriz $G = (g_{ij})$ autoadjunta tal que $A = G\overline{G}^T$. Assim,

$$\begin{aligned} A &= G\overline{G}^T \\ &= \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} \overline{g_{11}} & \overline{g_{12}} & \cdots & \overline{g_{1n}} \\ \overline{g_{21}} & \overline{g_{22}} & \cdots & \overline{g_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{g_{n1}} & \overline{g_{n2}} & \cdots & \overline{g_{nn}} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{p=1}^{n} g_{1p}\overline{g_{1p}} & \sum_{p=1}^{n} g_{1p}\overline{g_{2p}} & \cdots & \sum_{p=1}^{n} g_{1p}\overline{g_{np}} \\ \sum_{p=1}^{n} g_{2p}\overline{g_{1p}} & \sum_{p=1}^{n} g_{2p}\overline{g_{2p}} & \cdots & \sum_{p=1}^{n} g_{2p}\overline{g_{np}} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{p=1}^{n} g_{np}\overline{g_{1p}} & \sum_{p=1}^{n} g_{np}\overline{g_{2p}} & \cdots & \sum_{p=1}^{n} g_{np}\overline{g_{np}} \end{pmatrix} \end{aligned}$$

Dessa forma, temos que

$$a_{ij} = \sum_{p=1}^{n} g_{ip}\overline{g_{jp}}, \quad \text{para} \quad i,j = 1, \cdots, n.$$

Vamos mostrar que a matriz $C = (a_{ij}b_{ij})$ é não negativa definida, para isso, considere $c_1, \cdots, c_n \in \mathbb{C}$.

$$\sum_{i,j=1}^{n} c_i\overline{c_j}a_{ij}b_{ij} = \sum_{i,j=1}^{n}\sum_{p=1}^{n} c_i\overline{c_j}g_{ip}\overline{g_{jp}}b_{ij} = \sum_{i,j=1}^{n}\sum_{p=1}^{n} c_i g_{ip}\overline{c_j g_{jp}}b_{ij} = \sum_{i,j=1}^{n} d_i\overline{d_j}b_{ij}.$$

Onde $d_i = \sum_{i,j=1}^{n}\sum_{p=1}^{n} c_i g_{ip}$ e $\overline{d_j} = \sum_{i,j=1}^{n}\sum_{p=1}^{n} \overline{c_j g_{jp}}$. Como B é uma matriz não negativa definida, podemos concluir que:

$$\sum_{i,j=1}^{n} c_i\overline{c_j}a_{ij}b_{ij} \geq 0.$$

iii. Como K_n converge para K segue que,

$$\sum_{i,j=1}^{n} c_i\overline{c_j}K(x_i,x_j) = \lim_{p\to\infty}\sum_{i,j=1}^{n} c_i\overline{c_j}K_p(x_i,x_j) \geq 0$$

$\square$

3.1.1 Núcleo de Mercer

Um tipo especial de núcleo positivo definido é o núcleo Mercer que recebe esse nome em homenagem a J. Mercer, autor do artigo [17]. Esse artigo deu origem a diversos estudos sobre propriedades espectrais de operadores integrais gerados por núcleos positivos definidos [14]. Para a definição de núcleo de Mercer precisamos da condição de somalidade, que está explicitada na definição a seguir.

Definição 3.1.23 *Seja I um conjunto enumerável de índices e X um espaço topológico munido de uma medida de Borel ν. Considere $\{\lambda_k\}_{k\in I}$ uma sequência de números reais positivos e $\{f_k : X \to \mathbb{C}\}_{k\in I}$ um conjunto $L^2(X,\nu)$-ortonormal de funções contínuas. Dizemos que ambas satisfazem a condição de somalidade quando:*

$$\sum_{k\in I} \lambda_k |f_k(x)|^2 < \infty, \quad x \in X.$$

Um núcleo que se encaixa nas condições da definição anterior é dito núcleo de Mercer.

Definição 3.1.24 *Se $\{\lambda_k\}_{k\in I}$ e $\{f_k(x)\}_{k\in I}$ satisfazem a condição de somalidade então o núcleo $\phi : X \times X \to \mathbb{C}$ dado por:*

$$\phi(x,y) = \sum_{k\in I} \lambda_k \overline{f_k(x)} f_k(y), \quad x,y \in X$$

é um núcleo de Mercer.

Note que

$$\sum_{k\in I} \lambda_k \overline{f_k(x)} f_k(y) < \infty, \quad x,y \in X.$$

De fato, a condição de somalidade garante que as sequências $\left\{\lambda_k^{\frac{1}{2}}\overline{f_k(x)}\right\}_{k\in I}$ e $\left\{\lambda_k^{\frac{1}{2}} f_k(y)\right\}_{k\in I}$ pertencem a l^2, logo são sequências convergentes, pois l^2 é um espaço de Banach. Usando a desigualdade de Cauchy-Schwarz, obtemos:

$$\sum_{k\in I}\lambda_k\overline{f_k(x)}f_k(y) \leq \left(\sum_{k\in I}\lambda_k|\overline{f_k(x)}|^2\right)^{\frac{1}{2}}\left(\sum_{k\in I}\lambda_k|f_k(y)|^2\right)^{\frac{1}{2}}$$

Daí, a condição de somalidade leva a conclusão desejada. Dessa forma, essa função está bem definida.

Teorema 3.1.25 *Seja ϕ um núcleo de Mercer definido pelas sequências $\{\lambda_k\}_{k\in I}$ e $\{f_k(x)\}_{k\in I}$. O núcleo ϕ é positivo definido.*

Demonstração: Sejam n um inteiro positivo, $\{x_1,x_2,\cdots,x_n\}\subset X$ e $\{c_1,c_2,\cdots,c_n\}\subset\mathbb{C}$, então:

$$\begin{aligned}
\sum_{i,j=1}^{n} c_i\overline{c_j}\phi(x_i,x_j) &= \sum_{i,j=1}^{n} c_i\overline{c_j}\sum_{k\in I}\lambda_k\overline{f_k(x_i)}f_k(x_j)\\
&= \sum_{k\in I}\lambda_k\sum_{i,j=1}^{n} c_i\overline{f_k(x_i)}\overline{c_j}f_k(x_j)\\
&= \sum_{k\in I}\lambda_k\left|\sum_{i=1}^{n} c_i\overline{f_k(x_i)}\right|^2\\
&\geq 0
\end{aligned}$$

Portanto, o núcleo ϕ é positivo definido. □

Observação 3.1.26 *A condição de $L^2(X,\nu)$-ortonormalidade da nossa definição de núcleo de Mercer não é necessária para a demonstração dos resultados anteriores, como o leitor pode facilmente verificar. Essa condição vem da versão original do teorema de Mercer de 1909, encontrada no artigo [17]. Versões atuais do teorema de Mercer (veja por exemplo [14, p.25]) dão condições para que um núcleo seja um núcleo de Mercer. Para isso utilizam o teorema espectral para operadores compactos e autoadjuntos atuando sobre o espaço $L^2(X,\nu)$, o que justifica tal condição na nossa definição.*

3.1.2 Núcleo Gaussiano

Um dos núcleos positivos definidos mais conhecidos foi descrito por Carl Frederich Gauss, por volta de 1809, em seu livro sobre movimento de corpos celestes. O núcleo Gaussiano também é conhecido como função distribuição normal e é dado pela expressão:

$$K(x,y) = e^{-\epsilon^2|x-y|^2},$$

onde $x,y \in \mathbb{R}$ e $\epsilon > 0$.

Antes de mostrarmos que este núcleo é positivo definido, vamos apresentar outro núcleo positivo definido:

$$K_1(x,y) = e^{2\varepsilon^2 xy} = \sum_{l=0}^{\infty} \frac{(2\varepsilon^2)^l}{l!} x^l y^l \quad x,y \in X \subset \mathbb{R}.$$

Este núcleo é positivo definido, já que

$$\begin{aligned}
\sum_{i,j=0}^{n} \overline{c_j} c_i K_1(x_j,x_i) &= \sum_{i,j=0}^{n} \overline{c_j} c_i \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} x_j^k x_i^k \\
&= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} \sum_{i,j=0}^{n} \overline{c_j} c_i x_j^k x_i^k \\
&= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} \left| \sum_{i=1}^{n} c_i x_i^k \right|^2 \\
&\geq 0.
\end{aligned}$$

Note ainda que,

$$\sum_{i,j=1}^{n} \overline{c_i} c_j x_i^k x_j^k = \left| \sum_{i=1}^{n} c_i x_i^k \right|^2 = 0,\ k = 0, 1\ \ldots \Rightarrow c_i = 0.$$

Quando a expressão acima for igual a zero, teremos que cada c_i será igual a zero, ou seja, o núcleo K_1 é de fato estritamente positivo definido. Para dizer isso usamos a unicidade de solução do sistema, que é garantida pelas propriedades da matriz de Vandermonde. Isso está relacionado com a existência e unicidade de um polinômio interpolador para o problema (veja [13, p.214]).

Com isso, podemos ver que o núcleo gaussiano pode ser reescrito como:

$$\begin{aligned} K(x,y) &= e^{-\epsilon^2|x-y|^2} \\ &= e^{-\epsilon^2x^2-\epsilon^2y^2+2\epsilon^2xy} \\ &= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} e^{-\epsilon^2x^2} e^{-\epsilon^2y^2} x^k y^k, \qquad x,y \in \mathbb{R} \end{aligned}$$

e ainda,

$$\begin{aligned} \sum_{i,j=0}^{n} \overline{c_j} c_i K(x_j,x_i) &= \sum_{i,j=0}^{n} \overline{c_j} c_i \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} e^{-\epsilon^2x_j^2} e^{-\epsilon^2x_i^2} x_j^k x_i^k \\ &= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} \sum_{i,j=0}^{n} \overline{c_j} c_i e^{-\epsilon^2x_j^2} e^{-\epsilon^2x_i^2} x_j^k x_i^k \\ &= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} \sum_{i,j=1}^{n} \overline{d_j} d_i x_j^k x_i^k \\ &= \sum_{k=0}^{\infty} \frac{(2\epsilon^2)^k}{k!} \left| \sum_{i=1}^{n} d_i x_i^l \right|^2 \\ &\geq 0. \end{aligned}$$

Onde, $d_i = c_i e^{-\epsilon^2 x_i^2}$, $i = 1, \cdots, n$. Nesse caso, $d_i = 0$ apenas quando $c_i = 0$. Segue então que esse núcleo é estritamente positivo definido.

Cabe notar que essa representação em série não é a representação dada pelo Teorema de Mercer ([14, p. 25]).

Usaremos um raciocínio semelhante para o caso em que $x,y \in \mathbb{R}^n$. Assim, considere o núcleo

$$K_1(x,y) = e^{2\varepsilon^2 xy}, \quad x,y \in X^n \subset \mathbb{R}^n.$$

onde xy denota o produto escalar entre x e y.

Podemos representar esse núcleo da seguinte forma:

$$\begin{aligned} e^{2\varepsilon^2xy} &= e^{2\varepsilon^2(x_1y_1+\cdots+x_ny_n)} \\ &= e^{2\varepsilon^2x_1y_1} \times \cdots \times e^{2\varepsilon^2x_ny_n}. \end{aligned}$$

Segue que esse núcleo é positivo definido porque é o produto de núcleos positivos definidos, pelo Teorema 3.1.22.

Dessa forma, escrevendo o núcleo gaussiano, para $x,y \in \mathbb{R}^n$, obtemos:

$$K(x,y) = e^{-\varepsilon^2|x-y|^2} = \left(e^{-\epsilon^2\|x\|^2}e^{-\epsilon^2\|y\|^2}\right) \times e^{2\epsilon^2 xy}.$$

Como a multiplicação de núcleos positivos definidos também é um núcleo positivo definido obtemos a conclusão desejada.

Para terminar essa seção e esse capítulo apresentamos mais alguns exemplos de núcleos. Mais alguns serão dados quando falarmos de espaços de Hilbert de reprodução, no capítulo seguinte.

Exemplo 3.1.27 *O núcleo dado por $K(x,y) = cos(|x| - |y|)$ é positivo definido.*

Primeiramente, note que:

$$K(x,y) = cos(|x| - |y|) = cos|x|cos|y| + sen|x|sen|y|.$$

Observe ainda que:

$$\sum_{i,j=1}^{n} \overline{c_j}c_i cos|x_j|cos|x_i| = \left|\sum_{i=1}^{n} c_i cos|x_i|\right|^2 \geq 0$$

e,

$$\sum_{i,j=1}^{n} \overline{c_j}c_i sen|x_j|sen|x_i| = \left|\sum_{i=1}^{n} c_i sen|x_i|\right|^2 \geq 0.$$

Como a soma de núcleos positivos definidos é um núcleo positivo definido, podemos concluir o desejado.

Exemplo 3.1.28 *O núcleo dado por $K(x,y) = cos(|x| + |y|)$ não é positivo definido.*

Similar ao exemplo anterior, temos que:

$$K(x,y) = cos(|x| + |y|) = cos|x|cos|y| - sen|x|sen|y|.$$

Agora, note que:

$$\sum_{i,j=1}^{n} \overline{c_j}c_i(cos|x_j|cos|x_i| - sen|x_j|sen|x_i|) = \left|\sum_{i=1}^{n} c_i cos|x_i|\right|^2 - \left|\sum_{i=1}^{n} c_i sen|x_i|\right|^2.$$

Observe que a expressão acima nem sempre é positiva. Logo, o núcleo $K(x,y) = cos(|x| + |y|)$ não é positivo definido.

Exemplo 3.1.29 *O núcleo dado por* $K(x,y) = cos(|x| - |y|)e^{-\epsilon^2|x-y|^2}$ *é positivo definido.*

Basta notar que $K_1(x,y) = cos(|x| - |y|)$ e $K_2(x,y) = e^{-\epsilon^2|x-y|^2}$ são positivos definidos e a multiplicação de núcleos positivos definidos é um núcleo pertencente a $PD(X)$, como visto no Teorema 3.1.22.

4 ESPAÇOS DE REPRODUÇÃO

A teoria de núcleos positivos definidos e espaços de reprodução caminham lado a lado, com algumas ligações entre elas. Os espaços de reprodução são espaços que possuem um núcleo positivo definido. A partir de um núcleo positivo definido podemos gerar um espaço vetorial com produto interno que por sua vez pode ser completado formando um espaço de Hilbert de reprodução. Nesta seção, faremos algumas definições e proposições importantes, também daremos alguns exemplos de espaços de Hilbert de reprodução.

4.1 ESPAÇOS DE REPRODUÇÃO

O conceito espaços de Hilbert de reprodução foi introduzido, por volta de 1950, independentemente por Nachman Aronszajn e Stefan Bergman [12]. Posteriormente, Moore relacionou os núcleos positivos definidos a matrizes hermitianas. Iniciaremos essa seção com o conceito de função avaliação que será necessário para a definição de espaço de Hilbert de reprodução.

Definição 4.1.1 *Seja $\mathcal{H}$ um espaço de funções $f : X \to \mathbb{K}$ definidas em um conjunto não vazio X. Para cada $x \in X$, a função $\sigma_x : \mathcal{H} \to \mathbb{K}$ dada por $\sigma_x(f) = f(x)$ é chamada de função avaliação em x.*

Definição 4.1.2 *Um espaço de Hilbert de funções $f : X \to \mathbb{K}$, definidas em um conjunto não vazio X é um espaço de Hilbert de reprodução (EHR) se σ_x é contínua para todo $x \in X$.*

A definição acima é equivalente ao cumprimento de duas condições, estas estão enunciadas no próximo teorema.

Teorema 4.1.3 *Um espaço de Hilbert W é um espaço de Hilbert de reprodução se, e somente se, existe um núcleo positivo definido K, chamado de núcleo de reprodução, tal que:*

a) $\forall x \in X$, $K(\cdot,x) \in W$, onde $K(\cdot,x)$ denota a função $y \in X \to K(y,x)$;

b) (propriedade de reprodução) $\forall x \in X$ *e* $\forall f \in W$, $\langle f,K(\cdot,x)\rangle = f(x)$.

O núcleo K *é único.*

Demonstração: $\Rightarrow$) Suponhamos $\sigma_x \in W^*$. Pelo Teorema da Representação de Riesz, existe $h_x \in W$ tal que

$$\sigma_x(f) = \langle f,h_x\rangle_W \quad \forall f \in W.$$

Defina $K(y,x) = h_x(y)$, $\forall x,y \in X$. Então:

$$K(\cdot,x) = h_x \quad \in W \quad \text{e} \quad \langle f, K(\cdot,x)\rangle_W = \sigma_x(f) = f(x).$$

Logo, K é um núcleo associado a W e satisfaz os itens do teorema. Mais ainda,

$$K(x,y) = \langle h_y,h_x\rangle_W, \quad x,y \in X,$$

é positivo definido conforme o Lema (3.1.12).

$\Leftarrow$) Suponhamos que o espaço de Hilbert W tenha um núcleo associado K. Assim, pela desigualdade de Cauchy-Schwarz, $\forall f \in W$, temos:

$$\begin{aligned}
|\sigma_x(f)| &= |f(x)| \\
&= |\langle f, K(\cdot,x)\rangle_W| \\
&\leq \|K(\cdot,x)\|_W \|f\|_W \\
&= \sqrt{\langle K(\cdot, x), K(\cdot,x)\rangle_W}\|f\|_W \\
&= \sqrt{K(x,x)}\|f\|_W
\end{aligned}$$

Dessa forma, $\sigma_x : W \to \mathbb{K}$ é limitado. Logo é contínuo e linear e portanto W é um espaço de Hilbert de reprodução. $\square$

A partir de um núcleo positivo definido podemos criar um espaço de Hilbert de reprodução. Para isso, usamos o processo de completamento de um espaço vetorial construído a partir do próprio núcleo. Como este processo é construtivo garantimos a sua existência e pela unicidade do completamento, garantimos também que este espaço é único.

A definição de núcleo positivo definido possibilita definirmos um produto interno no espaço vetorial formado pelas funções

$$g(x) = \sum_{i=1}^{n} c_i K(x,x_i), \quad x \in X,$$

para todo $n \geq 1$, $\{x_1,...,x_n\} \subset X$ e $\{c_1,...,c_n\} \subset \mathbb{C}$. Esse produto interno é definido por

$$\langle g,h \rangle = \sum_{i,j=1}^{n} c_i \overline{d_j} K(x_j,x_i),$$

para

$$h(x) = \sum_{j=1}^{n} d_j K(x,x_j), \quad x \in X.$$

O completamento desse espaço é um espaço de Hilbert $\mathcal{H}_K$ que chamamos de espaço de reprodução, pois $K(\cdot,x) \in \mathcal{H}_K$ e $\langle f, K(.,x) \rangle = f(x)$, para toda $f \in \mathcal{H}_K$ e $x \in X$.

Um resultado muito importante sobre espaços de reprodução é que a convergência implica em convergência pontual.

Teorema 4.1.4 *Seja $\mathcal{H}$ um espaço de reprodução e $\{f_n\}$ uma sequência em $\mathcal{H}$. Se $f_n \to f$ então $f_n(x) \to f(x)$, para todo $x \in X$. Em particular, a convergência é uniforme em todo conjunto $A \subset X$, para o qual $sup_{x \in A} K(x,x) < \infty$.*

Demonstração: Se $x \in X$ então,

$$\begin{aligned}
|f_n(x) - f(x)| &= |\sigma_x(f_n) - \sigma_x(f)| \\
&= |\sigma_x(f_n - f)| \\
&\leq \|\sigma_x\| \|f_n - f\| \\
&= \sqrt{K(x,x)} \|f_n - f\|
\end{aligned}$$

Como por hipótese $f_n \to f$ e $\|\sigma_x\|$ refere-se a norma do funcional no espaço $\mathcal{H}$, podemos concluir que $f_n(x) \to f(x)$, $\forall\ x \in X$. □

A seguir apresentamos dois espaços de reprodução mas não faremos demonstrações. O leitor poderá adptar argumentos de outros exemplos da seção seguinte.

Definição 4.1.5 *Dizemos que uma função $f : X \to \mathbb{R}$ é absolutamente contínua se para todo $\epsilon > 0$ existe $\delta > 0$ tal que*

$$\sum_{i=1}^{n} |x_i - y_i| < \delta \Rightarrow \sum_{i=1}^{n} |f(x_i) - f(y_i)| < \epsilon.$$

Exemplo 4.1.6 *O espaço vetorial de Sobolev dado por:*

$$\mathcal{H} = \left\{ f : [0,1] \to \mathbb{R}; \quad f \text{ é absolutamente contínua}, \quad f(0) = f(1) = 0, \int_0^1 [f'(x)]^2 dx < \infty \right\}$$

é um espaço de Hilbert de reprodução com núcleo associado:

$$R(x,y) = \begin{cases} (1-y)x, & x \leq y \\ (1-x)y, & x > y \end{cases}$$

Com $x, y \in [0,1]$.

O produto interno entre duas funções f, g no espaço de Sobolev é dado por:

$$\langle f, g \rangle = \int_0^1 f'(x)g'(x)dx.$$

Exemplo 4.1.7 *O espaço de Hilbert*

$$\mathcal{H} = \left\{ f(x) =: [0, \infty) \to \mathbb{R} \;\middle|\; \int_0^\infty |f'(x)|^2 \rho(t)dt < \infty, \quad f(0) = 0 \right\},$$

com produto interno dado por

$$\langle f,g \rangle = \int_0^\infty f'(x)g'(x)\rho(x)dx, \qquad \rho(t) = \frac{e^t}{t},$$

é um espaço de reprodução. É fácil verificar que

$$K(x,y) = \int_0^{min(x,y)} \xi e^{-\xi} d\xi, \quad x,y \in [0,\infty),$$

é o núcleo de reprodução. Esse espaço é útil no estudo da inversão da transformada de Laplace em [16].

4.2 EXEMPLOS IMPORTANTES

Para a resolução das equações de Volterra se faz necessário o uso de dois espaços de Hilbert de reprodução $W[0,1]$ e $W_1[0,1]$. Desde que esses espaços de reprodução foram construídos por M.G. Cui, em 1986, a teoria de espaços de reprodução tem sido aplicada com sucesso na resolução de problemas lineares e não lineares, como em equações diferenciais, modelos populacionais e em muitas outras equações da física e engenharia [4].

Definição 4.2.1 *O espaço $W[0,1]$ é definido por:*

$$W[0,1] = \{f(x) | f'(x) \textit{é absolutamente contínua no intervalo } [0,1] \textit{ e } f''(x) \in \mathcal{L}^2[0,1]\}.$$

Seu produto interno é dado por:

$$\langle f(x), g(x) \rangle_W = f(0)g(0) + f'(0)g'(0) + \int_0^1 f''(x)g''(x)dx.$$

Definição 4.2.2 *O espaço $W_1[0,1]$ é definido por:*

$$W_1[0,1] = \{f(x) | f(x) \textit{é absolutamente contínua no intervalo } [0,1] \textit{ e } f'(x) \in \mathcal{L}^2[0,1]\}.$$

Seu produto interno é dado por:

$$\langle f(x), g(x) \rangle_{W_1} = f(0)g(0) + \int_0^1 f'(x)g'(x)dx.$$

Observação 4.2.3 *Podemos mostrar que $f \in W_1[0,1]$ se, e somente se, existe $h \in \mathcal{L}^2[0,1]$ tal que*

$$f(x) = f(0) + \int_0^x h(s)ds.$$

Essa função h é chamada de derivada de f, no sentido de que vale a regra de integração por partes. Ou seja,

$$\int_0^1 f(x)\phi(x)dx = f(x)\int_0^x \phi(s)ds - \int_0^1 \left(\int_0^x \phi(s)ds\right) h(x)dx,$$

sempre que ϕ for contínua em $[0,1]$. Dessa forma, podemos usar a notação $h(x) = f'(x)$

(veja [11, p. 106]).

Note ainda que o espaço do Exemplo 4.1.6 é subespaço de $W_1[0,1]$.

Teorema 4.2.4 *O espaço* $W[0,1]$ *é um espaço de Hilbert.*

Demonstração: Seja $\{f_n\}$ uma sequência de Cauchy em $W[0,1]$, ou seja, dado $\epsilon > 0$ existe N tal que se $n,m > N$, então

$$\|f_n - f_m\|^2 = (f_n(0) - f_m(0))^2 + (f_n'(0) - f_m'(0))^2 + \int_0^1 (f_n''(x) - f_m''(x))^2 dx < \epsilon.$$

Note que $\{f_n(0)\}$ é uma sequência de Cauchy em $\mathbb{R}$ e, como $\mathbb{R}$ é completo, existe $f(0) \in \mathbb{R}$ tal que $f_n(0)$ converge para $f(0)$. Dessa forma, $\epsilon > 0$, para n suficientemente grande vale

$$|f_n(0) - f(0)|^2 < \epsilon.$$

Da mesma forma, $\{f_n'(0)\}$ é uma sequência de Cauchy em $\mathbb{R}$ e existe $f'(0) \in \mathbb{R}$ tal que $f_n'(0)$ converge para $f'(0)$. Assim, para n suficientemente grande, temos

$$|f_n'(0) - f'(0)|^2 < \epsilon.$$

Note ainda que $\{f_n''\}$ é uma sequência de Cauchy em $\mathcal{L}^2[0,1]$, que é completo. Logo essa sequência converge para algum $f'' \in \mathcal{L}^2[0,1]$, ou seja, para n suficientemente grande, temos:

$$\int_0^1 (f_n''(x) - f''(x))^2 dx < \epsilon.$$

Tome

$$f'(x) = f'(0) + \int_0^x f''(s) ds$$

e

$$f(x) = f(0) + \int_0^x f'(s) ds.$$

Consequentemente, $f \in W[0,1]$ e

$$\|f_n - f\|^2 < 3\epsilon,$$

para n suficientemente grande. A demonstração termina. $\square$

Na realidade o espaço $W[0,1]$ é um espaço de Hilbert de reprodução, como mostra o teorema a seguir.

Teorema 4.2.5 *O espaço $W[0,1]$ é um espaço de reprodução com função núcleo:*

$$K(x,y) = \begin{cases} 1 + xy + \dfrac{xy^2}{2} - \dfrac{y^3}{6}, & y \leq x \\ \\ 1 + xy + \dfrac{x^2 y}{2} - \dfrac{x^3}{6}, & y > x \end{cases}$$

Demonstração: Para que K seja um núcleo de reprodução associado a $W[0,1]$, devemos ter:

a $\forall x \in [0,1]$, $K(\cdot,x) \in W[0,1]$;

b) $\forall x \in [0,1]$ e $\forall f \in W[0,1]$ vale a propriedade de reprodução,

$$\langle f(x), K_1(\cdot,x)\rangle_W = f(x).$$

a) Observe que, se $g(x) = K(x,y)$, para $y \in [0,1]$ fixo, então

$$g'(x) = \begin{cases} y + \dfrac{y^2}{2}, & y \leq x \\ y + xy - \dfrac{x^2}{2}, & y > x \end{cases}, \qquad g''(x) = \begin{cases} 0, & y \leq x \\ y - x, & y > x \end{cases}$$

estão em $\mathcal{L}^2[0,1]$ e

$$g'(x) = g'(0) + \int_0^x g''(s)ds, \quad g(x) = g(0) + \int_0^x g'(s)ds,$$

ou seja, $K(\cdot,y) \in W[0,1]$.

b) Cálculos diretos nos dizem que, se $f \in W[0,1]$, então

$$\begin{aligned} \langle f, K(\cdot,y)\rangle_W &= f(0)K(0,y) + f'(0)K'(0,y) + \int_0^1 f''(x)K''(x,y)dx \\ &= f(0).1 + f'(0)y + \int_0^y f''(x)K''(x,y)dx + \int_y^1 f''(x)K''(x,y)dx \\ &= f(0) + yf'(0) + \int_0^y f''(x)(y-x)dx + \int_y^1 f''(x).0\, dx. \end{aligned}$$

Assim,

$$\begin{aligned}\langle f,K(\cdot,y)\rangle_W &= f(0)+yf'(0)+[(y-x)f'(x)+f(x)]|_0^y \\ &= f(0)+yf'(0)+[(y-y)f'(y)+f(y)]-[(y-0)f'(0)+f(0)] \\ &= f(0)+yf'(0)+f(y)-yf'(0)-f(0) \\ &= f(y).\end{aligned}$$

□

O mesmo argumento que usamos nos dois últimos resultados pode ser usado para $W_1[0,1]$, como faremos a seguir.

Teorema 4.2.6 *O espaço* $W_1[0,1]$ *é um espaço de Hilbert.*

Demonstração: Seja $\{f_n\}$ uma sequencia de Cauchy em $W_1[0,1]$. Isto é, dado $\varepsilon > 0$, existe N tal que, se $n,m > N$, então

$$\|f_n - f_m\|^2 = |f_n(0) - f_m(0)|^2 + \int_0^1 |f_n'(x) - f_m'(x)|^2 dx < \varepsilon.$$

Note que $\{f_n(0)\}$ é uma sequência de Cauchy em $\mathbb{R}$ e $\{f_n'\}$ é uma sequência de Cauchy em $\mathcal{L}^2[0,1]$. Como esses espaços são completos, existem $f(0) \in \mathbb{R}$ e $f' \in \mathcal{L}^2[0,1]$ tais que,

$$\lim_{n\to\infty} f_n(0) = f(0), \quad \lim_{n\to\infty} \int_0^1 |f_n'(x) - f'(x)|^2 dx = 0$$

Ou seja, existe N_1 tal que, se $n > N_1$ então:

$$|f_n(0) - f(0)|^2 < \varepsilon \quad \text{e} \quad \int_0^1 |f_n'(x) - f'(x)|^2 dx < \varepsilon.$$

Tome

$$f(x) = f(0) + \int_0^x f'(s)ds.$$

Consequentemente, $f \in W_1[0,1]$ e segue que

$$\|f_n - f\|^2 < 2\varepsilon.$$

A demonstração termina. □

O espaço $W_1[0,1]$ também é um espaço de Hilbert de reprodução, como mostra o teorema a seguir.

Teorema 4.2.7 *O espaço $W_1[0,1]$ é um espaço de reprodução com função núcleo:*

$$K_1(x,y) = 1 + \min(x,y) = \begin{cases} 1+y, & y \leq x \\ 1+x, & y > x \end{cases}$$

Demonstração: Para que K_1 seja um núcleo de reprodução associado a $W_1[0,1]$, devemos ter:

a) $\forall x \in [0,1]$, $K_1(\cdot,x) \in W_1[0,1]$;

b) $\forall x \in [0,1]$ e $\forall f \in W_1[0,1]$ vale a propriedade de reprodução

$$\langle f(x), K_1(\cdot,x)\rangle_{W_1} = f(x).$$

a) Observe que, se $g(x) = K_1(x,y)$, para $y \in [0,1]$ fixo, então

$$g'(x) = \begin{cases} 0, & y \leq x \\ 1, & y > x \end{cases}$$

está em $\mathcal{L}^2[0,1]$ e

$$g(x) = g(0) + \int_0^x g'(s)ds,$$

ou seja, $K_1(\cdot,y) \in W_1[0,1]$.

b) Cálculos diretos nos dizem que, se $f \in W_1[0,1]$, então

$$\begin{aligned} \langle f, K_1(\cdot,y)\rangle_{W_1} &= f(0)K_1(0,y) + \int_0^1 f'(x)K_1'(x,y)dx \\ &= f(0)(1+0) + \int_0^y f'(x)K_1'(x,y)dx + \int_y^1 f'(x)K_1'(x,y)dx \\ &= f(0) + \int_0^y f'(x)(1+x)'dx + \int_x^1 f'(x)(1+y)'dx \end{aligned}$$

Logo,

$$
\begin{aligned}
\langle f,K_1(\cdot,y)\rangle_{W_1} &= f(0)+\int_0^y f'(x).1\,dx+\int_y^1 f'(x).0\,dx\\
&= f(0)+f(x)|_0^y\\
&= f(0)+f(y)-f(0)\\
&= f(y).
\end{aligned}
$$

$\square$

5 EQUAÇÕES DE VOLTERRA

Neste capítulo apresentaremos as equações integrais de Volterra e alguns conceitos necessários para a resolução de um sistema de equações integrais de Volterra.

Vito Volterra foi um famoso matemático italiano que publicou, em 1896, um trabalho sobre equações integrais, que conhecemos hoje como equações integrais de Volterra. Os trabalhos de Volterra juntamente com os de Ivar Fredholm, matemático sueco, marcaram o começo do estudo da análise funcional [1].

5.1 EXISTÊNCIA E UNICIDADE DE SOLUÇÕES

As equações diferenciais e integrais nos intrigam quanto a existência e unicidade de solução. Nesta seção, daremos as condições suficientes para que um sistema de equações integrais de Volterra tenha solução única. Desta forma, garantimos que a solução analítica que será apresentada no final deste capítulo é a única solução do problema.

Definição 5.1.1 *Dado um intervalo* $[0,T]$ *e conhecendo as funções* $g(t)$ *e* $r(t,s,u)$, *uma equação integral de Volterra é definida da seguinte forma:*

a) Primeiro Tipo

$$\int_0^t r(t,s,f(s))ds = g(t), \quad t \in [0,T].$$

b) Segundo Tipo

$$f(t) = g(t) + \int_0^t r(t,s,f(s))ds, \quad t \in [0,T].$$

Observação 5.1.2 *Uma equação integral de Volterra é chamada de linear quando*

$$r(t,s,f(s)) = r_1(t,s)f(s).$$

Observação 5.1.3 *Note que uma equação diferencial ordinária de primeira ordem, dada* *por:*

$$f'(x) = a(x)f(x) + b(x), \quad x \in [0,1]$$

é equivalente a equação integral linear de Volterra do segundo tipo

$$f(x) = f(0) + \int_0^x b(t)dt + \int_0^x a(t)f(t)dt.$$

Nesse caso, $g(x) = f(0) + \int_0^x b(t)dt$ *e o núcleo* $r(x,t) = a(t)$ *não depende de* x*.*

Por simplicidade de notação, nosso objetivo nesta seção é estudar o sistema de equações integrais lineares de Volterra da forma:

$$\begin{cases} a_{11}(x)f_1(x) - b_{11}\displaystyle\int_0^x r_{11}(x,t)f_1(t)dt + a_{12}(x)f_2(x) - b_{12}\int_0^x r_{12}(x,t)f_2(t)dt &= u_1(x) \\ \\ a_{21}(x)f_1(x) - b_{21}\displaystyle\int_0^x r_{21}(x,t)f_1(t)dt + a_{22}(x)f_2(x) - b_{22}\int_0^x r_{22}(x,t)f_2(t)dt &= u_2(x) \end{cases}$$

Porém, o método pode ser adaptado para tratar de sistemas maiores. O método pode também ser adaptado para sistemas não lineares, conforme será discutido ao final deste capítulo (veja por exemplo [18]).

É claro que podemos escrever esse sistema na forma matricial como

$$A(x)F(x) + \int_0^x R(x,t)F(t)dt = U(x). \tag{5.1.1}$$

Assumiremos que a equação (5.1.1) tem solução única (veja o Lema 5.1.5). Sob condições adicionadas a seguir, podemos considerar esse sistema como:

$$U(x) = VF(x), \tag{5.1.2}$$

onde

$$V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} : W[0,1] \oplus W[0,1] \to W_1[0,1] \oplus W_1[0,1],$$

com

$$v_{ij}(f_j)(x) = a_{ij}(x)f_j(x) - b_{ij}\int_0^x r_{ij}(x,t)f_j(t)dt,$$

$$U(x) = \begin{pmatrix} u_1(x) \\ u_2(x) \end{pmatrix} \in W_1[0,1] \oplus W_1[0,1],$$

e

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} \in W[0,1] \oplus W[0,1].$$

Exemplo 5.1.4 *Considere o sistema de equações lineares de Volterra:*

$$\begin{cases} f_1(x) - \displaystyle\int_0^x (x^2-t) f_1(t)dt - \int_0^x (x^2-t) f_2(t)dt = x + \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{3} \\ \\ -\displaystyle\int_0^x t f_1(t)dt + f_2(x) - \int_0^x t f_2(t)dt = x^2 - \frac{x^3}{3} - \frac{x^4}{4} \end{cases}$$

Nesse sistema temos

$$A(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R(x,t) = -\begin{pmatrix} x^2-t & x^2-t \\ t & t \end{pmatrix}, \quad U(x) = \begin{pmatrix} x + \dfrac{x^3}{3} - \dfrac{x^4}{4} - \dfrac{x^5}{3} \\ \\ x^2 - \dfrac{x^3}{3} - \dfrac{x^4}{4} \end{pmatrix}.$$

Como exemplo temos, nesse caso,

$$v_{11}(f_1)(x) = f_1(x) - \int_0^x (x^2-t) f_1(t)dt, \qquad v_{12}(f_2)(x) = -\int_0^x (x^2-t) f_2(t)dt$$

e

$$v_{21}(f_1)(x) = -\int_0^x t f_1(t)dt, \qquad v_{22}(f_2)(x) = f_2(x) - \int_0^x t f_2(t)dt.$$

Podemos verificar por substituição que a solução do sistema é dada por:

$$F(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}.$$

Definiremos o produto interno em $W[0,1] \oplus W[0,1]$ como a soma dos produtos internos de cada coordenada, ou seja,

$$\langle F, G \rangle = \sum_{i=1}^{2} \langle f_i, g_i \rangle_W.$$

O produto interno em $W_1[0,1] \oplus W_1[0,1]$ será definido da mesma forma, como a soma dos produtos internos de cada coordenada pertencente a $W_1[0,1]$.

Para garantir que a solução de um sistema de equações de Volterra é única usaremos o Teorema do Ponto Fixo de Banach 2.3.4. Este fato será melhor detalhado no próximo lema.

Lema 5.1.5 *Sejam $A(x)$, $F(x)$, $U(x)$ e $R(x,y)$ funções contínuas, para $x,y \in [0,1]$, com $A(x)$ invertível. Então, o sistema (5.1.1) tem solução única.*

Demonstração: Como $A(x)$ é invertível, podemos reescrever o sistema

$$A(x)F(x) + \int_0^x R(x,t)F(t)dt = U(x) \tag{5.1.3}$$

como

$$F(x) + \int_0^x R_1(x,t)F(t)dt = U_1(x) \tag{5.1.4}$$

Vamos definir o operado $T_u(F) : C[0,1] \to C[0,1]$, dado por:

$$T_u(F)(x) = U_1(x) - \int_0^x R_1(x,t)F(t)dt \tag{5.1.5}$$

Como $C[0,1]$ é um espaço de Banach ([1, p. 3]) vamos mostrar que T_u^n é uma contração, para n grande, com a finalidade de aplicarmos o Teorema do Ponto Fixo de Banach. Assim,

$$\begin{aligned}
\|T_u(F)(x) - T_u(G)(x)\| &= \left\| \int_0^x R_1(x,t)(F(t) - G(t))dt \right\| \\
&\leq \int_0^x \|R_1(x,t)\| \|F(t) - G(t)\| dt \\
&\leq \int_0^x M\|F - G\| dt \\
&\leq M\|F - G\|
\end{aligned}$$

onde

$$M = \max_{x,y\in[0,1]} \|R_1(x,y)\|, \qquad \|F - G\| = \max_{x\in[0,1]} \|F(x) - G(x)\|$$

e $\|R_1(x,y)\|$ é uma norma matricial.

Note que M não necessariamente pertence ao intervalo (0,1). Mas,

$$\begin{aligned}
\|T_u^2(F)(x) - T_u^2(G)(x)\| &= \left\| \int_0^x R_1(x,t) \left(\int_0^t R_1(t,z)(F(z) - G(z))dz \right) dt \right\| \\
&\leq \int_0^x \int_0^t M^2 \|F - G\| dz\, dt \\
&\leq M^2 \|F - G\| \int_0^x t\, dt \\
&\leq \frac{M^2}{2} \|F - G\|
\end{aligned}$$

Por indução chegamos a

$$\|T_u^n(F)(x) - T_u^n(G)(x)\| \leq \frac{M^n}{n!} \|F - G\|.$$

Observe que, para n suficientemente grande, vale a desigualdade $\dfrac{M^n}{n!} < 1$. Logo, T_u^n é uma contração e, pelo Corolário (2.3.5), T_u possui um único ponto fixo. Portanto, o sistema (5.1.1) possui solução única. □

Observação 5.1.6 *É possível adaptar a demonstração do Lema 5.1.5 e verificar que o mesmo é ainda verdadeiro quando supormos que o sistema 5.1.1 é não linear, ou seja, que*

$$F(x) + \int_0^x R_1(x, t, F(t))dt = U_1(x),$$

com $F(x), U(x)$ *e* $R_1(x, y, F(y))$ *funções contínuas e*

$$|R_1(x, y, v) - R(x, y, w)| \leq L|v - w|,$$

para $x, y \in [0, 1]$, $v, w \in \mathbb{R}^2$ *e* $L > 0$. *Note ainda que o mesmo vale se trabalharmos com um intervalo* $[a, b]$, *em vez de* $[0, 1]$. *Na realidade isso se aplica a grande parte deste trabalho, com os devidos ajustes de notação.*

5.2 MÉTODO DE RESOLUÇÃO

A seguir apresentaremos alguns conceitos envolvendo a teoria de espaços de reprodução que usaremos para encontrar uma aproximação da solução da equação (5.1.1).

Lema 5.2.1 *Seja* $r(x,t) \in C([0,1] \times [0,1])$*. Se* $\dfrac{\partial}{\partial x} r(x,t) \in \mathcal{L}^2[0,1]$*, então o operador* $T : W[0,1] \to W_1[0,1]$ *dado por* $T(f) = \displaystyle\int_0^x r(x,t)f(t)dt$ *é limitado.*

Primeiramente, vamos garantir que T está bem definida, ou seja, que a imagem de T está em $W_1[0,1]$. Mas isso segue do fato de

$$(T(f))'(x) = r(x,x)f(x) + \int_0^x \frac{\partial}{\partial x} r(x,t)f(t)dt$$

ser um elemento de $\mathcal{L}^2[0,1]$ e da definição de $W_1[0,1]$.

Como $K(\cdot,x)$ é um núcleo de reprodução de $W[0,1]$, $\|K(\cdot,x)\|_W = \sqrt{K(x,x)} \leq \sqrt{3}$ e $T(f)(0) = 0$, podemos concluir que

$$\|T(f)\|_{W_1}^2 = \int_0^1 |(T(f))'(x)|^2 dx$$

Assim,

$$\begin{aligned}
\|T(f)\|_{W_1}^2 &\leq \int_0^1 \left| r(x,x)f(x) + \int_0^x \frac{\partial}{\partial x} r(x,t)f(t)dt \right|^2 dx \\
&\leq \int_0^1 \left| 3M\|f\|_W + 3\|f\|_W \left\| \frac{\partial}{\partial x} r \right\|_{L^1} \right|^2 dx \\
&\leq \left| 3M + 3 \left\| \frac{\partial}{\partial x} r \right\|_{L^1} \right|^2 \|f\|_W^2, \qquad f \in W[0,1].
\end{aligned}$$

onde $M = sup_{x \in [0,1]} |r(x,x)|$. Seque então que T é limitado, com

$$\|T\| \leq 3 \left| M + \left\| \frac{\partial}{\partial x} r \right\|_{L^1} \right|.$$

□

Lema 5.2.2 *Seja $v : W[0,1] \to W_1[0,1]$ um operador linear limitado e $v^* : W_1[0,1] \to W[0,1]$ o seu adjunto. Então*

$$(v^* K_1(\cdot,x))(y) = (vK(\cdot,y))(x), \quad x,y \in X.$$

Demonstração: Como v_{ij}^* é operador adjunto de v_{ij}, então:

$$\langle v^* K_1(\cdot,y), K(\cdot,x)\rangle_W = \langle K_1(\cdot,y), vK(\cdot,x)\rangle_{W_1}, \quad x,y \in X.$$

Logo, para cada $x,y \in X$, temos que

$$\begin{aligned} (v^* K_1(\cdot,y))(x) &= \langle v^* K_1(\cdot,y), K(\cdot,x)\rangle_W \\ &= \langle K_1(\cdot,y), vK(\cdot,x)\rangle_{W_1} \\ &= \langle vK(\cdot,x), K_1(\cdot,y)\rangle_{W_1} \\ &= (vK(\cdot,x))(y) \end{aligned}$$

□

Lema 5.2.3 *Seja $V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$, como na equação (5.1.2). Então seu adjunto é da forma $V^* = \begin{pmatrix} v_{11}^* & v_{21}^* \\ v_{12}^* & v_{22}^* \end{pmatrix}$, onde v_{ij}^* é o adjunto de v_{ij}.*

Demonstração: Sejam $F \in W[0,1] \oplus W[0,1]$ e $G \in W_1[0,1] \oplus W_1[0,1]$. Então

$$\begin{aligned} \langle VF, G\rangle &= \left\langle \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} v_{11} f_1 + v_{12} f_2 \\ v_{21} f_1 + v_{22} f_2 \end{pmatrix}, \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\rangle \\ &= \langle v_{11} f_1, g_1\rangle_{W_1} + \langle v_{12} f_2, g_1\rangle_{W_1} + \langle v_{21} f_1, g_2\rangle_{W_1} + \langle v_{22} f_2, g_2\rangle_{W_1} \\ &= \langle f_1, v_{11}^* g_1\rangle_W + \langle f_2, v_{21}^* g_1\rangle_W + \langle f_1, v_{21}^* g_2\rangle_W + \langle f_2, v_{22}^* g_2\rangle_W \end{aligned}$$

De onde temos que,

$$
\begin{aligned}
\langle VF, G\rangle &= \left\langle \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} v_{11}^* g_1 + v_{21}^* g_2 \\ v_{12}^* g_1 + v_{22}^* g_2 \end{pmatrix} \right\rangle \\
&= \left\langle \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} v_{11}^* & v_{21}^* \\ v_{12}^* & v_{22}^* \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\rangle \\
&= \langle F, V^* G\rangle
\end{aligned}
$$

□

Teorema 5.2.4 *Sejam v_{ij} como na equação (5.1.2) e $\{x_i\}_{i=1}^{\infty}$ um conjunto denso no intervalo $[0,1]$. Se essa equação possuir solução única, então $\{\psi_{ij}\}_{(1,1)}^{(\infty,2)}$ é linearmente independente em $W[0,1] \oplus W[0,1]$, onde*

$$
\psi_{i1}(x) = \begin{pmatrix} (v_{11}K(\cdot,x))(x_i) \\ (v_{12}K(\cdot,x))(x_i) \end{pmatrix} \quad e \quad \psi_{i2}(x) = \begin{pmatrix} (v_{21}K(\cdot,x))(x_i) \\ (v_{22}K(\cdot,x))(x_i) \end{pmatrix}.
$$

Demonstração: Suponhamos, por absurdo, que $\{\psi_{ij}\}_{(1,1)}^{(h,2)}$ seja linearmente dependente. Ou seja, $\displaystyle\sum_{j=1}^{2}\sum_{i=1}^{h} c_{ij}\psi_{ij} = 0$ sem que todos os escalares c_{ij} sejam nulos. Dessa forma, seja $F \in W[0,1] \oplus W[0,1]$. Note que:

$$
\begin{aligned}
\left\langle VF, \begin{pmatrix} \sum_{i=1}^{h} c_{i1}K_1(\cdot, x_i) \\ \sum_{i=1}^{h} c_{i2}K_1(\cdot, x_i) \end{pmatrix} \right\rangle &= \left\langle F, V^* \begin{pmatrix} \sum_{i=1}^{h} c_{i1}K_1(\cdot, x_i) \\ \sum_{i=1}^{h} c_{i2}K_1(\cdot, x_i) \end{pmatrix} \right\rangle \\
&= \left\langle F, \begin{pmatrix} v_{11}^* \sum_{i=1}^{h} c_{i1}K_1(\cdot, x_i) \\ v_{12}^* \sum_{i=1}^{h} c_{i1}K_1(\cdot, x_i) \end{pmatrix} \right\rangle \\
&+ \left\langle F, \begin{pmatrix} v_{21}^* \sum_{i=1}^{h} c_{i2}K_1(\cdot, x_i) \\ v_{22}^* \sum_{i=1}^{h} c_{i2}K_1(\cdot, x_i) \end{pmatrix} \right\rangle
\end{aligned}
$$

Segue do Lema (5.2.2) e da nossa hipótese que

$$\left\langle VF, \begin{pmatrix} \sum_{i=1}^{h} c_{i1}K_1(\cdot,x_i) \\ \sum_{i=1}^{h} c_{i2}K_1(\cdot,x_i) \end{pmatrix} \right\rangle = \left\langle F, \sum_{j=1}^{2}\sum_{i=1}^{h} c_{ij}\psi_{ij} \right\rangle = 0.$$

Como V é injetivo e sobrejetivo e F é um elemento qualquer de $W[0,1] \oplus W[0,1]$, podemos escolhê-lo de tal forma que

$$VF = \begin{pmatrix} \sum_{i=1}^{h} c_{i1}K_1(\cdot,x_i) \\ \sum_{i=1}^{h} c_{i2}K_1(\cdot,x_i) \end{pmatrix},$$

o que garante que $VF = F = 0$.

Para concluir a demonstração, vamos definir a função $u_k \in W_1[0,1]$ de tal forma que:

$$u_k(x) = \begin{cases} 1, & x = x_k \\ 0, & x = x_i, i \neq k, i = 1, \cdots, h. \end{cases}$$

Pela propriedade de reprodução, podemos escrever $c_{kj} = \left\langle u_k, \sum_{i=1}^{h} c_{ij}K_1(\cdot,x_i) \right\rangle = 0$, e segue que $c_{kj} = 0$. Isso é uma contradição, pois temos como suposição que os escalares c_{ij} não são todos nulos. Assim, concluímos que $\{\psi_{ij}\}_{(1,1)}^{(h,2)}$ é linearmente independente em $W[0,1] \oplus W[0,1]$. $\square$

Teorema 5.2.5 *Seja* $\{x_i\}_{i=1}^{\infty}$ *um conjunto denso no intervalo* $[0,1]$. *O completamento do espaço gerado pelo conjunto* $\{\psi_{ij}\}_{(1,1)}^{(h,2)}$ *é* $W[0,1] \oplus W[0,1]$, *onde* ψ_{ij} $i = 1, 2, \cdots$, $j = 1, 2$ *é dado por:*

$$\psi_{i1}(x) = \begin{pmatrix} (v_{11}K_1(\cdot,x))(x_i) \\ (v_{12}K_1(\cdot,x))(x_i) \end{pmatrix}, \qquad \psi_{i2}(x) = \begin{pmatrix} (v_{21}K_1(\cdot,x))(x_i) \\ (v_{22}K_1(\cdot,x))(x_i) \end{pmatrix}.$$

Demonstração: No teorema (5.2.4) vimos que $\{\psi_{ij}\}_{(1,1)}^{(h,2)}$ é linearmente independente em $W[0,1] \oplus W[0,1]$. Queremos mostrar que esse conjunto é denso em $W[0,1] \oplus W[0,1]$. Dessa forma, se $F(x) = (f_1(x),f_2(x))^T \in W[0,1] \oplus W[0,1]$, é tal que $\langle F(x),\psi_{ij}(x)\rangle = 0$, para

$i = 1,2,\cdots$ e $j = 1,2.$, basta concluímos que $F = 0$, este resultado refere-se ao teorema 2.1.15. Usando o Lema 5.2.2 podemos escrever:

$$\begin{aligned}
0 &= \langle F,\psi_{i1}\rangle \\
&= \left\langle \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} v_{11}K(\cdot,x_i) \\ v_{12}K(\cdot,x_i) \end{pmatrix} \right\rangle \\
&= \langle f_1,v_{11}K(\cdot,x_i)\rangle + \langle f_2,v_{12}K(\cdot,x_i)\rangle \\
&= \langle f_1,[v_{11}^*K_1(\cdot,x_i)]\rangle + \langle f_2,[v_{12}^*K_1(\cdot,x_i)]\rangle \\
&= v_{11}f_1(x_i) + v_{12}f_2(x_i)
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
0 &= \langle F,\psi_{i2}\rangle \\
&= \left\langle \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} v_{21}K_1(\cdot,x_i) \\ v_{22}K_1(\cdot,x_i) \end{pmatrix} \right\rangle \\
&= \langle f_1,v_21K_1(\cdot,x_i)\rangle + \langle f_2(x),v_22K_1(\cdot,x_i)\rangle \\
&= \langle f_1,[v_{21}^*K(\cdot,x_i)]\rangle + \langle f_2,[v_{22}^*K(\cdot,x_i)]\rangle \\
&= v_{21}f_1(x_i) + v_{22}f_2(x_i)
\end{aligned}$$

Da densidade de $\{x_i\}$ e da unicidade de solução da equação (5.1.2) temos a validade das igualdades anteriores apenas quando $(f_1(x),f_2(x)) = (0,0)$. Isso garante que o completamento do conjunto gerado por $\{\psi_{i1}(x),\psi_{i2}(x)\}_{i=1}^{\infty}$ é $W[0,1] \oplus W[0,1]$. $\square$

Como o completamento do conjunto gerado por $\{\psi_{i1}(x),\psi_{i2}(x)\}_{i=1}^{\infty}$ é $W[0,1] \oplus W[0,1]$, podemos usar o processo de ortogonalização de Gram-Schmidt para obter uma base ortogonal de $W[0,1] \oplus W[0,1]$ (veja o Teorema 2.1.21). Para isso, vamos reclassificar a sequência $\{\psi_{i1}(x),\psi_{i2}(x)\}_{i=1}^{\infty}$ da seguinte forma:

$$A_i(x) = \begin{cases} \psi_{(\frac{i+1}{2})1}, & i \text{ é ímpar} \\ \psi_{(\frac{i}{2})2}, & i \text{ é par.} \end{cases}$$

Aplicando o processo de ortogonalização de Gram-Schmidt:

$$\begin{aligned} A'_1 &= A_1 \\ A'_2 &= A_2 - \frac{\langle A_2, A'_1 \rangle}{\|A'_1\|^2} A'_1 \\ &\vdots \\ A'_k &= A_k - \sum_{j=1}^{k-1} \frac{\langle A_k, A'_j \rangle}{\|A'_j\|^2} A'_j \end{aligned}$$

Assim, obtemos $\{A'_1, A'_2, A'_3, \cdots\}$, que pelos resultados anteriores é uma base ortogonal de $W[0,1] \oplus W[0,1]$. Logo, conforme feito na Seção (2.1.1), todo elemento $w \in W[0,1] \oplus W[0,1]$ pode ser escrito como

$$w = \sum_{j=1}^{\infty} \frac{\langle w, A'_j \rangle}{\|A'_j\|^2} A'_j.$$

Em seguida, dividindo cada vetor pela sua norma, obtemos uma base ortonormal. Seja $B = \{\overline{A_1}, \overline{A_2}, \overline{A_3} \cdots\}$ tal base, dada por:

$$\overline{A_i}(x) = \sum_{k=1}^{i} \beta_{ik} A_k(x), \quad i = 1,2, \cdots \tag{5.2.1}$$

Teorema 5.2.6 *Se* $\{x_i\}_{i=1}^{\infty}$ *é denso no intervalo* $[0,1]$*, então a única solução da equação (5.1.1) é dada por:*

$$F(x) = \sum_{i=1}^{\infty} \left(\sum_{k=1}^{i} \beta_{ik} \alpha_k \right) \overline{A_i}(x) \tag{5.2.2}$$

Onde

$$\alpha_k = \langle F, A_k \rangle = \begin{cases} u_1(x_{\frac{k+1}{2}}), & k \text{ é ímpar} \\ u_2(x_{\frac{k}{2}}), & k \text{ é par.} \end{cases}$$

Demonstração: Suponhamos que $F(x)$ seja a solução da equação (5.1.1). Vimos que $B = \{\overline{A_1}, \overline{A_2}, \overline{A_3} \cdots\}$ é uma base ortonormal de $W[0,1] \oplus W[0,1]$, onde $\overline{A}(x)$ é conforme a equação (5.2.1). Assim, Pela Identidade de Parserval (veja o Teorema 2.1.20), podemos escrever:

$$F(x) = \sum_{i=1}^{\infty} \langle F, \overline{A_i} \rangle \overline{A_i}(x).$$

Agora, note que:

$$\begin{aligned}
\langle F,\overline{A_i}\rangle &= \left\langle F,\sum_{k=1}^{i}\beta_{ik}A_k\right\rangle \\
&= \sum_{k=1}^{i}\beta_{ik}\langle F,A_k\rangle \\
&= \sum_{k=1}^{i}\beta_{ik}\alpha_k
\end{aligned}$$

Assim,

$$\begin{aligned}
\langle F,\overline{A_i}\rangle &= \sum_{k=1,\text{k é ímpar}}^{i}\beta_{ik}v_{1k}f_1(x_{\frac{k+1}{2}}) + \sum_{k=1,\text{k é par}}^{i}\beta_{ik}v_{2k}f_2(x_{\frac{k}{2}}) \\
&= \sum_{k=1,\text{k é ímpar}}^{i}\beta_{ik}u_1(x_{\frac{k+1}{2}}) + \sum_{k=1,\text{k é par}}^{i}\beta_{ik}u_2(x_{\frac{k}{2}})
\end{aligned}$$

□

O resultado seguinte segue da propriedade de reprodução de $W[0,1]$, mas faremos aqui uma outra demonstração.

Teorema 5.2.7 *Dada uma função absolutamente contínua $u(x) \in W[0,1]$, temos que:*

$$|u(x)| \leq 3\|u\|.$$

Demonstração: Seja $u(x) \in W[0,1]$, logo $u'(x)$ é absolutamente contínua e $u''(x) \in \mathcal{L}^2[0,1]$. Assim, podemos escrever:

$$u'(x) = u'(0) + \int_0^x u''(s)ds.$$

Integrando ambos os lados da equação, obtemos:

$$\begin{aligned}
\int_0^x u'(s)ds &= \int_0^x u'(0)ds + \int_0^x\int_0^t u''(s)dsdt \\
u(x) - u(0) &= xu'(0) + \int_0^x\int_0^t u''(s)dsdt \\
u(x) &= u(0) + xu'(0) + \int_0^x\int_0^t u''(s)dsdt.
\end{aligned}$$

Dessa forma,

$$\begin{aligned}|u(x)| &= \left|u(0) + xu'(0) + \int_0^x \int_0^t u''(s)dsdt\right| \\ &\leq |u(0)| + |x||u'(0)| + \left|\int_0^x \int_0^t u''(s)dsdt\right|.\end{aligned}$$

Agora, analisando cada termo separadamente, note que:

$$|u(0)| = \sqrt{u^2(0)} \leq \sqrt{u^2(0) + (u'(0))^2 + \int_0^1 (u''(x))^2 dx} = \|u\|_W$$

$$|u'(0)| = \sqrt{(u'(0))^2} \leq \sqrt{u^2(0) + (u'(0))^2 + \int_0^1 (u''(x))^2 dx} = \|u\|_W$$

$$\begin{aligned}\left|\int_0^x \int_0^t u''(s)dsdt\right| &\leq \int_0^1 1dt \int_0^1 |u''(s)|ds \\ &= \int_0^1 |u''(s)|.1ds \\ &\leq \int_0^1 (u''(s))^2 ds| \\ &\leq \sqrt{u^2(0) + (u'(0))^2 + \int_0^1 (u''(x))^2 dx} \\ &= \|u\|_W.\end{aligned}$$

Portanto, $|u(x)| \leq 3\|u\|$. □

Teorema 5.2.8 *A série dada por*

$$F_n(x) = \sum_{i=1}^{n} \left(\sum_{k=1}^{i} \beta_{ik}\alpha_k \right) \overline{A}_i(x) = (f_{1n}(x), f_{2n}(x))^T$$

converge uniformemente para a solução exata

$$F(x) = (f_1(x), f_2(x))^T.$$

Demonstração: Segue da Desigualdade de Bessel e da Identidade de Parseval (veja os teoremas 2.1.18 e 2.1.20) que $\|F_n - F\|^2 \searrow 0$, ou seja, a sequência converge para 0 de forma não crescente, quando $n \to \infty$. Observe que:

$$\|F_n - F\|^2 = \sum_{i=1}^{2} \|f_{in} - f_i\|^2 \searrow 0.$$

Usando o teorema anterior, obtemos:

$$|f_{in}(x) - f_i(x)| \leq \sqrt{3}\|f_{in} - f_i\|.$$

Dessa forma, $f_{in}(x)$ converge uniformemente para $f_i(x)$. Assim, $F_n(x)$ converge uniformemente para $F(x)$ no intervalo [0,1]. □

Corolário 5.2.9 *A sequência dada por $\|F_n - F\|$ é não crescente.*

Demonstração: Isso segue diretamente da Desigualdade de Bessel (2.1.18) e da Identidade de Parseval (2.1.20), já que:

$$\|F_n - F\|^2 = \sum_{i=n+1}^{\infty} \left| \sum_{k=1}^{i} \beta_{ik}\alpha_k \right|^2.$$

□

5.3 GENERALIZAÇÃO DO MÉTODO

Nesta seção faremos um estudo da adaptação do que discutimos até o momento para tratar de problemas mais gerais. Como exemplo, queremos tratar de adaptar o método da seção anterior para aproximar soluções de sistemas não lineares de Volterra da forma

$$A(x)F(x) + \int_0^x R(x,t)G(F(t))dt = U(x), \tag{5.3.1}$$

uma vez que é esse tipo de equação que aparece frequentemente em problemas de modelagem de fenômenos biológicos ([19]).

Pretendemos olhar para um caso ainda mais geral, onde estamos interessados em resolver o sistema de equações não lineares de Fredholm-Volterra (veja por exemplo [18] que trata de apenas uma equação), dado por

$$\begin{aligned} U(x) &= A(x)F'(x) + B(x)G(x)N(x,F(x)) + C(x)\int_0^1 R_1(x,t)G_1(F(t))dt \\ &\quad + D(x)\int_0^x R_2(x,t)G_2(F(t))dt. \end{aligned} \tag{5.3.2}$$

Para resolvermos os sistemas (5.3.1) e (5.3.2) vamos denotar esses sistemas por

$$V(F)(x) = N(x,F(x)), \tag{5.3.3}$$

onde $V : \mathcal{H}_1 \to \mathcal{H}_2$ é linear.

Note que o sistema (5.3.1) pode ser escrito como

$$V(F)(x) = A(x)F(x) = N(x, F(x)), \quad N(x, F(x)) = U(x) - \int_0^x R(x,t)G(F(t))dt.$$

Como feito anteriormente, vamos supor que as equações (5.3.1) e (5.3.2), ou de forma mais geral a equação (5.3.3), possuem solução única (veja a Observação 5.1.6). Por simplicidade de notação, consideramos os sistemas 2×2, onde:

$$V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} : W[0,1] \oplus W[0,1] \to W_1[0,1] \oplus W_1[0,1],$$

com $v_{ij} : W[0,1] \to W_1[0,1]$ linear,

$$U(x) = \begin{pmatrix} u_1(x) \\ u_2(x) \end{pmatrix} \in W_1[0,1] \oplus W_1[0,1],$$

e

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} \in W[0,1] \oplus W[0,1].$$

Exemplo 5.3.1 *Considere o sistema de equações não lineares de Volterra:*

$$\begin{cases} (x+1)f_1(x) - \displaystyle\int_0^x (x^2 - t)(f_1(t))^2 dt - \int_0^x (x^2 - t)cos(f_2(t))dt = x + \frac{x^3}{2} + \frac{x^4}{12} - \frac{x^5}{5} \\ \\ -\displaystyle\int_0^x t(f_1(t))^3 dt + (e^x + 4)f_2(x) - \int_0^x tsen(f_2(t))dt = x^2 - \frac{x^3}{3} - \frac{x^4}{4} \end{cases}$$

para $x \in [0,1]$.

Nesse sistema temos

$$V(F)(x) = \begin{pmatrix} x+1 & 0 \\ 0 & e^x+4 \end{pmatrix} \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix},$$

e

$$N(x,F(x)) = \begin{pmatrix} \displaystyle\int_0^x (x^2-t)(f_1(t))^2dt + \int_0^x (x^2-t)cos(f_2(t))dt + x + \frac{x^3}{2} + \frac{x^4}{12} - \frac{x^5}{5} \\ \displaystyle\int_0^x t(f_1(t))^3dt - (e^x+4)f_2(x) + \int_0^x tsen(f_2(t))dt + x^2 - \frac{x^3}{3} - \frac{x^4}{4} \end{pmatrix},$$

segue da observação 5.1.6 que esta equação possui solução única.

Exemplo 5.3.2 *Para terminar a seção apresentamos um sistema de equações integro-diferenciais não lineares de Fredholm-Volterra (adaptado do artigo [18]). Para* $x \in [0;1]$, *o sistema*

$$\begin{cases} f_1'(x) + f_1(x) - \dfrac{1}{4}\displaystyle\int_0^1 t(f_1(t))^3dt + \dfrac{1}{2}\int_0^x x(f_1(t))dt = \dfrac{x^4}{6} + x^2 + 2x - \dfrac{1}{32} \\ \\ f_2'(x) + x(f_2(x))^2 - \displaystyle\int_0^1 (x+t)(1+(f_2(t))^2)dt - \int_0^x xcos(f_2(t))dt = -xsen(x) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + x^3 - \dfrac{4x}{3} + \dfrac{3}{4} \\ \\ f_1(0) = 0 \\ f_2(0) = 0 \end{cases},$$

tem solução

$$F(x) = \begin{pmatrix} x^2 \\ x \end{pmatrix}.$$

Nesse sistema temos

$$V(F)(x) = \begin{pmatrix} f_1'(x) + f_1(x) \\ f_2'(x) \end{pmatrix},$$

$$N(x,F(x)) = \begin{pmatrix} \dfrac{1}{4}\displaystyle\int_0^1 t(f_1(t))^3 dt - \dfrac{1}{2}\int_0^x x(f_1(t))dt + \dfrac{x^6}{10} + x^2 + 2x - \dfrac{1}{32} \\ \\ -x(f_2(x))^2 + \displaystyle\int_0^1 (x+t)(1+(f_2(t))^2)dt + \int_0^x xcos(f_2(t))dt - \\ \hfill -xsen(x) + x^3 - \dfrac{4x}{3} + \dfrac{1}{4} \end{pmatrix}$$

5.3.1 Método de resolução: extensão para o caso geral

A seguir vamos adaptar o método envolvendo a teoria de espaços de reprodução, que usamos para encontrar uma aproximação da solução da Equação (5.1.1), para aproximar a solução da Equação (5.3.3). Em particular, podemos aproximar a solução de (5.3.1) e (5.3.2).

Observação 5.3.3 *Note que os lemas 5.2.2 e 5.2.3 e os teoremas 5.2.4, 5.2.5 e 5.2.8 e o corolário 5.2.9 continuam válidos para as equações (5.3.1) e (5.3.2), já que estamos supondo a unicidade de solução. O teorema 5.2.6 precisa apenas de um ajuste no enunciado, que faremos a seguir, mas tem demonstração análoga.*

Teorema 5.3.4 *Se* $\{x_i\}_{i=1}^{\infty}$ *é denso em* $[0,1]$, *então a única solução da equação (5.3.3) é dada por:*

$$F(x) = \sum_{i=1}^{\infty}\left(\sum_{k=1}^{i}\beta_{ik}\alpha_k\right)\overline{A_i}(x) \tag{5.3.4}$$

Onde

$$\alpha_k = \langle F, A_k\rangle = \begin{cases} N_1\left(x_{\frac{k+1}{2}}, F(x_{\frac{k+1}{2}})\right), & k \text{ é ímpar} \\ N_2\left(x_{\frac{k}{2}}, F(x_{\frac{k}{2}})\right), & k \text{ é par.} \end{cases}$$

Demonstração: Suponhamos que $F(x)$ seja a solução da equação (5.3.3). Vimos que $B = \{\overline{A_1}, \overline{A_2}, \cdots\}$ é uma base ortonormal de $W[0,1] \oplus W[0,1]$, onde $\overline{A}(x)$ é conforme a equação (5.2.1). Assim, pela identidade de Parseval (Veja o teorema 2.1.20), podemos escrever

$$F(x) = \sum_{i=1}^{\infty}\langle F, \overline{A_i}\rangle\overline{A_i}(x).$$

Agora, note que:

$$
\begin{aligned}
\langle F,\overline{A_i}\rangle &= \left\langle F, \sum_{k=1}^{i} \beta_{ik} A_k \right\rangle \\
&= \sum_{k=1}^{i} \beta_{ik} \langle F,A_k\rangle \\
&= \sum_{k=1}^{i} \beta_{ik}\alpha_k \\
&= \sum_{k=1,\text{k é ímpar}}^{i} \beta_{ik} v_{1k} f_1(x_{\frac{k+1}{2}}) + \sum_{k=1,\text{k é par}}^{i} \beta_{ik} v_{2k} f_2(x_{\frac{k}{2}}) \\
&= \sum_{k=1,\text{k é ímpar}}^{i} \beta_{ik} N_1(x_{\frac{k+1}{2}}, F(x_{\frac{k+1}{2}})) + \sum_{k=1,\text{k é par}}^{i} \beta_{ik} N_2(x_{\frac{k}{2}}, F(x_{\frac{k}{2}}))
\end{aligned}
$$

$\square$

5.3.2 Método iterativo para o caso não linear

Observamos que o teorema 5.3.4 nos dá uma forma direta de obter uma solução para a equação (5.3.3), para o caso em que $N(x, F(x)) = U(x)$, ou seja, quando N não depende de F, que é o caso em que a equação em questão é linear. Essa aproximação é dada por

$$F_n(x) = \sum_{i=1}^{n} \left(\sum_{k=1}^{i} \beta_{ik}\alpha_k \right) \overline{A_i}(x) = \begin{pmatrix} f_{1n}(x) \\ f_{2n}(x), \end{pmatrix}$$

com convergência uniforme para a solução exata $F(x)$, como no teorema 5.2.8, com α_k dado pelo teorema 5.3.4 e o corolário 5.2.9. O problema que vamos discutir nessa seção final é o caso em que N depende de F. Neste caso, vamos utilizar um método iterativo para obter uma boa aproximação de α_k e obter assim uma boa aproximação de $F_n(x)$ e por sua vez de $F(x)$. De acordo com o teorema 5.3.4, a representação da solução de (5.3.1) pode ser denotada por

$$F(x) = \sum_{i=1}^{\infty} C_i \overline{A_i}(x),$$

onde $C_i = \sum_{k=1}^{i} \beta_{ik}\alpha_k$. De fato, α_k é desconhecido e nós aproximaremos C_i usando uma função conhecida B_i. Definiremos inicialmente $F_0(x_1) = 0$ e os próximos termos de F_n serão dados por:

$$F_n(x) = \sum_{i=1}^{n} B_i \overline{A_i},$$

onde os coeficientes B_i são dados por

$$\begin{aligned}
B_1 &= \beta_{11} N(x_1, F_0(x_1)) \\
F_1(x) &= B_1 \overline{A_1}(x), \\
B_2 &= \sum_{k=1}^{2} \beta_{2k} N(x_k, F_{k-1}(x_k)), \\
F_2(x) &= \sum_{i=1}^{2} B_i \overline{A_i}(x), \\
&\vdots \\
F_{n-1}(x) &= \sum_{i=1}^{n-1} B_i \overline{A_i}(x), \\
B_n &= \sum_{k=1}^{n} \beta_{nk} N(x_k, F_{k-1}(x_k)).
\end{aligned}$$

Podemos obter uma aproximação de $F(x)$ truncando a série obtida por esse método iterativo. Dessa forma, a aproximação é dada por

$$F_n^J(x) = \sum_{i=1}^{J} \sum_{k=1}^{i} \beta_{ik} N(x_k, F_{n-1}(x_k)) \overline{A_i}(x). \quad (5.3.5)$$

A seguir, provaremos que $F_n(x)$ conforme obtido no método iterativo converge para a solução exata $F(x)$.

Teorema 5.3.5 *Seja $N(x, u)$ contínua para $x \in [0,1]$ e $u \in (-\infty, +\infty)$. Se $\|F_n(x) - F(x)\| \to 0$ quando $x_n \to y$ e $n \to \infty$ então*

$$N(x_n, F_{n-1}(x_n)) \to N(x, F(x)).$$

Demonstração: Note que $f_{n-1}(x_n) \to f(y)$. De fato,

$$\begin{aligned}
|f_{n-1}(x_n) - f(y)| &= |f_{n-1}(x_n) - f_{n-1}(y) + f_{n-1}(y) - f(y)| \\
&\leq |f_{n-1}(x_n) - f_{n-1}(y)| + |f_{n-1}(y) - f(y)| \\
&\leq \langle f_{n-1}, K(\cdot,x_n) - K(\cdot,y)\rangle + |f_{n-1}(y) - f(y)| \\
&\leq \|f_{n-1}\| \|K(\cdot,x_n) - K(\cdot,y)\| + |f_{n-1}(y) - f(y)|
\end{aligned}$$

Pela simetria do núcleo K temos que $\|K(\cdot,x_n) - K(\cdot,y)\| \to 0$. Portanto, podemos concluir que $|f_{n-1}(x_n) - f(y)| \to 0$, ou seja, $f_{n-1}(x_n) \to f(y)$.

Pela continuidade da função N podemos concluir o desejado. $\square$

Teorema 5.3.6 *Da maneira que foi definido em* (5.3.5) *temos que* $\{F_n\}_{i=1}^{\infty}$ *é não decrescente.*

Demonstração: Observe que

$$\begin{aligned}
\|F_n\|^2 &= \sum_{i=1}^{n} |\left\langle B_i\overline{A}_i, \overline{A}_i \right\rangle|^2 \\
&= \sum_{i=1}^{n} B_i^2 |\langle \overline{A}_i, \overline{A}_i\rangle|^2 \\
&= \sum_{i=1}^{n} B_i^2 \|\overline{A}_i\|^2 \\
&= \sum_{i=1}^{n} (B_i)^2.
\end{aligned}$$

Dessa forma, $\|F_n\|$ é não decrescente. $\square$

Teorema 5.3.7 *Nas mesmas condições anteriores* $V(F_n)(x_j) = N(x_j,F_{j-1}(x_j))$, $j \leq n$. *Além disso,* $V(F_n)(x_j) = V(F)(x_j)$, $j \leq n$.

Demonstração: Se $j \leq n$ então

$$\begin{aligned}
V(F_n)(x_j) &= V\left(\sum_{i=1}^{n} B_i\overline{A_i}\right)(x_j) \\
&= \sum_{i=1}^{n} B_i V(\overline{A}_i)(x_j)
\end{aligned}$$

Assim,

$$\begin{aligned} V(F_n)(x_j) &= \sum_{i=1}^{n} B_i \langle V(\overline{A}_i), K(x_j, \cdot) \rangle \\ &= \sum_{i=1}^{n} B_i \langle \overline{A}_i, V^*(K(x_j, \cdot)) \rangle \\ &= \sum_{i=1}^{n} B_i \langle \overline{A}_i, A_j \rangle \end{aligned}$$

Ou seja,

$$V(F_n)(x_j) = \sum_{i=1}^{n} B_i \langle \overline{A}_i, A_j \rangle.$$

Multiplicando a igualdade acima por β_{ji} e usando a ortogonalidade de $\overline{A}_i$, temos que

$$\begin{aligned} \sum_{l=1}^{j} \beta_{jl} V(F_n)(x_l) &= \sum_{l=1}^{j} \beta_{jl} \sum_{i=1}^{n} B_i \langle \overline{A}_l, A_j \rangle \\ &= \sum_{i=1}^{n} B_i \langle \overline{A}_i, \sum_{l=1}^{j} \beta_{jl} A_l \rangle \\ &= \sum_{i=1}^{n} \sum_{l=1}^{i} B_i \langle \overline{A}_i, \overline{A}_l \rangle \\ &= B_j \\ &= \sum_{l=1}^{j} \beta_{jl} N(x_l, F_{l-1}(x_l)). \end{aligned}$$

Se $j = 1$, então

$$V(F_n)(x_1) = N(x_1, F_0(x_1)).$$

Se $j = 2$, então

$$\begin{aligned} V(F_n)(x_2) &= \beta_{21} N(x_1, F_0(x_1)) + \beta_{22} N(x_2, F_1(x_2)) \\ &= \beta_{21} V(F_n)(x_1) + \beta_{22} V(F_n)(x_2). \end{aligned}$$

Ou seja, comparando com o caso $j = 1$, temos

$$V(F_n)(x_2) = N(x_2, F_1(x_2)).$$

Prosseguindo com esse raciocínio podemos concluir que

$$V(F_n)(x_j) = N(x_j,F_{j-1}(x_j)).$$

$\square$

Teorema 5.3.8 *Suponha que $\{F_n\}$ seja limitada. Se $\{x_i\}_{i=1}^{\infty}$ é denso em $[0,1]$ então $F_n(x)$ converge para a solução exata $F(x)$.*

Demonstração: Observe que $F_n(x)$ converge, pois

$$F_{n+1}(x) = F_n(x) + B_{n+1}\overline{A}_{n+1}(x)$$

Dessa forma,

$$\begin{aligned} \|F_{n+1}\|^2 &= \|F_n\|^2 + (B_{n+1})^2 \\ &= \|F_{n-1}\|^2 + (B_n)^2 + (B_{n+1})^2 \\ &= \vdots \\ &= \|F_0\|^2 + \sum_{i=1}^{n+1}(B_i)^2. \end{aligned}$$

Pelo teorema 5.3.6, e como $\{F_n\}$ limitada segue que $\{F_n\}$ é convergente. Assim, existe uma constante c tal que $\sum_{i=1}^{\infty}(B_i)^2 = c$. Segue que

$$B_i = \sum_{i=1}^{i} \beta_{ik} N(x_k,F_{k-1}(x_k)) \in l^2, i = 1,2\cdots .$$

Seja $m > n$, temos que $(F_m - F_{m-1}) \perp (F_{m-1} - F_{m-2}) \perp \cdots \perp (F_{n+1} - F_n)$. Em consequência, temos que

$$\begin{aligned} \|F_m - F_n\|^2 &= \|F_m - F_{m-1} + F_{m-1} - \cdots + F_{n+1} - F_n\|^2 \\ &= \|F_m - F_{m-1}\|^2 + \cdots + \|F_{n+1} - F_n\|^2. \end{aligned}$$

Note que $\|F_m - F_{m-1}\|^2 = (B_m)^2$, assim podemos escrever

$$\|F_m - F_n\|^2 = \sum_{i=n+1}^{m} (B_i)^2.$$

E ainda, $\displaystyle\sum_{i=n+1}^{m} (B_i)^2 \to 0$ quando $m,n \to \infty$. Dessa forma, como $W[0,1]$ é um espaço completo, existe $F(x) \in W[0,1]$ tal que $F_n(x) \to F(x)$. Agora, nos resta mostrar que $F(x)$ é a solução de (5.3.2).

Por hipótese, $\{x_i\}$ é denso em [0,1]. Então, existe uma subsequência $\{x_{n_j}\}$ tal que $x_{n_j} \to x$. Usando o teorema (5.3.7 podemos concluir que $V(F)(x_{n_j}) = N(x_{n_j}, F_{n_j-1}(x_{n_j}))$. Então, fazendo $j \to \infty$ e usando o teorema 5.3.5, temos que $V(F)(x) = N(x,F(x))$. Portanto, $F_n(x)$ converge para a solução $F(x)$. $\square$

Teorema 5.3.9 *Seja $r_n(x)$ a diferença entre a solução aproximada $F_n(x)$ e a solução exata $F(x)$. Então $r_n(x)$ é não crescente.*

Demonstração: Note que

$$\begin{aligned}
\|r_n\|^2 &= \|F_n - F\|^2 \\
&= \sum_{i=n+1}^{\infty} \left|\left\langle B_i\overline{A_i}, \overline{A_i}\right\rangle\right|^2 \\
&= \sum_{i=n+1}^{\infty} (B_i)^2 |\langle \overline{A_i}, \overline{A_i}\rangle|^2 \\
&= \sum_{i=n+1}^{\infty} (B_i)^2 \|\overline{A_i}\|^2 \\
&= \sum_{i=n+1}^{\infty} (B_i)^2.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\|r_{n-1}\|^2 &= \|F_{n-1} - F\|^2 \\
&= \sum_{i=n}^{\infty} \left|\left\langle B_i\overline{A_i}, \overline{A_i}\right\rangle\right|^2
\end{aligned}$$

Segue que,

$$
\begin{aligned}
\|r_{n-1}\|^2 &= \sum_{i=n}^{\infty}(B_i)^2|\langle\overline{A_i},\overline{A_i}\rangle|^2 \\
&= \sum_{i=n}^{\infty}(B_i)^2\|\overline{A_i}\|^2 \\
&= \sum_{i=n}^{\infty}(B_i)^2.
\end{aligned}
$$

Dessa forma,

$$\|r_n\|^2 \leq \|r_{n-1}\|^2.$$

Consequentemente, $r_n(x)$ é não crescente. $\square$

6 Conclusão

As equações diferenciais e integrais modelam diversos fenômenos físicos e biológicos. Logo a resolução de tais equações torna-se necessária para diversos estudos e pesquisas. Os métodos tradicionais de resolução de equações não abrangem sua totalidade. Como alternativa, temos o estudo qualitativo das soluções, porém apenas o comportamento das soluções é estudado neste caso e sua forma analítica é desconhecida. Neste trabalho, a solução analítica de equações integrais de Volterra foi encontrada através da teoria de núcleos positivos definidos e espaços de reprodução. Em um primeiro momento, abordamos apenas equações lineares e então expandimos o conhecimento adquirido para equações não lineares. Embora tenhamos ficado na parte teórica em um primeiro momento, de acordo com [3], a aproximação das soluções de equações integrais obtidas com a teoria de espaços de Hilbert de reprodução e núcleos positivos definidos tem uma precisão maior do que a aproximação obtida com outros métodos. Além disso, sua programação é mais simples do que as de outros métodos. Assim, a teoria de espaços de Hilbert de reprodução é de grande valia na resolução de equações diferenciais e integrais, visto que fornece uma aproximação satisfatória e de fácil programação. Entendemos porém que esse método de resolução de equações é bastante recente e, por isso, ainda pode contribuir muito para o entendimento e resolução de diversos problemas conhecidos e por surgir. Mesmo assim, a gama de resultados e problemas que estão relacionados a espaços de reprodução é muito ampla e é necessário mais tempo que o de um mestrado para poder visualizar bem toda a extensão dessa área, que inclui aspectos teóricos, de computação e aplicações, que estão intimamente ligados (veja [12], por exemplo).

Dessa forma, dentre as possibilidades que surgiram para este trabalho, simulação numérica e estudos teóricos de equações não lineares, optamos pela segunda opção. Ou seja, continuar com os estudos teóricos, fazendo então o estudo de métodos adaptados para tratar de equações não lineares, seguindo por exemplo o trabalho [18]. Isso é importante pelo ponto de vista de possíveis aplicações e problemas reais, como em modelos Físicos e Biológicos que são dados por equações não lineares (veja um modelo de epidemia tratado em [19]).

REFERÊNCIAS

[1] OLIVEIRA, C. R., *Introdução à análise funcional*. Rio de Janeiro: IMPA, 2012. 257p.

[2] ARONSZAJN, N., Theory of reproducing kernels. *Trans. Amer. Math Soc.*, EUA, v.68, p. 337-404, 1950.

[3] YANG, L.H.; SHEN, J.H.;WANG, Y.; The reproducing kernel method for solving the system of the linear Volterra integral equations with variable coefficients, *Journal of Computational and Applied Mathematics*, v.236, Issue 9, p. 2398-2405, 2012.

[4] JIANG, W.; CHEN, Z.; Solving a system of linear Volterra integral equations using the new reproducing kernel method, *Applied Mathematics and Computation*, China, v.219, p. 10225-10230, 2003.

[5] YANG, L.H.; LIN, Y.; Reproducting Kernel Methods for Solving Linear Initial-Boundary-Value Problems. *Electronic Journal of Differential Equations*, China, v.2008, n.29, p. 1-11, 2008.

[6] YANG, L.H.; LI, H.Y.; WANG, J.R.; Solving a system o linear Volterra integral equation using the modified reproducing kernel method, *Abstract and Applied Analysis*, China, v.2013, ID 196308, p. 1-5, 2013.

[7] ISNARD, C. *Introdução à medida e integração*. Rio de Janeiro: IMPA, 2007. 314p.

[8] LIMA, E.L. *Espaços Métricos*. Rio de Janeiro: IMPA, 2009. 299p.

[9] COELHO, F.U.; LOURENÇO, L.L., *Um curso de álgebra linear*. 2.ed. São Paulo: Editora da universidade de São Paulo, 2013. 272p.

[10] FERNANDEZ, R.; *Introdução à teoria da medida*. Projeto Euclides. Rio de Janeiro: IMPA, 2002.

[11] FOLLAND, G. B.; *Real analysis: Modern techniques and their applications*, John Wiley & Sons, 1999.

[12] FASSHAUER, G. E. *Positive Definite Kernels: Past, Present and Future*, Dolomites Research Notes on Approximation, v. 4, Special Issue, p. 21-63, 2011.

[13] RUGGIERO, M. A. G, LOPES, V. L. R, *Cálculo Numérico e Aspectos Teóricos e Computacionais*, São Paulo: Makron Books, 1997.

[14] FERREIRA, J. C. *Operadores integrais positivos e espaços de Hilbert de reprodução*. 2010. 121f. Tese (Doutorado em Matemática) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2010.

[15] ZILL, D.G. *Equações Diferenciais com aplicações em modelagem*. São Paulo: Pioneira Thomson Learnig, 2003.

[16] MATSUURA, T.; SAITOH, S; Real inversion formulas and numerical experiments of the Laplace transform by using the theory of reproducing kernels. *Procedia Social and Behavioral Sciences*, Japão, v.2, p. 111-119, 2010.

[17] MERCER, J., Functions of positive and negative type and their connection with the theory of integral equations, *Philosophical Transsaction of the Royal Society of London*, Ser. A, v.209, p. 415-446, 1909.

[18] ARQUB, O.A.; SMADI, M., A.; MOMANI, S.; Application of Reproducing Kernel Method for Solving Nonlinear Fredholm-Volterra Integrodifferential Equations, *Abstract and Applied Analysis*, v. 2012, Article ID 839836.

[19] HETHCOTE, H.W.; TUDOR, D.W.; Integral Equation Models for Endemic Infectious Diseases. *Jornal of Mathematical Biology*. Springer-Verlag. Departament of Mathematics. USA. 1980.