

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Métodos de aprendizado profundo para a escolha do
parâmetro de regularização**

Emanuel Oliveira Souza

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências
de Computação e Matemática Computacional (PPG-C²MC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Emanuel Oliveira Souza

Métodos de aprendizado profundo para a escolha do parâmetro de regularização

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

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Elias Salomão Helou Neto

USP – São Carlos
Julho de 2024

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

O48m Oliveira Souza, Emanuel
Métodos de aprendizado profundo para a escolha
do parâmetro de regularização / Emanuel Oliveira
Souza; orientador Elias Salomão Helou Neto. -- São
Carlos, 2024.
81 p.

Dissertação (Mestrado - Programa de Pós-Graduação
em Ciências de Computação e Matemática
Computacional) -- Instituto de Ciências Matemáticas
e de Computação, Universidade de São Paulo, 2024.

1. Problemas Inversos. 2. Regularização. 3.
Aprendizado Profundo. I. Salomão Helou Neto, Elias,
orient. II. Título.

Emanuel Oliveira Souza

Deep Learning methods for choosing the regularization
parameter

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

Concentration Area: Computer Science and
Computational Mathematics

Advisor: Prof. Dr. Elias Salomão Helou Neto

USP – São Carlos
July 2024

AGRADECIMENTOS

A Deus, pelo seu infinito amor e por guiar meus passos.

Aos meus pais, pelo amor, incentivo e confiança ao longo desses anos.

Aos meus familiares e amigos, pelo apoio e amizade apesar do distanciamento.

Aos meus colegas de pós-graduação, por todos os momentos compartilhados.

A todos os professores com os quais tive o prazer de estudar.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao DM Service pelo apoio financeiro.

Ao meu orientador, Elias, por ter me dado suporte indispensável e nunca ter me deixado desanimar.

Muito obrigado pela oportunidade e dedicação.

Por fim, a todos o meu muito obrigado.

“O futuro pertence àqueles que se prepararam hoje para ele.”
(Malcolm X)

RESUMO

SOUZA, E. O. **Métodos de aprendizado profundo para a escolha do parâmetro de regularização**. 2024. 81 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

Os problemas inversos costumam aparecer em várias áreas das ciências e engenharias e, de certo modo, consistem em determinar causas desconhecidas a partir de efeitos observados. Alguns exemplos de aplicações de problemas inversos são encontrados em reconstrução/restauração de imagens, engenharia civil e finanças quantitativas. As principais características dos problemas inversos são a não unicidade e a instabilidade de suas soluções. Uma abordagem frequente é utilizar técnicas chamadas de regularização, especialmente a regularização de Tikhonov, uma das mais utilizadas nesses casos. Geralmente, os métodos de regularização estabelecem uma família de operadores parametrizada por um escalar, conhecido como parâmetro de regularização. Uma dificuldade existente nestes métodos é a escolha do parâmetro de regularização adequado que fornece estabilidade e ajuste à solução regularizada. Na literatura, alguns métodos na busca desse parâmetro são utilizados, dentre os quais podemos citar: o princípio da discrepância, validação cruzada generalizada e o método da curva L. Com o acelerado desenvolvimento da Inteligência Artificial, o surgimento de novas técnicas baseadas em redes neurais são aplicadas com sucesso em vários problemas reais. Sendo assim, neste trabalho é proposto a criação de uma nova técnica que utiliza aprendizado profundo para selecionar o melhor parâmetro de regularização. Os resultados mostram que este tipo de abordagem supera as técnicas tradicionais de seleção de parâmetro. Isto representa um avanço na área de seleção de parâmetros. Sendo assim, a abordagem utilizada proporcionou melhorias significativas nos experimentos numéricos utilizados.

Palavras-chave: Problemas inversos, regularização, aprendizado profundo.

ABSTRACT

SOUZA, E. O. **Deep Learning methods for choosing the regularization parameter**. 2024. 81 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

Inverse problems tend to appear in various areas of science and engineering and, in a way, consist of determining unknown causes from observed effects. Some examples of applications of inverse problems are found in image reconstruction/restoration, civil engineering and quantitative finance. The main characteristics of inverse problems are the non-uniqueness and instability of their solutions. A common approach is to use techniques called regularization, especially Tikhonov regularization, one of the most used in the cases. Generally, regularization methods establish a family of operators parameterized by a scalar, known as the regularization parameter. A difficulty in these methods is the choice of the appropriate regularization parameter that provides stability and adjustment to the regularized solution. In the literature, some methods are used to search for this parameter, among which we can mention: the discrepancy principle, generalized cross-validation and the L-curve method. With the accelerated development of Artificial Intelligence, the emergence of new techniques based on neural networks are successfully applied to several real problems. Therefore, this work proposes the creation of a new technique that uses deep learning to select the best regularization parameter. The results show that this type of approach outperforms traditional parameter selection techniques. This represents a step forward in the area of parameter selection. Therefore, the approach used provided significant improvements in the numerical experiments used.

Keywords: Inverse problems, regularization, Deep Learning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Geometria da transformada de Radon.	27
Figura 2 – Exemplo de uma plotagem da curva-L	42
Figura 3 – Neurônio artificial	46
Figura 4 – Estrutura de uma rede neural convolucional simples.	47
Figura 5 – A parte esquerdo da figura demonstra o erro relativo no conjunto de validação para os dois métodos. A parte direita ilustra a dispersão das previsões dos valores do parâmetro de regularização.	53
Figura 6 – Erro relativo entre os métodos de seleção do parâmetro de regularização para o problema inverso do calor.	54
Figura 7 – Exemplos de imagens contidas no conjunto BSD68.	55
Figura 8 – <i>Patches</i> obtidos de uma imagem do conjunto BSD68.	55
Figura 9 – Arquitetura proposta para o problema de estimar o parâmetro de regularização para o problema de <i>deblurring</i>	56
Figura 10 – Dispersão dos valores estipulados pela rede para o problema de <i>deblurring</i> no conjunto de validação para a função ℓ_1 (esquerdo) e ℓ_2 (direito).	57
Figura 11 – Imagens de testes utilizadas para o problema de <i>deblurring</i>	57
Figura 12 – Resultados obtidos para a imagem <i>peppers</i> nas duas abordagens utilizando um erro de 2,5%.	58
Figura 13 – Resultados obtidos para a imagem <i>butterfly</i> nas duas abordagens utilizando um erro de 2,5%.	58
Figura 14 – Gráficos de densidade para a imagem <i>butterfly</i> . O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.	58
Figura 15 – Gráficos de densidade para a imagem <i>camera-man</i> . O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.	59
Figura 16 – Gráficos de densidade para a imagem <i>peppers</i> . O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.	59
Figura 17 – Exemplos de <i>Sheep-Logan</i> contidas no conjunto de dados.	60
Figura 18 – Exemplos de sinogramas contidas no conjunto de dados.	60
Figura 19 – Arquitetura proposta para o problema de estimar o parâmetro de regularização para o problema de tomografia computadorizada.	61

Figura 20 – Dispersão dos valores estipulados pela rede para o problema de tomografia computadorizada no conjunto de validação para a função ℓ_1 (esquerdo) e ℓ_2 (direito).	61
Figura 21 – Imagem do <i>camera man</i> utilizado para o problema de tomografia computadorizada.	62
Figura 22 – Resultados obtidos na reconstrução do <i>camera man</i> para o sinograma com ruído de 5%.	62
Figura 23 – Gráfico de densidade para a tomografia computadorizada do <i>camera man</i> . A densidade da métrica SSIM é exibida à esquerda, enquanto a da métrica PSNR é mostrada à direita.	62

LISTA DE ALGORITMOS

Algoritmo 1 – Seleção do parâmetro de regularização utilizando aprendizado profundo. 50

LISTA DE TABELAS

Tabela 1 – Resumo dos hiperparâmetros para o problema do calor.	53
Tabela 2 – Resumo dos hiperparâmetros para o problema de <i>deblurring</i>	57
Tabela 3 – Resumo dos hiperparâmetros para o problema de tomografia computadorizada.	61

SUMÁRIO

1	INTRODUÇÃO	21
2	PROBLEMAS INVERSOS	23
2.1	Definição	23
2.2	Equação do calor	24
2.3	<i>Deblurring</i>	25
2.4	Tomografia Computadorizada	26
3	REGULARIZAÇÃO	29
3.1	Operador inverso de Moore-Penrose	29
3.2	Operador de regularização	37
3.3	Regularização de Tikhonov	39
3.4	Métodos para escolha do parâmetro de regularização	41
3.4.1	<i>Princípio da discrepância</i>	41
3.4.2	<i>Validação cruzada generalizada</i>	42
3.4.3	<i>Curva-L</i>	43
4	APRENDIZADO DE MÁQUINA	45
5	ABORDAGEM PROPOSTA	49
6	EXPERIMENTOS NUMÉRICOS	51
6.1	Especificações	51
6.2	Calor	51
6.3	<i>Deblurring</i>	54
6.4	Tomografia Computadorizada	59
7	CONCLUSÕES	65
7.1	Considerações finais	65
7.2	Limitações e melhorias futuras	65
	REFERÊNCIAS	67
	APÊNDICE A FUNDAMENTOS TEÓRICOS	71
A.1	Espaços vetoriais normados	71

A.2	Operadores Lineares	76
------------	--------------------------------------	-----------

INTRODUÇÃO

Os problemas inversos pertencem a uma área da Matemática interessada em criar modelos baseados em problemas reais. Além disso, sempre estão associados a um problema direto. Com isso, definimos que dois problemas são inversos um do outro se a formulação de um é a inversa do outro. No entanto, não existe um consenso qual é o problema direto e, qual é o inverso. A referência Keller (1976) define que o problema direto é o que foi estudado mais cedo e, o outro é o inverso. Em outras palavras, o objetivo do problema direto é determinar a consequência, conhecendo-se a causa e o modelo. Em razão disso, o inverso associado ao direto é obter a causa, observando-se a consequência.

Por exemplo, em tomografia por raios-x, o problema direto é determinar as imagens projetadas pelos raios-x atravessando um corpo físico do qual a estrutura interna é conhecida. O problema inverso é reconstruir a estrutura interna de um corpo físico desconhecido a partir de informações tiradas dos raios-x de diferentes posições. Atualmente, existem muitos exemplos de problemas inversos no nosso cotidiano, destacamos o problema de restauração de imagens (OSHER *et al.*, 2005; NATTERER; WÜBBELING, 2001), ressonância magnética (HASHEMI; BRADLEY; LISANTI, 2010) e de reconstruções de imagens tomográficas (KAK; SLANEY, 2001).

A principal dificuldade dos problemas inversos é oriunda do fato que a maioria deles é de problemas ditos mal-postos no sentido de Hadamard (HADAMARD, 2003). No geral, os problemas inversos lutam contra a instabilidade devido aos dados fornecidos terem erros de medida, um fato muito comum por se tratar de dados reais. Uma abordagem usual para contornar este problema é utilizar técnicas de regularização (ENGL; HANKE; NEUBAUER, 1996) para calcular uma solução. Um método de regularização bastante utilizado é a regularização de Tikhonov (TIKHONOV; ARSENIN, 1977), que substitui o modelo original por um problema que necessita de um escalar chamado de parâmetro de regularização. A escolha deste parâmetro é crucial para fornecer uma solução regularizada conveniente (HANSEN, 1992). Muitos métodos

foram criados para selecionar esse parâmetro, destacamos o princípio da discrepância (PHILLIPS, 1962), validação cruzada generalizada (GOLUB; HEATH; WAHBA, 1979) e curva-L (HANSEN; O'LEARY, 1993). Métodos diferentes aplicados em um mesmo problema produzem parâmetros distintos. Assim, escolher um método específico é bastante difícil.

O rápido desenvolvimento do *hardware* e das tecnologias de *software* permitiram o uso continuado e bem-sucedido do aprendizado de máquina em aplicações práticas. Isto forneceu uma nova ferramenta para se construir novos métodos para selecionar o parâmetro de regularização utilizando conceitos de aprendizado de máquina. Na literatura, é possível encontrar trabalhos recentes que utilizam estes conceitos, pode-se destacar Liu e Zhang (2021) que usa um conjunto de dados sintéticos e um modelo de regressão pré-treinado para obter o parâmetro em problemas inversos sísmicos; e Afkham, Chung e Chung (2021), que utilizaram redes neurais profundas para selecionar o parâmetro de regularização em diferentes aplicações.

Neste trabalho, o objetivo é fornecer um novo algoritmo para selecionar o parâmetro da regularização de Tikhonov utilizando redes neurais profundas e compará-lo com os métodos que utilizam o mesmo princípio. Esta abordagem leva grande vantagem em relação aos métodos tradicionais, como pode ser vistas em Liu e Zhang (2021) e Afkham, Chung e Chung (2021).

O restante deste trabalho está organizado como segue. No Capítulo 2 apresentamos a definição de problema inverso e alguns dos exemplos mais conhecidos. No Capítulo 3 definimos o operador de Moore-Penrose e introduzimos o conceito de regularização. Posteriormente, no Capítulo 4 é inserida uma pequena introdução sobre aprendizado de máquina. No Capítulo 5 é apresentado o método proposto neste trabalho. Finalmente, no Capítulo 6 são apresentados os resultados numéricos e as conclusões no Capítulo 7.

PROBLEMAS INVERSOS

Neste capítulo apresentamos uma definição de problemas inversos e como motivação introduzimos os seguintes exemplos de problemas inversos: equação do calor, *deblurring* e a tomografia computadorizada.

2.1 Definição

Como visto anteriormente, a maioria dos problemas inversos estudados são mal-postos. Dizemos que um problema é bem-posto no sentido de Hadamard ([HADAMARD, 2003](#)) se satisfizer as seguintes condições:

- i - Para todos os dados admissíveis, a solução existe;
- ii - Para todos os dados admissíveis, a solução é única;
- iii - A solução depende continuamente dos dados.

Se um determinado problema não satisfaz alguma das condições acima, ele é dito mal-posto. Segundo [Hadamard \(2003\)](#), a violação das condições (i) e (ii) ocorre com mais frequência. Uma das maneiras para se tratar o problema com a restrição (i) é relaxar um pouco o conceito da existência de solução, por exemplo, tomar como base uma aproximação da solução. No caso em que a solução não é única, a estratégia que pode ser escolhida é atribuir informações adicionais ao modelo, com o intuito de obter uma solução específica, ou simplesmente escolher a solução que contém a menor norma. Agora, se o problema não depende continuamente dos dados, isto torna o problema sensível e surgirão muitas dificuldades quando as medidas tiverem incertezas.

A partir de agora, consideraremos um problema inverso linear definido em espaços de Hilbert para ilustrar a estrutura matemática que será usada. Para isso, consideramos dois espaços de Hilbert \mathcal{X} e \mathcal{Y} e um operador linear limitado $T : \mathcal{X} \rightarrow \mathcal{Y}$. O problema agora consiste em encontrar $\boldsymbol{x} \in \mathcal{X}$ (causa) que satisfaz a equação:

$$T\mathbf{x} = \mathbf{y},$$

onde $\mathbf{y} \in \mathcal{Y}$ (consequência) é dado. Em suma, resolver o problema inverso corresponde em encontrar o operador inverso, caso exista, e aplicá-lo sobre \mathbf{y} .

Se consideramos a primeira condição de Hadamard, ela é somente satisfeita se $\mathbf{y} \in R(T)$, onde $R(T)$ é o conjunto imagem de T . No caso da segunda, é cumprida se $N(T) = \mathbf{0}$, onde $N(T)$ é o núcleo de T . A terceira condição é mais complicada, pois envolve características do operador T , por exemplo, se T for compacto, o operador inverso T^{-1} existe, mas ele pode não ser contínuo, exceto se \mathcal{X} e \mathcal{Y} forem de dimensão finita. Introduziremos o operador de Moore-Penrose que conseguirá fornecer uma solução que satisfaz as duas primeiras condições e a terceira condição será satisfeita se o operador for contínuo.

2.2 Equação do calor

A equação do calor pertence a uma classe de fenômenos físicos modelados por um processo de difusão. A difusão de calor é um processo contínuo que tem um efeito crescente no tempo e se manifesta como um espalhamento a partir de uma distribuição inicial. Pode-se escrever este processo a partir de um operador:

$$T\mathbf{f} = \mathbf{g}, \tag{2.1}$$

onde $T : \mathcal{X} \rightarrow \mathcal{Y}$ é o operador que representa o processo, \mathbf{f} é a distribuição inicial e \mathbf{g} a distribuição final. No entanto, o problema da equação do calor pode ser representado mais naturalmente por uma equação diferencial parabólica, tendo a forma geral

$$\frac{\partial u(x,t)}{\partial t} = c\nabla^2 u(x,t), \tag{2.2}$$

onde o termo do lado esquerdo denota a taxa de variação da temperatura em relação ao tempo e, o termo do lado direito é o termo de difusão e $c (\geq 0)$ representa a difusividade térmica. Repare que c não precisa ser necessariamente constante, por exemplo, ela pode ser definida por uma função que dependa do tempo (t) ou da posição (x). Caso c seja uma função dependente da distribuição u , isto resulta numa equação não-linear, caso que não será tratado neste trabalho. Utilizando as informações da Eq. (2.1), conseguimos algumas condições:

$$\begin{aligned} u(x,0) &= \mathbf{f}, \\ u(x,T) &= \mathbf{g}. \end{aligned}$$

Geralmente, os problemas clássicos da equação do calor consistem em uma distribuição inicial $u(x,0)$ sendo governada por um modelo no formato da Eq.(2.2). Estes problemas são

resolvidos utilizando métodos numéricos, por exemplo, diferenças finitas e elementos finitos (ver (RECKTENWALD, 2004) e (GOSZ, 2017) para exemplos).

Ponderamos aqui o mesmo problema inverso da equação do calor considerado por Beck, Blackwell e Clair (1985) como sendo o processo de determinar a distribuição inicial em um material a partir da distribuição após um intervalo de tempo decorrido. Com isso, o problema inverso é descrito pela equação

$$\begin{aligned}\frac{\partial u(x,t)}{\partial t} &= c^2 \frac{\partial^2 u(x,t)}{\partial x^2}, \quad x \in (0,L], \quad t \in (0,T]; \\ u(0,t) &= 0, \quad t \in (0,T]; \\ u(L,t) &= 0, \quad t \in (0,T]; \\ u(x,0) &= u(x), \quad x \in [0,L].\end{aligned}$$

onde $u(0,t)$ e $u(L,t)$ são as condições de contorno, $u(x)$ é a função que precisa ser encontrada. Este problema pode ser reformulado por uma equação integral de Fredholm de primeiro tipo (BECK; BLACKWELL; CLAIR, 1985), ficando da forma

$$Au(x) = \int_0^L k(x,y)u(y)dy, \quad (2.3)$$

onde $k(x,y)$ é o *kernel* definido pela série infinita

$$k(x,y) = \frac{2}{L} \sum_{n=1}^{\infty} e^{-\frac{(n\pi c)^2 T}{L^2}} \operatorname{sen}\left(\frac{n\pi x}{L}\right) \operatorname{sen}\left(\frac{n\pi y}{L}\right). \quad (2.4)$$

Se definimos o lado direito da equação por $g(x)$, reescrevemos o problema como uma equação linear da forma $Au = g$.

2.3 Deblurring

Muitas aplicações de astronomia, medicina e biologia utilizam imagens digitais necessárias para responder importantes questionamentos provenientes de problemas reais. No entanto, geralmente estas imagens podem não ter sido geradas com perfeição e vir com degradação ou com desfoque. Uma das técnicas de pós-processamento utilizadas é o *deblurring*, modelado por um problema inverso.

Normalmente, uma imagem em escala de cinza pode ser representada por uma função $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, onde $g(x,y)$ representa a intensidade da luz em (x,y) . Se denotarmos por $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a intensidade real do objeto que estamos estudando, o problema de *deblurring* (BERTERO; BOCCACCI; MOL, 2021) é descrito pela equação

$$g(x,y) := \int \int_D k(x-x', y-y') f(x',y') dx' dy'.$$

Isto estabelece que a imagem borrada g é atingida pela convolução da intensidade do objeto real f , definido em um domínio $D \subset \mathbb{R}^2$, com um *kernel* k chamado de função de espalhamento pontual (PSF, do inglês *point spread function*). Existem vários *kernels* utilizados atualmente, dentre eles podemos citar, *Gaussian blur*, *horizontal motion blur* e *vertical motion blur*.

Com este problema, uma aplicação interessante é resolver o problema inverso associado, isto é, dado o *kernel* k e a imagem borrada g , temos que estimar qual é a imagem original f . Sendo assim, podemos reformular o problema por

$$Kf = g,$$

onde K é um operador compacto (VOGEL, 2002).

2.4 Tomografia Computadorizada

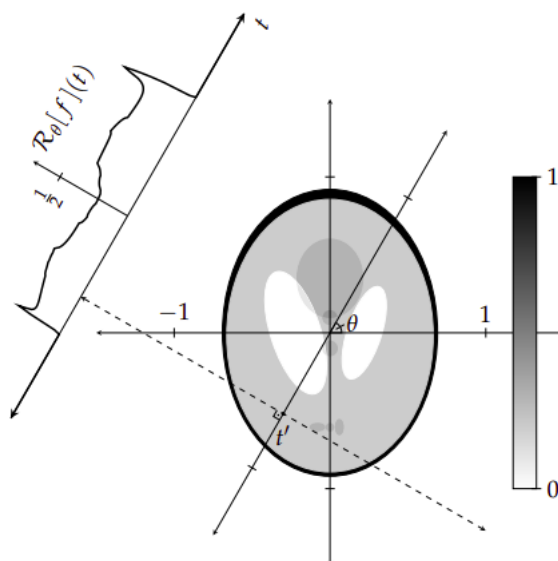
Um dos problemas inversos mais conhecidos é o da Tomografia Computadorizada (CT, do inglês *Computed Tomography*), cuja finalidade é obter a representação de estruturas internas de uma região de forma não-invasiva. O funcionamento da CT é baseado no fato que objetos distintos possuem diferentes taxas de absorção se um raio atravessa ele. Por exemplo, CT pode ser realizada utilizando raio x, raio gama, nêutrons, etc. Para isto, um sensor é utilizado para capturar a intensidade do raio que atravessou o objeto e os dados são processados para reconstruir a estrutura buscada. Uma das principais aplicações da CT é na área médica, onde é usada para escanear alguma parte transversal do corpo para fins de diagnósticos médicos.

O grande problema da CT é reconstruir o objeto, a partir de um conjunto de radiografias. A visualização utilizando apenas uma radiografia não é suficiente para obter uma boa aproximação da estrutura interna de uma região. Sendo assim, o objeto deve ser radiografado em todas as direções. Conforme as propriedades do objeto que está sendo trabalhado, o mesmo apresentará uma resistência a passagem do feixe e o sinal emitido final sofrerá uma atenuação. Com todas as projeções realizadas, o problema de reconstruir se resume à representação das radiografias medidas para o espaço da função $f(x, y)$.

A aquisição das projeções é descrita pela transformada de Radon (RADON, 1917) da função $f(x, y)$, sendo definida pela integral de linha ao longo de uma reta inclinada num dado ângulo θ (ver figura 1). A transformada de Radon pode ser expressa pelo operador:

$$\mathbf{R}[f](\theta, t) := \int_{\mathbb{R}} f\left(t \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + s \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}\right) ds.$$

Figura 1 – Geometria da transformada de Radon.



Fonte: Retirado de Neto (2009)

Com isso, o problema inverso relacionado a CT é obter a função $f(x,y)$ a partir de suas integrais de linhas calculadas em vários ângulos. A transformada de Radon possui uma inversa única com fórmula fechada, porém a unicidade é somente garantida mediante um número infinito de projeções. Para mais detalhes sobre a transformada de Radon e suas propriedades, consulte a referência Natterer (2001).

A função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ pode ser visualizada como uma imagem em diferentes níveis de tons de cinza, variando entre branco e preto. Na prática, é impossível obter uma integral de linha a partir de todas as direções. Essa limitação no processo de medição, implica em uma imagem discretizada. Portanto, o problema de CT pode ser formulado como um sistema de equações lineares

$$A\mathbf{x} = \mathbf{b},$$

onde a matriz A simula o processo feito na CT, o vetor \mathbf{x} representa as intensidades da imagem em escala de cinza e o vetor \mathbf{b} representa o conjunto das projeções medidas, conhecido como sinograma quando representado no plano $\theta \times t$.

Consideramos que a imagem pode ser discretizada em $m \times n$ píxeis e para um ângulo qualquer, assumiremos que o número de projeções varia de $\{1, 2, \dots, n_p\}$. Agora para dado n_a ângulos diferentes, a matriz $A \in \mathbb{R}^{n_p \cdot n_a \times m \cdot n}$, $\mathbf{x} \in \mathbb{R}^{m \cdot n}$ e $\mathbf{b} \in \mathbb{R}^{n_p \cdot n_a}$.

A grande dificuldade é a construção da estrutura e armazenamento da matriz A , visto que é impraticável armazenar computacionalmente e as propriedades da imagem que será reconstruída depende exclusivamente da matriz. Existem alguns métodos comumente utilizados para o cálculo de A como o de Joseph (JOSEPH, 1982) e Siddon (SIDDON, 1985). Neste trabalho utilizaremos o método de Siddon.

REGULARIZAÇÃO

Neste capítulo, definiremos o operador inverso de Moore-Penrose, métodos de regularização e métodos para a escolha do parâmetro de regularização. Primeiramente, é feita uma discussão sobre o operador inverso de Moore-Penrose. Em seguida, abordaremos a regularização, uma abordagem para estabilizar soluções em problemas mal-postos. Especificamente, detalharemos a regularização de Tikhonov. Finalmente, explicitaremos métodos para a seleção do parâmetro de regularização, em particular, o método da discrepância, curva-L e validação cruzada generalizada. Demonstrações e conceitos fundamentais adicionais estão disponíveis no Apêndice A.

3.1 Operador inverso de Moore-Penrose

Lidaremos agora somente com as duas primeiras condições de Hadamard. Para isto, consideraremos o operador $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ e a equação $T\mathbf{x} = \mathbf{y}$. Se $\mathbf{y} \notin R(T)$, $T\mathbf{x} = \mathbf{y}$ não tem solução, sendo assim procuraremos um $\mathbf{x} \in \mathcal{X}$ que minimiza $\|T\mathbf{x} - \mathbf{y}\|_{\mathcal{Y}}$. Agora, se $N(T) \neq \{0\}$, a equação tem infinitas soluções, sendo assim, pegaremos a solução que tiver a menor norma.

Definição 1. Seja $T : \mathcal{X} \rightarrow \mathcal{Y}$ um operador linear limitado, o ponto $\mathbf{x}^* \in \mathcal{X}$ é denominado:

1. solução de quadrados mínimos de $T\mathbf{x} = \mathbf{y}$ se

$$\|T\mathbf{x}^* - \mathbf{y}\|_{\mathcal{Y}} = \inf\{\|T\mathbf{z} - \mathbf{y}\|_{\mathcal{Y}} \mid \mathbf{z} \in \mathcal{X}\};$$

2. melhor solução aproximada de $T\mathbf{x} = \mathbf{y}$, se \mathbf{x}^* é solução de quadrados mínimos de $T\mathbf{x} = \mathbf{y}$ e

$$\|\mathbf{x}^*\|_{\mathcal{X}} = \inf\{\|\mathbf{z}\|_{\mathcal{X}} \mid \mathbf{z} \text{ é solução de quadrados mínimos de } T\mathbf{x} = \mathbf{y}\}.$$

Obviamente, se o operador T é bijetivo, $\mathbf{x}^* = T^{-1}\mathbf{y}$, é a única solução de quadrados mínimos, portanto, a que tem a menor norma. Agora, se T não for bijetivo, temos que especificar para qual $\mathbf{y} \in \mathcal{Y}$ existe uma solução de norma mínima. Para isto, introduziremos um novo operador, conhecido como operador de Moore-Penrose, ou pseudo-inversa, o qual mapeia \mathbf{y} para uma solução de norma mínima correspondente. Isto é executado restringindo o domínio e a imagem do operador T de modo que ele seja invertível.

Definição 2. A inversa generalizada de Moore-Penrose T^\dagger de $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ é definida como a única extensão linear de \tilde{T}^{-1} para

$$D(T^\dagger) := R(T) \oplus R(T)^\perp,$$

onde

$$\tilde{T} := T|_{N(T)^\perp} : N(T)^\perp \rightarrow R(T),$$

com

$$N(T^\dagger) = R(T)^\perp.$$

A restrição de T ao subespaço $N(T)^\perp \subset \mathcal{X}$ e o fato de o contradomínio ser $R(T)$, é suficiente para verificar que \tilde{T} é injetivo e sobrejetivo, ou seja, \tilde{T}^{-1} existe. Ou seja, para qualquer $\mathbf{y} \in D(T^\dagger)$, conseguimos obter uma decomposição única com $\mathbf{y}_1 \in R(T)$ e $\mathbf{y}_2 \in R(T)^\perp$, onde $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$. Como $N(T^\dagger) = R(T)^\perp$,

$$T^\dagger \mathbf{y} := T^\dagger(\mathbf{y}_1 + \mathbf{y}_2) = T^\dagger \mathbf{y}_1 = \tilde{T}^{-1} \mathbf{y}_1.$$

Com isto, T^\dagger está bem definido em todo seu domínio e, ainda, é linear em $R(T)$. A seguir, mostraremos algumas propriedades importantes da pseudo-inversa que serão necessárias no decorrer do trabalho.

Proposição 1. Sejam P_N e $P_{\overline{R}}$ projetores ortogonais sobre os subespaços $N(T)$ e $\overline{R(T)}$, respectivamente. Então $R(T^\dagger) = N(T)^\perp$ e as quatro equações de Moore-Penrose (ENGL; HANKE; NEUBAUER, 1996) são:

$$TT^\dagger T = T \tag{1}$$

$$T^\dagger TT^\dagger = T^\dagger \tag{2}$$

$$T^\dagger T = I - P_N \tag{3}$$

$$TT^\dagger = P_{\overline{R}}|_{D(T^\dagger)} \tag{4}$$

Demonstração. Primeiramente, é necessário verificar que $R(T^\dagger) = N(T)^\perp$. Pela definição de T^\dagger , temos que para todo $\mathbf{y} \in D(T^\dagger)$, existem $\mathbf{y}_1 \in R(T)$ e $\mathbf{y}_2 \in R(T)^\perp$ com $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$, assim

$$T^\dagger \mathbf{y} = T^\dagger(\mathbf{y}_1 + \mathbf{y}_2) = T^\dagger \mathbf{y}_1 + T^\dagger \mathbf{y}_2 = T^\dagger \mathbf{y}_1 = T^\dagger P_{\overline{R}} \mathbf{y}.$$

Como $\mathbf{y} \in D(T^\dagger) = R(T) \oplus R(T)^\perp$, isto implica que $P_{\bar{R}}\mathbf{y} \in R(T)$. Portanto, $T^\dagger\mathbf{y} \in R(\tilde{T}^{-1}) = N(T)^\perp$, assim $R(T^\dagger) \subset N(T)^\perp$.

Por outro lado, $T^\dagger T\mathbf{x} = \tilde{T}^{-1}\tilde{T}\mathbf{x} = \mathbf{x}$, $\forall \mathbf{x} \in N(T)^\perp$, ou seja, $\mathbf{x} \in R(T^\dagger)$, assim $N(T)^\perp \subset R(T^\dagger)$. Com isto, concluímos que $R(T^\dagger) = N(T)^\perp$.

4) Para todo $\mathbf{y} \in D(T^\dagger)$ e pelo fato de $R(T^\dagger) = N(T)^\perp$, temos que

$$TT^\dagger\mathbf{y} = TT^\dagger P_{\bar{R}}\mathbf{y} = T\tilde{T}^{-1}P_{\bar{R}}\mathbf{y} = \tilde{T}\tilde{T}^{-1}P_{\bar{R}}\mathbf{y} = P_{\bar{R}}\mathbf{y}.$$

3) Pela definição de T^\dagger , temos que $T^\dagger T\mathbf{x} = \tilde{T}^{-1}T\mathbf{x}$, $\forall \mathbf{x} \in \mathcal{X}$. Com isto,

$$T^\dagger T\mathbf{x} = \tilde{T}^{-1}T\mathbf{x} = \tilde{T}^{-1}T(P_N\mathbf{x} + (I - P_N)\mathbf{x}) = \tilde{T}^{-1}TP_N\mathbf{x} + \tilde{T}^{-1}T(I - P_N)\mathbf{x} = (I - P_N)\mathbf{x}.$$

2) Usando o que foi provado em (4), temos que para todo $\mathbf{y} \in D(T^\dagger)$

$$T^\dagger TT^\dagger\mathbf{y} = T^\dagger P_{\bar{R}}\mathbf{y}|_{D(T)^\perp} = T^\dagger\mathbf{y}.$$

1) Usando o que foi provado em (3), temos que para todo $\mathbf{x} \in \mathcal{X}$

$$TT^\dagger\mathbf{x} = T(I - P_N)\mathbf{x} = T\mathbf{x} - TP_N\mathbf{x} = T\mathbf{x}$$

□

Com as equações de Moore-Penrose definidas é, possível verificar que de fato a pseudo-inversa leva diretamente à solução de norma mínima.

Teorema 1. Seja $\mathbf{y} \in D(T^\dagger)$. Então, $T\mathbf{x} = \mathbf{y}$ tem uma única melhor solução aproximada, que é dada por

$$\mathbf{x}^\dagger := T^\dagger\mathbf{y}.$$

Além disso, o conjunto de todas as soluções de quadrados mínimos é dado por $\mathbf{x}^\dagger + N(T)$.

Demonstração. Seja um conjunto S , onde

$$S = \{\mathbf{z} \in \mathcal{X} \mid T\mathbf{z} = P_{\bar{R}}\mathbf{y}\}.$$

Como $P_{\bar{R}}\mathbf{y} \in R(T)$ para qualquer $\mathbf{y} \in D(T^\dagger)$, temos que $T\mathbf{z} = P_{\bar{R}}\mathbf{y}$ tem pelo menos uma solução, assim $S \neq \emptyset$. Sejam $\mathbf{x} \in \mathcal{X}$ e $\mathbf{z} \in S$ quaisquer, assim

$$\|T\mathbf{z} - \mathbf{y}\|_{\mathcal{Y}} = \|P_{\bar{R}}\mathbf{y} - \mathbf{y}\|_{\mathcal{Y}} \leq \|T\mathbf{x} - \mathbf{y}\|_{\mathcal{Y}}.$$

Por (1), $z \in S$ é uma solução de quadrados mínimos de $Tx = y$. Como z foi escolhido arbitrariamente, todo elemento pertencente a S é uma solução de quadrados mínimos de $Tx = y$. Agora, seja z uma solução qualquer de quadrados mínimos de $Tx = y$,

$$\begin{aligned} \|P_{\overline{R}}y - y\|_{\mathcal{Y}} &\leq \|Tz - y\|_{\mathcal{Y}} = \inf\{\|Tv - y\|_{\mathcal{Y}} \mid v \in \mathcal{X}\} \\ &= \inf\{\|u - y\|_{\mathcal{Y}} \mid u \in R(T)\} \\ &= \|P_{\overline{R}}y - y\|_{\mathcal{Y}}. \end{aligned}$$

Isto verifica que $Tz = P_{\overline{R}}y$. Portanto, S é o conjunto de todas as soluções de quadrados mínimos de $Tx = y$.

O próximo passo é usar o fato que toda solução de quadrados mínimos também é solução de $Tx = P_{\overline{R}}y$. Como $N(T)$ é fechado, toda solução de quadrados mínimos x pode ser escrita da forma $x = x_1 + x_2$, com $x_1 \in N(T)^\perp$ e $x_2 \in N(T)$. Como T é injetiva em $N(T)^\perp$, segue que

$$\begin{aligned} \|x\|_{\mathcal{X}}^2 &= \|x_1 + x_2\|_{\mathcal{X}}^2 = \|x_1\|_{\mathcal{X}}^2 + 2\langle x_1, x_2 \rangle_{\mathcal{X}} + \|x_2\|_{\mathcal{X}}^2 = \|x_1\|_{\mathcal{X}}^2 + \|x_2\|_{\mathcal{X}}^2 \\ &\geq \|x_1\|_{\mathcal{X}}^2 \end{aligned}$$

Sendo assim, a única solução de norma mínima é $x^\dagger := x_1$. Por fim, temos que $x^\dagger \in N(T)^\perp$ e $Tx^\dagger = P_{\overline{R}}y$. Com o auxílio das equações (1) e (3) da pseudo-inversa

$$x^\dagger = P_{N^\perp}x^\dagger = T^\dagger T x^\dagger = T^\dagger P_{\overline{R}}y = T^\dagger T T^\dagger y = T^\dagger y.$$

□

Existe outra maneira de caracterizar as soluções de quadrados mínimos, e que será posteriormente utilizada para desenvolver outros resultados.

Teorema 2. Seja $y \in D(T^\dagger)$. Então, $x \in \mathcal{X}$ é uma solução de quadrados mínimos de $Tx = y$ se, e somente se, a equação normal

$$T^*Tx = T^*y$$

é válida.

Demonstração. Pelo Teorema 1, temos o seguinte:

$$\begin{aligned} x \text{ é solução de quadrados mínimos} &\Leftrightarrow Tx = P_{\overline{R}}y \\ &\Leftrightarrow Tx \in \overline{R(T)}, Tx - y \in \overline{R(T)}^\perp \\ &\Leftrightarrow Tx - y \in \overline{R(T)}^\perp = N(T^*) \\ &\Leftrightarrow T^*(Tx - y) = 0. \end{aligned}$$

□

A solução de norma mínima \mathbf{x}^\dagger é também solução da equação normal do teorema acima, ou seja,

$$\mathbf{x}^\dagger = (T^*T)^\dagger T^* \mathbf{y}. \quad (3.1)$$

Na construção, foi definido que $D(T^\dagger) = R(T) \oplus R(T)^\perp$. Deste modo,

$$\overline{D(T^\dagger)} = \overline{R(T)} \oplus R(T)^\perp = N(T^*)^\perp \oplus N(T^*) = \mathcal{Y}.$$

Em vista disso, $D(T^\dagger)$ é denso em \mathcal{Y} . Se $R(T)$ for fechado, então $D(T^\dagger) = \mathcal{Y}$ e vice-versa. Sendo assim, temos três pontos principais:

1. se $\mathbf{y} \in R(T)^\perp = N(T^\dagger)$, temos que a solução de norma mínima é $\mathbf{x}^\dagger = 0$;
2. se $\mathbf{y} \in \overline{R(T)}$, T^\dagger é contínuo;
3. se existir pelo menos um $\mathbf{y} \in \overline{R(T)} \setminus R(T)$, isto é suficiente para afirmar que T^\dagger não é contínuo.

Teorema 3. A inversa generalizada T^\dagger tem um gráfico fechado $gr(T^\dagger)$. Além disso, T^\dagger é limitado (i.e., contínuo) se, e somente se, $R(T)$ é fechado.

Demonstração. Primeiro, é necessário verificar que $gr(T^\dagger)$ é fechado. Pela definição de \tilde{T} e pelo fato de ser bijetiva vale a seguinte relação

$$\{(\mathbf{y}_1, \tilde{T}^{-1} \mathbf{y}_1) \mid \mathbf{y}_1 \in R(T)\} = \{(T\mathbf{x}, \mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\} \cap \{\mathcal{Y} \times N(T)^\perp\}. \quad (3.2)$$

Considere o gráfico de T^\dagger definido por:

$$\begin{aligned} gr(T^\dagger) &= \{(\mathbf{y}, T^\dagger \mathbf{y}) \mid \mathbf{y} \in D(T^\dagger)\} \\ &= \{(\mathbf{y}_1 + \mathbf{y}_2, T^\dagger(\mathbf{y}_1 + \mathbf{y}_2)) \mid \mathbf{y}_1 \in R(T), \mathbf{y}_2 \in R(T)^\perp\} \\ &= \{(\mathbf{y}_1 + \mathbf{y}_2, T^\dagger \mathbf{y}_1) \mid \mathbf{y}_1 \in R(T), \mathbf{y}_2 \in R(T)^\perp\} \\ &= \{(\mathbf{y}_1, \tilde{T}^{-1} \mathbf{y}_1) \mid \mathbf{y}_1 \in R(T)\} + (R(T)^\perp \times \{0\}) \end{aligned}$$

Utilizando o fato explicitado em (3.2), temos que

$$gr(T^\dagger) = [\{(T\mathbf{x}, \mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\} \cap \{\mathcal{Y} \times N(T)^\perp\}] + (R(T)^\perp \times \{0\}).$$

Todos os espaços do lado direito da igualdade acima são fechados, sendo assim $gr(T^\dagger)$ é fechado.

Assumindo que $R(T)$ é fechado, então $D(T^\dagger) = \mathcal{Y}$. Portanto, pelo Teorema 10, T^\dagger é contínuo. Reciprocamente, seja T^\dagger contínuo em $D(T^\dagger)$ e $D(T^\dagger)$ denso em \mathcal{Y} ($\mathcal{Y} \subset \overline{D(T^\dagger)}$). Pode-se estender continuamente T^\dagger em \mathcal{Y} , onde

$$\overline{T^\dagger} \mathbf{y} := \lim_{n \rightarrow \infty} T^\dagger \mathbf{y}_n \text{ para uma sequência } \{\mathbf{y}_n\} \subset D(T^\dagger), \text{ com } \mathbf{y}_n \rightarrow \mathbf{y} \in \mathcal{Y}.$$

Agora seja um $\mathbf{y} \in \overline{R(T)}$, então existe uma sequência $\{\mathbf{y}_n\} \subset R(T)$ com $\mathbf{y}_n \rightarrow \mathbf{y}$. Com isto,

$$\mathbf{y} = P_{\overline{R(T)}}\mathbf{y} = \lim_{n \rightarrow \infty} P_{\overline{R(T)}}\mathbf{y}_n = \lim_{n \rightarrow \infty} TT^\dagger\mathbf{y}_n = T\overline{T^\dagger}\mathbf{y} \in R(T).$$

Logo, $\overline{R(T)} \subset R(T)$. Portanto, $R(T) = \overline{R(T)}$ e por consequência $R(T)$ é fechado. \square

Com a definição de pseudo-inversa bem estabelecida, o intuito agora é apresentar uma maneira de construir a pseudo-inversa para operadores compactos $K \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ através de sistemas ortonormais. Para a realização disto é utilizado a decomposição espectral, porém é necessário que o operador K seja auto-adjunto. Existem casos que o operador pode não ser auto-adjunto, devido à Eq. (3.1) é possível considerar o operador K^*K , que neste caso é auto-adjunto.

Teorema 4. Para $K \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$, existem

1. uma sequência $\{\sigma_n\}_{n \in \mathbb{N}}$ com $\sigma_1 \geq \sigma_2 \geq \dots > 0$;
2. um sistema ortonormal completo $\{\mathbf{u}_n\}_{n \in \mathbb{N}} \subset \mathcal{Y}$ de $\overline{R(K)}$;
3. um sistema ortonormal completo $\{\mathbf{v}_n\}_{n \in \mathbb{N}} \subset \mathcal{X}$ de $\overline{R(K^*)}$.

Além disso,

$$\begin{aligned} K\mathbf{v}_n &= \sigma_n\mathbf{u}_n \text{ e } K^*\mathbf{u}_n = \sigma_n\mathbf{v}_n, \quad \forall n \in \mathbb{N}; \\ K\mathbf{x} &= \sum_{n \in \mathbb{N}} \sigma_n \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} \mathbf{u}_n, \quad \forall \mathbf{x} \in \mathcal{X}. \end{aligned}$$

Demonstração. Seja $K^*K : \mathcal{X} \rightarrow \mathcal{X}$ compacto e auto-adjunto, então, pelo Teorema 11, existem uma sequência $\{\lambda_n\} \subset [0, \infty)$ e um sistema ortonormal completo $\{\mathbf{v}_n\} \subset \mathcal{X}$ para as quais

$$K^*K\mathbf{x} = \sum \lambda_n \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} \mathbf{v}_n, \quad \forall \mathbf{x} \in \mathcal{X}.$$

Ordena-se toda sequência $\{\lambda_n\}$ decrescentemente, sendo assim

$$\lambda_n = \lambda_n \|\mathbf{v}_n\|_{\mathcal{X}}^2 = \langle \lambda_n \mathbf{v}_n, \mathbf{v}_n \rangle_{\mathcal{X}} = \langle K^*K\mathbf{v}_n, \mathbf{v}_n \rangle_{\mathcal{X}} = \langle K\mathbf{v}_n, K\mathbf{v}_n \rangle_{\mathcal{Y}} = \|K\mathbf{v}_n\|_{\mathcal{Y}}^2 > 0.$$

Com isto, definimos

$$\sigma_n := \sqrt{\lambda_n} \text{ e } \mathbf{u}_n := \frac{K\mathbf{v}_n}{\sigma_n}.$$

Agora é necessário mostrar que a sequência $\{\mathbf{u}_n\}$ forma um sistema ortonormal completo.

Sejam quaisquer \mathbf{u}_i e \mathbf{u}_j ,

$$\begin{aligned} \langle \mathbf{u}_i, \mathbf{u}_j \rangle_{\mathcal{Y}} &= \left\langle \frac{K\mathbf{v}_i}{\sigma_i}, \frac{K\mathbf{v}_j}{\sigma_j} \right\rangle_{\mathcal{Y}} = \frac{1}{\sigma_i \sigma_j} \langle K\mathbf{v}_i, K\mathbf{v}_j \rangle_{\mathcal{Y}} = \frac{1}{\sigma_i \sigma_j} \langle K^*K\mathbf{v}_i, \mathbf{v}_j \rangle_{\mathcal{X}} \\ &= \frac{\lambda_i}{\sigma_i \sigma_j} \langle \mathbf{v}_i, \mathbf{v}_j \rangle_{\mathcal{X}} \\ &= \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases} \end{aligned}$$

Portanto, temos que para todo $n \in \mathbb{N}$

$$K^* \mathbf{u}_n = K^* \frac{K \mathbf{v}_n}{\sigma_n} = \frac{1}{\sigma_n} K^* K \mathbf{v}_n = \frac{1}{\sigma_n} \lambda_n \mathbf{v}_n = \sigma_n \mathbf{v}_n.$$

Sabemos que $\{\mathbf{v}_n\}$ é um sistema ortonormal completo de $\overline{R(K^*K)}$. Portanto, basta provar que $\overline{R(K^*K)} = \overline{R(K^*)}$. Para isto, seja $\mathbf{u} \in \overline{R(K^*)}$, assim $\exists \{\mathbf{y}_n\} \subset Y$ tal que $K^* \mathbf{y}_n \rightarrow \mathbf{u}$. Tome em particular uma sequência $\{\mathbf{y}_n\} \subset N(K^*)^\perp = \overline{R(K)}$, deste modo para cada elemento \mathbf{y}_n existe $\{\mathbf{x}_n\} \subset X$ tal que $K \mathbf{x}_n \rightarrow \mathbf{y}_n$. Juntando as implicações, temos que $K^* K \mathbf{x}_n \rightarrow \mathbf{u}$ e, portanto, $\mathbf{u} \in \overline{R(K^*K)}$. Reciprocamente, seja $\mathbf{u} \in \overline{R(K^*K)}$, assim $\exists \{\mathbf{x}_n\} \subset X$ tal que $K^* K \mathbf{x}_n \rightarrow \mathbf{u}$. Com isto, formamos uma sequência $\{\mathbf{y}_n\}$ com $\mathbf{y}_n = K \mathbf{x}_n$. Portanto, $K^* \mathbf{y}_n \rightarrow \mathbf{u}$, assim $\mathbf{u} \in \overline{R(K^*)}$.

Mostramos que $\{\mathbf{v}_n\}$ é um sistema ortonormal completo de $\overline{R(K^*)} = N(K)^\perp$, assim para todo $\mathbf{x} \in \mathcal{X}$ temos que

$$K \mathbf{x} = KP_{N^\perp} \mathbf{x} = K \left(\sum_{n \in \mathbb{N}} \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} \mathbf{v}_n \right) = \sum_{n \in \mathbb{N}} \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} K \mathbf{v}_n = \sum_{n \in \mathbb{N}} \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} \sigma_n \mathbf{u}_n.$$

□

Com o teorema acima, a tripla $(\sigma_n, \mathbf{u}_n, \mathbf{v}_n)$ é denominada sistema singular, sendo apenas uma generalização da decomposição de valores singulares para operadores definido entre espaços de dimensões infinitas. Uma consequência é mostrada abaixo, que expressa a solução \mathbf{x}^\dagger conforme o sistema singular.

Teorema 5. Seja $K \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ com um sistema singular $(\sigma_n, \mathbf{u}_n, \mathbf{v}_n)$ e $\mathbf{y} \in \overline{R(K)}$. Então $\mathbf{y} \in R(K)$ se, e somente se, a condição de Picard

$$\sum_{n \in \mathbb{N}} \frac{|\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}|^2}{\sigma_n^2} < \infty$$

é satisfeita. Neste caso,

$$K^\dagger \mathbf{y} = \sum_{n \in \mathbb{N}} \frac{\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}} \mathbf{v}_n}{\sigma_n}$$

Demonstração. Seja $\mathbf{y} \in R(K)$, assim $\exists \mathbf{x} \in \mathcal{X}$ tal que $K \mathbf{x} = \mathbf{y}$. Com isso,

$$\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}} = \langle K \mathbf{x}, \mathbf{u}_n \rangle_{\mathcal{Y}} = \langle \mathbf{x}, K^* \mathbf{u}_n \rangle_{\mathcal{X}} = \langle \mathbf{x}, \sigma_n \mathbf{v}_n \rangle_{\mathcal{X}} = \sigma_n \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}}$$

Aplicando o módulo, elevando ao quadrado e somando todos os elementos em ambos os lados teremos

$$\sum_{n \in \mathbb{N}} \frac{|\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}|^2}{\sigma_n^2} = \sum_{n \in \mathbb{N}} |\langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}}|^2.$$

Pela desigualdade de Bessel (KREYSZIG, 1991, Teorema 3.4-6),

$$\sum_{n \in \mathbb{N}} \frac{|\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}|^2}{\sigma_n^2} = \sum_{n \in \mathbb{N}} |\langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}}|^2 \leq \|\mathbf{x}\|_{\mathcal{X}}^2 < \infty.$$

Reciprocamente, seja $\mathbf{y} \in \overline{R(K)}$ que satisfaz a condição de Picard. Construímos uma sequência da forma

$$\left\{ \sum_{n=1}^N \frac{|\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}|^2}{\sigma_n^2} \right\}_{N \in \mathbb{N}},$$

que conforme a hipótese, converge. Agora, definimos uma nova sequência $\{\mathbf{x}_n\}$, onde

$$\mathbf{x}_n = \sum_{k=1}^n \frac{\langle \mathbf{y}, \mathbf{u}_k \rangle_{\mathcal{Y}} \mathbf{v}_k}{\sigma_k}.$$

Esta sequência definida é de Cauchy, pois para $m, n \in \mathbb{N}$ temos

$$\|\mathbf{x}_n - \mathbf{x}_m\|_{\mathcal{X}}^2 = \left\| \sum_{k=m+1}^n \frac{\langle \mathbf{y}, \mathbf{u}_k \rangle_{\mathcal{Y}} \mathbf{v}_k}{\sigma_k} \right\|_{\mathcal{X}}^2 = \sum_{k=m+1}^n \left| \frac{\langle \mathbf{y}, \mathbf{u}_k \rangle_{\mathcal{Y}}}{\sigma_k} \right|^2 \rightarrow 0,$$

quando $n, m \rightarrow \infty$. Denotemos $\mathbf{x} := \lim_{n \rightarrow \infty} \mathbf{x}_n$. É sabido que $\{\mathbf{v}_n\} \subset \overline{R(K^*)}$, e por consequência $\{\mathbf{x}_n\} \subset \overline{R(K^*)}$. Como $\overline{R(K^*)}$ é fechado, segue para qualquer $\mathbf{x} \in \overline{R(K^*)} = N(K)^\perp$. Sendo assim, temos que

$$K\mathbf{x} = K \sum_{n \in \mathbb{N}} \frac{\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}} \mathbf{v}_n}{\sigma_n} = \sum_{n \in \mathbb{N}} \frac{\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}}{\sigma_n} K\mathbf{v}_n = \sum_{n \in \mathbb{N}} \langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}} \mathbf{u}_n = \mathbf{y}.$$

Isto mostra que $\mathbf{y} \in R(K)$. Se $K\mathbf{x} = \mathbf{y}$ para qualquer $\mathbf{x} \in N(K)^\perp$, pelo Teorema 1

$$\mathbf{x} = K^\dagger \mathbf{y} = \sum_{n \in \mathbb{N}} \frac{\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}} \mathbf{v}_n}{\sigma_n}$$

□

A condição de Picard estabelece que uma solução de norma mínima só existe se $\langle \mathbf{y}, \mathbf{u}_n \rangle_{\mathcal{Y}}$ decai mais rápido que os valores singulares σ_n . No entanto, suponhamos que algumas medidas de \mathbf{y} não sejam totalmente exatas, por exemplo, $\mathbf{y}^\delta = \mathbf{y} + \delta \mathbf{u}_n$ para algum $\delta > 0$, assim

$$\|K^\dagger \mathbf{y} - K^\dagger \mathbf{y}^\delta\|_{\mathcal{X}} = \delta \|K^\dagger \mathbf{u}_n\|_{\mathcal{X}} = \frac{\delta}{\sigma_n},$$

como $\sigma_n \rightarrow 0$, o erro é amplificado. Por outro lado, se $R(K)$ é dimensão finita, a sequência $\{\frac{1}{\sigma_n}\}$ é limitada e, assim o erro fica limitado. Além disso, o operador K^\dagger é contínuo. A decomposição do operador K^\dagger por sistema singular é uma ferramenta poderosa, porém o cálculo da decomposição é complexo e demanda muito tempo computacional (CLINE; DHILLON, 2006).

Teorema 6. Se $K \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ tem $R(K)$ de dimensão infinita, então K^\dagger não é contínuo.

Demonstração. A maneira mais trivial é demonstrar pela contrapositividade.

Seja K^\dagger contínuo, então pelo Teorema 3 $R(K)$ é fechado. Utilizando a mesma estratégia na definição da pseudo-inversa, definimos o operador $\tilde{K} := K|_{N(K)^\perp} : N(K)^\perp \rightarrow R(K)$. Consequentemente, \tilde{K}^{-1} é contínuo e com isso $K \circ \tilde{K}^{-1}$ é compacto. Sendo assim,

$$K\tilde{K}^{-1}\mathbf{y} = \mathbf{y}, \quad \forall \mathbf{y} \in R(K).$$

Este é o operador identidade $I : R(K) \rightarrow R(K)$, onde I é compacto. Por (KREYSZIG, 1991, Lema 8.1-2), $R(K)$ tem dimensão finita. \square

Observamos que equação $T\mathbf{x} = \mathbf{y}$ admite, para qualquer $\mathbf{y} \in D(T^\dagger)$, uma única solução de norma mínima $\mathbf{x}^\dagger = T^\dagger \mathbf{y}$. Na prática, não é possível obter \mathbf{y} exato, sendo disponível apenas uma aproximação \mathbf{y}^δ , para a qual

$$\|\mathbf{y} - \mathbf{y}^\delta\|_{\mathcal{Y}} \leq \delta,$$

onde $\delta > 0$ é definido como o nível de ruído.

Pelo Teorema 6, o operador T^\dagger pode não ser contínuo, por isso não é garantido que $T^\dagger \mathbf{y}^\delta$ produzirá uma boa aproximação para \mathbf{x}^\dagger (ver (SEIDMAN, 1980)). Por conta deste empecilho, o objetivo é construir uma aproximação \mathbf{x}_α^δ que satisfaça duas condições:

1. dependa continuamente de \mathbf{y}^δ ;
2. através da escolha de um $\alpha > 0$, se aproxime de \mathbf{x}^\dagger o máximo que o nível de ruído permita.

Um método que constrói esta aproximação é conhecido como regularização, cuja ideia é construir uma família de operadores contínuos que se aproximem pontualmente do operador T^\dagger .

3.2 Operador de regularização

No geral, a regularização é uma aproximação de um problema mal-posto por uma família de problemas bem-postos. A ideia padrão da regularização é construir uma família de operadores lineares $R_\alpha : \mathcal{Y} \rightarrow \mathcal{X}$ que aproximem pontualmente o operador T^\dagger . A teoria de regularização para problemas inversos é bem discutida e desenvolvida em Engl, Hanke e Neubauer (1996), Groetsch (1993). Uma definição mais precisa é a seguinte

Definição 3. Sejam $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ e $\alpha_0 \in (0, +\infty)$. Para todo $\alpha \in (0, \alpha_0)$, seja

$$R_\alpha : \mathcal{Y} \rightarrow \mathcal{X}$$

um operador contínuo. A família $\{R_\alpha\}$ é chamada de regularização ou operador de regularização para T^\dagger se, $\forall \mathbf{y} \in D(T^\dagger)$, existir uma regra de escolha do parâmetro $\alpha := \alpha(\delta, \mathbf{y}^\delta)$ tal que a condição

$$\limsup_{\delta \rightarrow 0} \{\|R_{\alpha(\delta, \mathbf{y}^\delta)} \mathbf{y}^\delta - T^\dagger \mathbf{y}\|_{\mathcal{X}} \mid \mathbf{y}^\delta \in \mathcal{Y}, \|\mathbf{y}^\delta - \mathbf{y}\|_{\mathcal{Y}} \leq \delta\} = 0$$

é satisfeita. Além disso,

$$\alpha : \mathbb{R}^+ \times \mathcal{Y} \rightarrow (0, \alpha_0)$$

é tal que

$$\limsup_{\delta \rightarrow 0} \{\alpha(\delta, \mathbf{y}^\delta) \mid \mathbf{y}^\delta \in \mathcal{Y}, \|\mathbf{y}^\delta - \mathbf{y}\|_{\mathcal{Y}} \leq \delta\} = 0.$$

O par (R_α, α) é considerado um método de regularização convergente se a definição acima for válida, ou seja, $x_\alpha := R_\alpha \mathbf{y}$ se aproxima o suficiente de x^\dagger . Uma etapa importante é definir uma regra para escolha do parâmetro de regularização α que possa garantir uma boa aproximação de $T^\dagger \mathbf{y}^\delta$. Distinguímos a regra de escolha em três grupos:

1. Regra de escolha *a priori* se depende somente de δ ($\alpha := \alpha(\delta)$);
2. Regra de escolha *a posteriori* se depende de δ e \mathbf{y}^δ ($\alpha := \alpha(\delta, \mathbf{y}^\delta)$);
3. Regra de escolha heurística se depende somente de \mathbf{y}^δ ($\alpha := \alpha(\mathbf{y}^\delta)$).

Dentre os três tipos de regras definidas, somente na terceira regra não estão bem estabelecida, sendo o principal empecilho o resultado mostrado por Bakushinskii (BAKUSHINSKII, 1984), onde é verificado que qualquer regra de escolha do parâmetro da forma $\alpha := \alpha(\mathbf{y}^\delta)$ não converge no pior caso. Entretanto, isso não impede que o uso de regras de escolha heurística possa fornecer uma boa aproximação para x_α^δ .

A expectativa agora é escolher um parâmetro que aproxime a solução regularizada da solução exata. Para isto, consideraremos o erro geral $\|R_\alpha \mathbf{y}^\delta - T^\dagger \mathbf{y}\|_{\mathcal{Y}}$ e efetuaremos algumas manipulações,

$$\begin{aligned} \|R_\alpha \mathbf{y}^\delta - T^\dagger \mathbf{y}\|_{\mathcal{Y}} &= \|R_\alpha \mathbf{y}^\delta + R_\alpha \mathbf{y} - R_\alpha \mathbf{y} - T^\dagger \mathbf{y}\|_{\mathcal{Y}} \\ &\leq \|R_\alpha \mathbf{y}^\delta - R_\alpha \mathbf{y}\|_{\mathcal{Y}} + \|R_\alpha \mathbf{y} - T^\dagger \mathbf{y}\|_{\mathcal{Y}} \\ &= \|R_\alpha (\mathbf{y}^\delta - \mathbf{y})\|_{\mathcal{Y}} + \|R_\alpha \mathbf{y} - T^\dagger \mathbf{y}\|_{\mathcal{Y}} \\ &\leq \delta \|R_\alpha\|_{\mathcal{Y}} + \|R_\alpha \mathbf{y} - T^\dagger \mathbf{y}\|_{\mathcal{Y}}. \end{aligned}$$

Com as operações feitas acima, o erro geral se divide em duas partes. A primeira parte, demonstra que a norma do operador R_α , que podemos entender como o “tamanho” do operador, é amplificada pelo erro dos dados δ . Além do que, podem ocorrer casos em que o termo $\|R_\alpha\|_{\mathcal{Y}} \rightarrow \infty$ quando $\alpha \rightarrow 0$, conforme mostrado por Engl, Hanke e Neubauer (1996). A segunda parte, representa o erro de aproximação que desaparece devido à convergência pontual do operador de regularização. Portanto, o desafio é construir uma regra de escolha para obter uma aproximação razoável para o operador T^\dagger e que o erro geral se anule quando $\delta \rightarrow 0$.

Geralmente, a convergência de um método de regularização é bastante lenta. No entanto, as taxas de convergência costumam ser dadas em subconjuntos de \mathcal{X} , ou seja, é preciso ter conhecimento a priori da solução exata. Estas informações são feitas em condições específicas em que a solução exata garante a convergência do método, estes termos são conhecidas como condições de fonte (ENGL; HANKE; NEUBAUER, 1996).

3.3 Regularização de Tikhonov

Dentre os numerosos operadores de regularização existentes, o mais tradicional e difundido é o operador desenvolvido pelo matemático russo Andrey Nikolayevich Tikhonov em [Tikhonov \(1963\)](#). Andrey Tikhonov tinha como objetivo resolver o problema $Tx = y$, onde T era um operador integral e, x e y eram funções. Ele sugeriu substituir o problema de quadrados mínimos

$$x^* = \arg \inf \{ \|Tz - x\|_2^2 \mid z \in \mathcal{H} \}$$

por

$$x_\alpha^* = \arg \inf \{ \|Tz - y\|_2^2 + \alpha^2 \Omega(z) \mid z \in \mathcal{H} \},$$

onde $\Omega(z) = \int_a^b p(s)z(s)^2 + q(s)z'(s)^2 ds$, p e q eram funções de peso positivo, com \mathcal{H} um espaço de funções apropriado. Posteriormente, [Tikhonov e Arsenin \(1977\)](#) mostraram que um operador de regularização pode ser composto por uma função de quadrados mínimos com a adição de informações extras.

A família de operadores para a regularização de Tikhonov é definida por

$$R_\alpha := (\alpha I + T^*T)^{-1}T^*.$$

A escolha destes operadores é válida, pois o operador T^*T é auto-adjunto, compacto e, tem autovalores não negativos. Como o parâmetro de regularização α é positivo por definição, $(T^*T + \alpha I)$ tem autovalores positivos e deste modo possui inversa limitada. Portanto, a equação

$$(T^*T + \alpha I)\hat{x}_\alpha = T^*y$$

é bem-posta, além disso, ela pode ser vista como uma equação normal regularizada.

Outra formulação apresentada na literatura para a regularização de Tikhonov é dada pela interpretação variacional

$$\hat{x}_\alpha = \arg \min_{x \in \mathcal{X}} \{ \|Tx - y\|_{\mathcal{Y}}^2 + \alpha \|x\|_{\mathcal{X}}^2 \}.$$

Teorema 7. Sejam $y \in \mathcal{Y}$ e $\alpha > 0$. Então $\hat{x}_\alpha := R_\alpha y$ é o único minimizador do funcional de Tikhonov $J_\alpha(x) := \|Tx - y\|_{\mathcal{Y}}^2 + \alpha \|x\|_{\mathcal{X}}^2$.

Demonstração. Para que $x^* \in \mathcal{X}$ seja um minimizador de J_α é necessário satisfazer $J_\alpha(x^*) \leq$

$J_\alpha(\mathbf{x}), \forall \mathbf{x} \in \mathcal{X}$. Seja a diferença de J_α para um \mathbf{x} arbitrário e $\hat{\mathbf{x}}_\alpha$,

$$\begin{aligned}
J_\alpha(\mathbf{x}) - J_\alpha(\hat{\mathbf{x}}_\alpha) &= \|T\mathbf{x} - \mathbf{y}\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 - \|T\hat{\mathbf{x}}_\alpha - \mathbf{y}\|_{\mathcal{Y}}^2 - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x}\|_{\mathcal{Y}}^2 + \|\mathbf{y}\|_{\mathcal{Y}}^2 - 2\langle T\mathbf{x}, \mathbf{y} \rangle_{\mathcal{Y}} + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 - \|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - \|\mathbf{y}\|_{\mathcal{Y}}^2 \\
&\quad + 2\langle T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x}\|_{\mathcal{Y}}^2 - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 - \|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x}\|_{\mathcal{Y}}^2 + \|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - \|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + 2\langle T\mathbf{x}, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&\quad - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 - \|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - 2\langle T\mathbf{x}, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - 2\|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + 2\langle T\mathbf{x}, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} \\
&\quad - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 - 2\|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + 2\langle T\mathbf{x}, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} + \alpha\|\mathbf{x}\|_{\mathcal{X}}^2 \\
&\quad - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 + \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 - \alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 - 2\alpha\langle \mathbf{x}, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} \\
&\quad + 2\alpha\langle \mathbf{x}, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 - 2\|T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + 2\langle T\mathbf{x}, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} \\
&\quad - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} - 2\alpha\|\hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 + 2\alpha\langle \mathbf{x}, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 + 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, T\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{Y}} \\
&\quad - 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, \mathbf{y} \rangle_{\mathcal{Y}} + 2\alpha\langle \mathbf{x} - \hat{\mathbf{x}}_\alpha, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 + 2\langle T\mathbf{x} - T\hat{\mathbf{x}}_\alpha, T\hat{\mathbf{x}}_\alpha - \mathbf{y} \rangle_{\mathcal{Y}} \\
&\quad + 2\alpha\langle \mathbf{x} - \hat{\mathbf{x}}_\alpha, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 + 2\langle \mathbf{x} - \hat{\mathbf{x}}_\alpha, T^*(T\hat{\mathbf{x}}_\alpha - \mathbf{y}) \rangle_{\mathcal{X}} \\
&\quad + 2\alpha\langle \mathbf{x} - \hat{\mathbf{x}}_\alpha, \hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + 2\langle \mathbf{x} - \hat{\mathbf{x}}_\alpha, T^*(T\hat{\mathbf{x}}_\alpha - \mathbf{y}) + \alpha\hat{\mathbf{x}}_\alpha \rangle_{\mathcal{X}} + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x} - T\hat{\mathbf{x}}_\alpha\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x} - \hat{\mathbf{x}}_\alpha\|_{\mathcal{X}}^2 \\
&\geq 0.
\end{aligned}$$

Para mostrar a unicidade, tomamos $\hat{\mathbf{x}} = \mathbf{x}^* + t\mathbf{z}$, com $t > 0$ e $\mathbf{z} \in \mathcal{X}$. Então,

$$\begin{aligned}
J_\alpha(\mathbf{x}^* + t\mathbf{z}) - J_\alpha(\mathbf{x}^*) &= \|T(\mathbf{x}^* + t\mathbf{z}) - \mathbf{y}\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x}^* + t\mathbf{z}\|_{\mathcal{X}}^2 - \|T\mathbf{x}^* - \mathbf{y}\|_{\mathcal{Y}}^2 \\
&\quad - \alpha\|\mathbf{x}^*\|_{\mathcal{X}}^2 \\
&= \|(T\mathbf{x}^* - \mathbf{y}) + tT\mathbf{z}\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x}^* + t\mathbf{z}\|_{\mathcal{X}}^2 - \|T\mathbf{x}^* - \mathbf{y}\|_{\mathcal{Y}}^2 \\
&\quad - \alpha\|\mathbf{x}^*\|_{\mathcal{X}}^2 \\
&= \|T\mathbf{x}^* - \mathbf{y}\|_{\mathcal{Y}}^2 + 2\langle T\mathbf{x}^* - \mathbf{y}, tT\mathbf{z} \rangle_{\mathcal{Y}} + \|tT\mathbf{z}\|_{\mathcal{Y}}^2 + \alpha\|\mathbf{x}^*\|_{\mathcal{X}}^2 \\
&\quad + 2\alpha\langle \mathbf{x}^*, t\mathbf{z} \rangle_{\mathcal{X}} + \alpha\|t\mathbf{z}\|_{\mathcal{X}}^2 - \|T\mathbf{x}^* - \mathbf{y}\|_{\mathcal{Y}}^2 - \alpha\|\mathbf{x}^*\|_{\mathcal{X}}^2 \\
&= 2\langle T\mathbf{x}^* - \mathbf{y}, tT\mathbf{z} \rangle_{\mathcal{Y}} + \|tT\mathbf{z}\|_{\mathcal{Y}}^2 + 2\alpha\langle \mathbf{x}^*, t\mathbf{z} \rangle_{\mathcal{X}} + \alpha\|t\mathbf{z}\|_{\mathcal{X}}^2 \\
&= 2t\langle T\mathbf{x}^* - \mathbf{y}, T\mathbf{z} \rangle_{\mathcal{Y}} + t^2\|T\mathbf{z}\|_{\mathcal{Y}}^2 + 2t\alpha\langle \mathbf{x}^*, \mathbf{z} \rangle_{\mathcal{X}} + \alpha t^2\|\mathbf{z}\|_{\mathcal{X}}^2 \\
&= 2t\langle T^*(T\mathbf{x}^* - \mathbf{y}), \mathbf{z} \rangle_{\mathcal{X}} + t^2\|T\mathbf{z}\|_{\mathcal{Y}}^2 + 2t\alpha\langle \mathbf{x}^*, \mathbf{z} \rangle_{\mathcal{X}}
\end{aligned}$$

$$\begin{aligned}
& + \alpha t^2 \|z\|_{\mathcal{X}}^2 \\
& = 2t \langle T^*(T\mathbf{x}^* - \mathbf{y}) + \alpha \mathbf{x}^*, z \rangle_{\mathcal{X}} + t^2 \|Tz\|_{\mathcal{Y}}^2 + \alpha t^2 \|z\|_{\mathcal{X}}^2.
\end{aligned}$$

Dividindo por $t > 0$ e passando o limite quanto $t \rightarrow 0$, temos que

$$2 \langle T^*(T\mathbf{x}^* - \mathbf{y}) + \alpha \mathbf{x}^*, z \rangle_{\mathcal{X}} \geq 0.$$

Como z foi tomado arbitrariamente, a única possibilidade é que $T^*(T\mathbf{x}^* - \mathbf{y}) + \alpha \mathbf{x}^* = 0$, com isso $\mathbf{x}_\alpha^\delta = \mathbf{x}^*$. Portanto, \mathbf{x}_α^δ é o único minimizador do funcional de Tikhonov. \square

A minimização no funcional de Tikhonov é um compromisso entre minimizar a norma $\|T\mathbf{x} - \mathbf{y}\|_{\mathcal{Y}}$ e conseguir tornar a $\|\mathbf{x}\|$ pequena e forçar a estabilidade. O parâmetro α tem grande influência neste compromisso, pois ele irá estabelecer o equilíbrio entre a estabilidade e a precisão da solução. Com a definição da solução regularizada pela regularização de Tikhonov existem resultados sobre a convergência e a estabilidade da abordagem selecionada, os quais podem ser vistos em [Engl, Hanke e Neubauer \(1996\)](#).

3.4 Métodos para escolha do parâmetro de regularização

Muitos métodos foram criados e, ainda são, para encontrar um valor apropriado para o parâmetro de regularização, onde cada um deles foram desenvolvidos e testados em diferentes contextos. Atualmente, uma vasta quantidade de métodos e uma parte deles pode ser encontrada em [Bauer e Lukas \(2011\)](#). Nesta subseção serão apresentados três métodos, alguns dos mais utilizados.

3.4.1 Princípio da discrepância

O método do princípio da discrepância foi proposto por Philips ([PHILLIPS, 1962](#)), porém foi desenvolvido e analisado por [Morozov \(1966\)](#). É um dos mais antigos e amplamente utilizados. A justificativa para utilização do método é simplesmente que, para uma boa solução regularizada, a norma do resíduo deve corresponder ao nível de ruído. A ideia é basicamente a seguinte: sejam $\mathbf{y} \in D(T^\dagger)$ e $\|\mathbf{y}^\delta - \mathbf{y}\|_{\mathcal{Y}} \leq \delta$ e considere o resíduo

$$\|TR_\alpha \mathbf{y}^\delta - \mathbf{y}^\delta\|_{\mathcal{Y}}.$$

Se $\mathbf{y} \in R(T)$ e $\|\mathbf{y} - \mathbf{y}^\delta\|_{\mathcal{Y}} = \delta$, então

$$\|T\mathbf{x}^\dagger - \mathbf{y}^\delta\|_{\mathcal{Y}} = \|\mathbf{y} - \mathbf{y}^\delta\|_{\mathcal{Y}} = \delta.$$

A grande motivação para o princípio da discrepância introduzido por Morozov é que o resíduo da regularização não pode ser menor que o nível de ruído. Sendo assim, para um $\delta > 0$ e $\|\mathbf{y}^\delta - \mathbf{y}\| \leq \delta$, escolhamos um $\alpha := \alpha(\delta, \mathbf{y}^\delta)$ tal que

$$\|TR_\alpha \mathbf{y}^\delta - \mathbf{y}^\delta\|_{\mathcal{Y}} \leq \tau \delta, \quad \text{para algum } \tau > 1.$$

Na prática, consiste em computar uma sequência nula¹ $\{\alpha_n\}$, calcular sucessivamente $R_{\alpha_n} \mathbf{y}^\delta$ e interromper o processo quando o princípio da discrepância for satisfeito. O método proposto por Morozov foi criado em um ambiente determinístico, porém há alguns trabalhos nos contextos discreto e estocástico (VOGEL, 2002; LUKAS, 1995). Para o início do método é necessária uma estimativa precisa do ruído, se existirem pequenas divergências sobre esta estimativa, o método pode levar a soluções ruins (HANSEN, 1998).

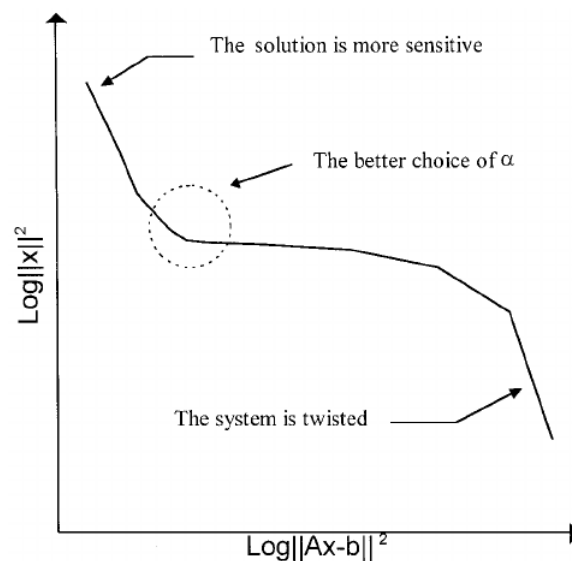
3.4.2 Validação cruzada generalizada

O método da validação cruzada generalizada (*generalized cross-validation* - GCV) foi proposto e desenvolvido por Golub, Heath e Wahba (1979), que consiste numa modificação da validação cruzada comum (ALLEN, 1974). É um método bastante popular para problemas práticos onde os dados são discretos e contêm ruído estocástico. O GCV parte da ideia que um bom método de escolha do parâmetro deve conseguir detectar quais dados foram perdidos. Sendo assim, o valor do parâmetro de regularização é produzido encontrando o mínimo da função

$$G(\alpha) = \frac{m \|AR_\alpha \mathbf{y} - \mathbf{y}\|_2}{(\text{tr}(I - AZ(\alpha)))^2},$$

onde A é a representação do operador T num espaço de dimensão finita, $\text{tr}(\cdot)$ é o traço da matriz, m é o número de linhas de A e $Z(\alpha) = (A^t A + \alpha I)^{-1} A^t$. Contudo, o custo computacional de calcular o mínimo da função acima é grande. Uma maneira de reduzir o custo é calcular a decomposição de valores singulares da matriz A , se estiver disponível.

Figura 2 – Exemplo de uma plotagem da curva-L



Fonte: Retirado de Chang, Yeih e Shieh (2001)

¹ Sequência nula é uma sequência que converge para zero.

3.4.3 Curva-L

O método da curva-L foi proposto por Per Christian Hansen (HANSEN, 1992). Este método é baseado no fato que um gráfico log-log de $(\|T\mathbf{x}_\alpha^\delta - \mathbf{y}^\delta\|_{\mathcal{Y}}, \|\mathbf{x}_\alpha^\delta\|_{\mathcal{X}})$ muitas vezes tem forma de L. Um exemplo pode ser visto na Figura 2. O uso da curva-L mostra o equilíbrio entre a norma do resíduo da solução e a norma da solução. A finalidade do uso de eixos logarítmicos é deixar a curva próxima de um formato de L, conforme mostrado por Hansen e O’Leary (1993). Além disso, a utilização oferece uma ajuda na visualização do parâmetro de regularização ótimo que está o mais próximo possível da curvatura máxima. Os pontos da parte vertical correspondem a α grande, onde temos as soluções sub-suavizadas. Os pontos da parte horizontal correspondem a α pequeno, onde as soluções são super-suavizadas. Por isso, é sugerido que o “ponto da quina” da curva-L é um valor apropriado para α . Devido à sua simplicidade, o método tornou-se bastante popular e, por isso, diversas variantes do método foram criadas (REGIŃSKA, 1996).

APRENDIZADO DE MÁQUINA

O estudo das redes neurais começou em 1943, no momento em que o matemático Walter Pitts e o neurocientista Warren Sturgis McCulloch apresentaram um modelo matemático que mostrava como um conjunto de neurônios poderia calcular qualquer função lógica e aritmética (MCCULLOCH; PITTS, 1943). Inicialmente, o modelo era bastante simplificado e se assemelhava bastante aos neurônios biológicos conhecidos na época. A grande ascensão da ideia proposta por Pitts e McCulloch veio através do livro (HEBB, 2012) publicado por Donald Hebb em 1949. Que mostrou que os neurônios quanto mais utilizados, mais forte ficam, isto é, quando estamos aprendendo algo novo, o nosso cérebro ativa os neurônios, que se conectam com outros, formando uma rede neural. A grande contribuição de Hebb é que essas conexões no começo são fracas e com o passar do tempo se fortalecem a cada novo estímulo. Alguns anos depois, surgiram vários trabalhos cuja ideia era propor métodos de aprendizado para redes neurais (ROSENBLATT, 1958; WIDROW; HOFF, 1962; KOHONEN, 1972; GROSSBERG, 1970). No entanto, Schmidhuber (2015) destaca que as primeiras redes neurais criadas eram apenas variações da regressão linear.

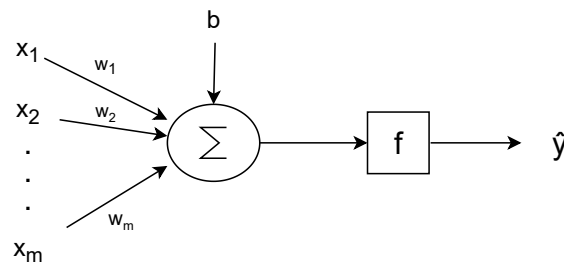
Apesar da denominação neurônio artificial para o principal elemento de uma rede neural, ele possui poucas semelhanças com um neurônio biológico. Aspectos comuns entre eles são a capacidade de receber e emitir e a capacidade de serem combinados com estruturas mais complexas.

Neste contexto, podemos escrever a saída desses neurônios por

$$\hat{y} = f\left(\sum_{i=1}^m w_i x_i + b\right),$$

onde $\mathbf{x} = [x_1, x_2, x_3, \dots, x_m]$ é o vetor que representa a entrada dos dados, $w = [w_1, w_2, w_3, \dots, w_m]$ é o vetor de pesos, b é chamado de *bias* e f é a função de ativação.

Figura 3 – Neurônio artificial



Fonte: Elaborada pelo autor.

A função de ativação tem como uma das finalidades evitar o aumento gradativo dos valores das saídas e facilitar o processo de aprendizagem. Existem diversas funções de ativação na literatura, as mais populares e mais usadas são:

- $f(z) = \frac{1}{1 + e^{-z}}$;
- $f(z) = \max(0, z)$;
- $f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$;
- $f(z) = z$.

O primeiro modelo de rede capaz de aprender foi apresentado por [Rosenblatt \(1958\)](#) e utilizou como função de ativação a função degrau. Em geral, este modelo é limitado a aprender somente funções lineares por ter variações bruscas na saída. No entanto, se mostrou bastante eficaz para aprender as portas lógicas AND e NAND.

Quando alteramos a quantidade de neurônios e camadas presentes num modelo, estamos criando estruturas complexas de redes, que denominamos arquiteturas de redes neurais. Um exemplo comum é um modelo no qual os neurônios são conectados uns aos outros, onde a saída de um neurônio pode se conectar a outro ou à sua própria entrada. Este tipo de arquitetura é conhecida por *feed-forward*, também podendo ser denominada como rede *multilayer-perceptron*. O termo *feed-forward* refere-se ao fato de que o fluxo dos dados de entrada flui diretamente para a saída sem efetuar *loops* na estrutura.

O processo de treinamento de uma rede neural é feito da mesma forma que em um neurônio simples, atualizando os pesos e os vieses por meio de um algoritmo de otimização. Se considerarmos um problema de aprendizado supervisionado, teremos uma rede neural que receberá uma entrada x e produzirá uma saída \hat{y} . Como queremos $\hat{y} \approx y$ (valor esperado), isto gerará um erro cujo valor podemos calcular usando o erro quadrático médio ou o erro logarítmico. Para minimizar o erro, o gradiente da função de erro é calculado usando o *backpropagation*, criado por [Rumelhart, Hinton e Williams \(1986\)](#). Com os gradientes calculados, um método de

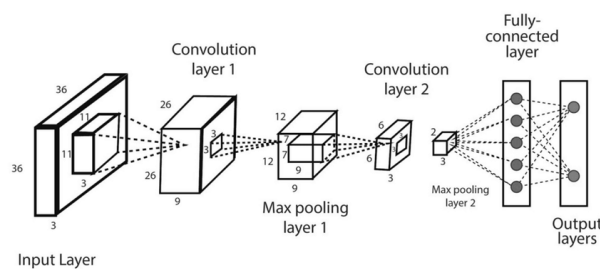
otimização especializados para encontrar o mínimo, como o gradiente descendente estocástico (GOODFELLOW; BENGIO; COURVILLE, 2016), AdaGrad (DUCHI; HAZAN; SINGER, 2011) e Adam (KINGMA; BA, 2014) são utilizados.

Outra arquitetura popular são as redes neurais convolucionais, propostas inicialmente por LeCun *et al.* (1998). O diferencial dessa arquitetura é poder ter como entrada matrizes ou imagens em vez de simplesmente vetores. Isso auxiliou a redução da quantidade de parâmetros que precisariam ser treinados. Além disso, permitiu utilizar certas propriedades que anteriormente não seriam possíveis, por exemplo, matrizes de convolução, que facilitam a extração de características do conjunto de dados. Esta capacidade de extrair características é fundamental na classificação de imagens, pois a localização de um píxel pode mudar completamente seu significado.

A estrutura de uma rede neural convolucional é composta por camadas (ver Figura 4), onde cada uma tem uma função particular, e cada etapa é responsável por aprender padrões encontrados no conjunto de entrada. Dentre estas camadas, as principais e mais utilizadas são:

- Camada de convolução: extrai informações da imagem através de filtros, o que ocorre por meio da aplicação deste filtro em pequenos blocos da imagem;
- *Pooling*: é utilizada para reduzir o tamanho espacial do modelo e possivelmente a quantidade de parâmetros da rede;
- Camada totalmente conectada: é utilizada no processo de classificar as informações extraídas das camadas anteriores. Esta camada achata o bloco anterior e transforma num vetor para realizar o processo de classificação.

Figura 4 – Estrutura de uma rede neural convolucional simples.



Fonte: Retirado de (SARKER, 2021).

ABORDAGEM PROPOSTA

Embora alguns métodos de seleção de parâmetro existentes sejam bastante utilizados, muitos deles não garantem uma boa aproximação da solução exata. Portanto, escolher um parâmetro “ótimo” é fundamental para a qualidade da solução aproximada.

Trabalhos anteriores foram realizados utilizando métodos de aprendizado profundo para o cálculo do parâmetro de regularização e mostraram grandes benefícios ao utilizar essa abordagem, como eficiência computacional e generalização. No entanto, essa abordagem pode ser melhorada ao utilizar novas maneiras de calcular o conjunto de dados. Para este trabalho introduziremos uma nova abordagem que utiliza novas métricas para criar novos conjuntos de dados. Com isso, podemos treinar uma rede neural para aproximar o mapeamento dos dados de entrada ao parâmetro de regularização. Uma vez que a rede já está treinada, podemos inserir novos dados e obter o parâmetro correspondente.

Para a abordagem, utilizaremos como base o trabalho de [Afkham, Chung e Chung \(2021\)](#), onde alteramos a forma de calcular o conjunto de dados utilizando novas métricas para medir a qualidade da solução. [Afkham, Chung e Chung \(2021\)](#) utilizam a seguinte abordagem para calcular o parâmetro de regularização para a criação do conjunto de dados

$$\alpha^* = \arg \min_{\alpha} \|\hat{x}(\alpha) - \mathbf{x}^*\|_2,$$

onde

$$\hat{x}(\alpha) = \arg \min_x \|T\mathbf{x} - \mathbf{y}\|_2^2 + \alpha \|\mathbf{x}\|_2^2$$

com \mathbf{x}^* sendo a solução exata. Essa maneira de calcular o parâmetro é interessante, porém em problemas de imageamento, a diferença entre píxeis pode não ser muito usual para avaliar a qualidade de uma imagem reconstruída. Por conta disso, a utilização de métricas de imagens no cálculo do parâmetro é uma excelente alternativa. Além disso, métricas de imagens podem considerar fatores importantes como nitidez, contraste e foco. Sendo assim, podem ter um impacto significativo na percepção na qualidade da imagem.

Neste trabalho não será utilizada unicamente uma função de erro para todo tipo de problema tratado. Sendo assim, calculamos o parâmetro da forma

$$\alpha^* = \arg \min_{\alpha} \ell(\hat{x}(\alpha), \mathbf{x}^*), \quad (5.1)$$

onde $\ell(\cdot, \cdot)$ é uma função de erro definida conforme o problema.

Com o auxílio das técnicas de aprendizado profundo, podemos utilizar os dados de treinamento disponíveis ou que foram gerados sinteticamente, para escolher o valor do parâmetro de regularização. A escolha da arquitetura de rede ideal dependerá do problema que está sendo tratado, por exemplo, para problemas em \mathbb{R}^2 uma arquitetura composta por redes neurais convolucionais tende a ser uma ótima alternativa. Então, temos que os dados de treinamento estarão no seguinte formato $\{\mathbf{y}^{(i)}, \alpha^{*(i)}\}_{i=1}^I$, onde I é a quantidade de dados e $\mathbf{y}^{(i)}$ é a observação ruidosa.

O processo de criação do conjunto de dados, treinamento da rede e estimação do parâmetro de regularização são divididos em duas fases. A primeira, conhecida como fase “Geração do conjunto de dados e treinamento da rede neural”, envolve a geração do conjunto de dados e treino da rede. Note que nesta fase é onde ocorre todo esforço computacional. A segunda fase, chamada de “Previsão do parâmetro”, consiste na propagação direta na rede para uma nova observação e obtenção do parâmetro de regularização. Nesta fase, quase não tem esforço computacional nenhum, visto que a propagação é executada rapidamente. Estas etapas estão descritas no Algoritmo 1.

Algoritmo 1 – Seleção do parâmetro de regularização utilizando aprendizado profundo.

- 1: **Geração do conjunto de dados e treinamento da rede neural:**
 - 2: Gere novas observações $\mathbf{y}^{(i)} = T \mathbf{x}^{*(i)} + \boldsymbol{\varepsilon}^{(i)}$, para $i = 1, 2, \dots, I$
 - 3: Defina a função de erro $\ell(\cdot, \cdot)$
 - 4: Calcule $\alpha^{*(i)}$ (resolva (5.1))
 - 5: Defina a arquitetura da rede neural
 - 6: Use os dados de treino para o treinamento da rede
 - 7: **Previsão do parâmetro:**
 - 8: Obtenha uma nova observação \mathbf{y}
 - 9: Propague \mathbf{y} na rede treinada para obter α
-

Com o parâmetro escolhido, é possível calcular a solução regularizada do problema inverso utilizando a regularização de Tikhonov. Nos experimentos numéricos presentes no próximo capítulo utilizaremos esta abordagem para estimar o parâmetro de regularização e mostraremos os benefícios aplicados em três problemas inversos.

EXPERIMENTOS NUMÉRICOS

Para confirmar alguns dos benefícios da abordagem definida no capítulo anterior, serão fornecidos alguns resultados obtidos para três problemas inversos: calor, *deblurring* e tomografia computadorizada. Para efeito de comparação, os resultados provenientes do problema calor serão comparados com três métodos de seleção de parâmetro: discrepância, validação cruzada generalizada e curva L. Nos demais problemas os resultados serão comparados com a abordagem de [Afkham, Chung e Chung \(2021\)](#), no qual se obteve vantagem em relação aos métodos convencionais. As estruturas das redes neurais foram determinadas empiricamente, mediante um processo de experimentação e ajuste contínuo. Várias configurações foram testadas, analisando o desempenho de cada uma, até que a melhor arquitetura fosse selecionada.

6.1 Especificações

Todos os experimentos foram realizados em um computador equipado com um processador Intel Core i7 de 2,80 GHz, 8 GB de memória RAM, uma placa de vídeo NVIDIA GeForce GTX 745 com 4 GB de memória DDR3 e executando o sistema operacional Linux. Todos os métodos foram implementados utilizando a linguagem de programação Python 3. Para a construção das redes neurais foi utilizado a biblioteca PyTorch.

6.2 Calor

Neste experimento consideramos o problema inverso do calor descrito na Seção 2.2. O problema tratado é contínuo, por conta disso é reduzido a um sistema finito de equação lineares. Para isto, o *kernel* descrito pela equação (2.4) é truncado para $n = 10$. A solução da integral de Fredholm descrito pela equação (2.3) é resolvida numericamente pela regra dos trapézios composta ([PHILLIPS, 1962](#)). A próxima etapa é discretizar o intervalo $[0, L]$ em $m - 1$

subintervalos de tamanho h definidos pelos pontos $x_i = ih$ e $y_j = jh$, onde $i = 0, 1, 2, \dots, m-1$, $j = 0, 1, 2, \dots, m-1$ e $h = \frac{L}{m}$.

Com as configurações acima, podemos reescrever o problema inverso do calor como um problema de equação lineares da forma

$$A\mathbf{u} = \mathbf{g},$$

onde \mathbf{u} é a solução inicial do problema, \mathbf{g} é a solução final no instante T e a matriz A é da forma

$$\frac{L}{m} \begin{bmatrix} k(x_0, y_0) & k(x_0, y_1) & k(x_0, y_2) & \cdots & k(x_0, y_{m-1}) \\ k(x_1, y_0) & k(x_1, y_1) & k(x_1, y_2) & \cdots & k(x_1, y_{m-1}) \\ k(x_2, y_0) & k(x_2, y_1) & k(x_2, y_2) & \cdots & k(x_2, y_{m-1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ k(x_{m-1}, y_0) & k(x_{m-1}, y_1) & k(x_{m-1}, y_2) & \cdots & k(x_{m-1}, y_{m-1}) \end{bmatrix}$$

Para utilizarmos a estratégia abordada, é necessário a definição de uma função de erro, para isto utilizaremos

$$\ell(\hat{\mathbf{x}}(\alpha), \mathbf{x}^*) = \|\mathbf{x}^* - \hat{\mathbf{x}}(\alpha)\|_2. \quad (6.1)$$

Na geração do conjunto de dados, consideraremos o seguinte problema

$$\begin{aligned} u_t &= u_{xx}, & 0 < x < \pi, & \quad t > 0; \\ u(0, t) &= 0, & t &\geq 0; \\ u(\pi, t) &= 0, & t &\geq 0; \\ u(x, 0) &= 2 \sin(3x) - 3 \sin(4x), & 0 < x < \pi, \end{aligned}$$

cuja a solução exata é dada pela função

$$u(x, t) = 2e^{-9t} \sin(3x) - 3e^{-16t} \sin(4x).$$

O espaço é discretizado utilizando $m = 250$, o tempo escolhido é $T = 0.1$ e os dados da solução exata são contaminados por um ruído $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$ para uma variância σ^2 escolhida aleatoriamente entre o intervalo $(10^{-3}, 10^{-1})$ distribuído uniformemente. Para encontrar o parâmetro de regularização para cada dado gerado, utilizamos dois métodos. O primeiro, é o método do gradiente conjugado para resolver a regularização de Tikhonov. O segundo, é o *golden-section* (KIEFER, 1953), utilizado para encontrar o mínimo de uma função escalar. Métodos que estão implementados, validados e disponíveis na biblioteca SciPy (VIRTANEN *et al.*, 2020). Com essas definições, geramos 400.000 dados, onde 80% é utilizado para treinamento e 20% para validação.

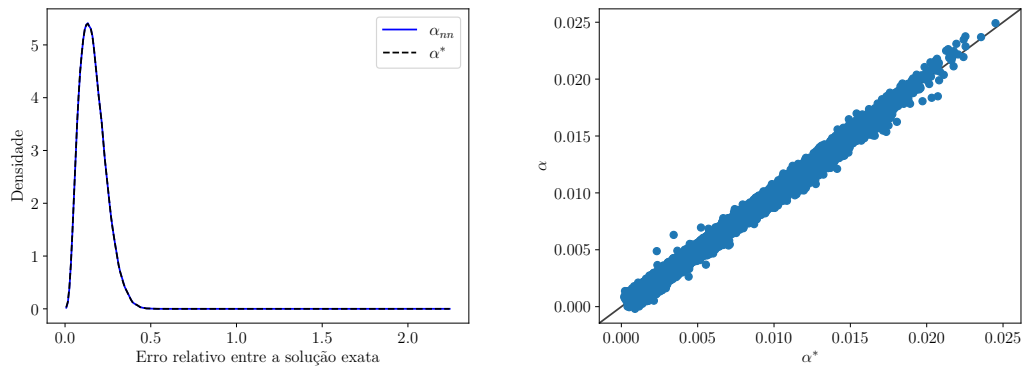
A rede neural utilizada é composta por 5 camadas ocultas com 75, 50, 25, 12 e 6 neurônios, respectivamente. A função de ativação escolhida é a ReLU, e os hiperparâmetros utilizados para esta rede podem ser visualizados na Tabela 1.

Tabela 1 – Resumo dos hiperparâmetros para o problema do calor.

Hiperparâmetros	
Épocas	10
Batch size	32
Otimizador	Adam
Parâmetro de aprendizado	10^{-4}
Função de perda	MSE

O desempenho do treinamento da rede neural realizado na fase de previsão é mostrado na Figura 5, onde é ilustrado no lado direito a dispersão das previsões dos parâmetros de regularização realizadas pela rede no conjunto de validação e, no lado esquerdo, o erro relativo entre a solução exata para o problema resolvido utilizando o parâmetro definido no conjunto e o estimado pela rede. A estratégia empregada, que utiliza técnicas de aprendizado profundo, demonstrou um bom desempenho quando comparada ao método selecionado para determinar o parâmetro α^* . Isso é evidente tanto na precisão da aproximação do valor quanto na discrepância entre a solução aproximada e a exata. É importante ressaltar que, embora essa abordagem tenha se mostrado eficaz para este problema específico, ainda é necessário avaliar seu desempenho em outros exemplos. Para verificar a generalização das redes neurais e fins de comparações com as

Figura 5 – A parte esquerdo da figura demonstra o erro relativo no conjunto de validação para os dois métodos. A parte direita ilustra a dispersão das previsões dos valores do parâmetro de regularização.



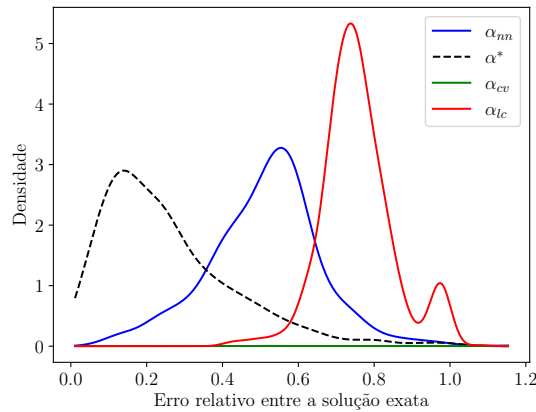
Fonte: Elaborada pelo autor.

técnicas de seleção de parâmetro tradicionais, utilizamos outra equação do calor para o próximo experimento. Consideramos a equação

$$\begin{aligned}
 u_t &= 2u_{xx}, & 0 < x < \pi, & \quad t > 0; \\
 u(0, t) &= 0, & t & \geq 0; \\
 u(\pi, t) &= 0, & t & \geq 0; \\
 u(x, 0) &= \sin(2x) & 0 < x < \pi,
 \end{aligned}$$

onde a solução exata é dada por $u(x,t) = \sin(2x)e^{-8t}$. Para esse exemplo, geramos 1000 dados ruidosos com a mesma configuração definida no conjunto de dados e utilizamos $T = 0.2$. Os métodos utilizados para comparação foram a validação cruzada generalizada, curva-L e o método utilizado para gerar o conjunto de dados da rede. O princípio da discrepância não foi utilizado por obter um desempenho ruim nos dados utilizados. Na Figura 6 providenciamos o erro relativo entre a solução exata para os métodos acima explicitados, onde α_{nn} é o parâmetro estimado pela rede, α^* é o parâmetro definido no conjunto de dados, α_{cv} é o parâmetro definido pelo método da validação cruzada e α_{lc} é o parâmetro gerado pelo método da curva L. Resumidamente, para

Figura 6 – Erro relativo entre os métodos de seleção do parâmetro de regularização para o problema inverso do calor.



Fonte: Elaborada pelo autor.

novos dados a abordagem utilizando rede neural obteve um desempenho superior em relação aos métodos tradicionais, onde o erro relativo se manteve menor do que tais métodos. Isto vem do fato da grande vantagem em utilizar redes neurais, a generalização. Como o principal esforço computacional da abordagem está contida na fase de geração do conjunto de dados e treinamento da rede, o tempo necessário para se calcular o parâmetro para um novo dado é significativamente pequeno em comparação aos métodos tradicionais.

6.3 Deblurring

Nesta subseção tratamos do problema inverso de *deblurring*, onde é desejado estimar a partir de uma imagem borrada, a imagem original. Para este experimento consideramos um *kernel* gaussiano de tamanho 5×5 com variância de 2.0. Matematicamente, este problema pode ser reformulado como um sistema linear da forma

$$Ax = b + \varepsilon,$$

onde $A \in \mathbb{R}^{m \times n}$ representa a matriz de convolução, $x \in \mathbb{R}^n$ a imagem original discretizada achatada, $b \in \mathbb{R}^m$ a imagem borrada discretizada achatada e $\varepsilon \in \mathbb{R}^m$ um ruído aditivo.

Neste problema, para gerar o conjunto de dados consideramos o conjunto de dados BSD68 (ROTH; BLACK, 2009). Este conjunto é formado por 68 imagens em tons de cinza de tamanhos 481×321 e 321×481 . Sendo composto por uma variedade de imagens, desde imagens naturais até de imagens de pessoas. Algumas destas imagens podem ser visualizadas na Figura 7. Com o intuito de aumentar o conjunto de dados, duplicamos cada imagem contida em BSD68

Figura 7 – Exemplos de imagens contidas no conjunto BSD68.



Fonte: Elaborada pelo autor.

com sua imagem rotacionada a 90° no sentido anti-horário. Outra importante manipulação, é a extração de 8 *patches* de tamanho 256×256 para cada imagem (ver Figura 8). Com todos esses passos, o conjunto de dados reformulado agora contém 1088 imagens. O nível de ruído é representado por um valor randômico selecionado uniformemente entre 1% e 5%. Onde, por exemplo, um ruído de 4% significa que $\frac{\|\epsilon\|_2}{\|Ax\|_2} = 0.04$, para um ϵ escolhido randomicamente. Com a configuração dos ruídos definida, para cada imagem do conjunto de dados realizamos a geração de 10 imagens ruidosas. Deste modo, o conjunto de dados final após realizado todas as alterações tem um tamanho total de 10.880 imagens em tons de cinza. Neste experimento,

Figura 8 – *Patches* obtidos de uma imagem do conjunto BSD68.



Fonte: Elaborada pelo autor.

definiremos duas funções de erro, a primeira, ℓ_1 proposta por Afkham, Chung e Chung (2021), a segunda, ℓ_2 proposta neste trabalho, onde a função é composta pela métrica SSIM (*structural similarity index measure*) (WANG *et al.*, 2004) e MSE (*mean squared error*). Sendo assim, as funções de erro são definidas da forma

$$\ell_1(\hat{x}(\alpha), x^*) = \|x^* - \hat{x}(\alpha)\|_2, \quad (6.2)$$

$$\ell_2(\hat{x}(\alpha), x^*) = \text{MSE}(\hat{x}(\alpha)) \cdot (1 - \text{SSIM}(\hat{x}(\alpha), x^*)). \quad (6.3)$$

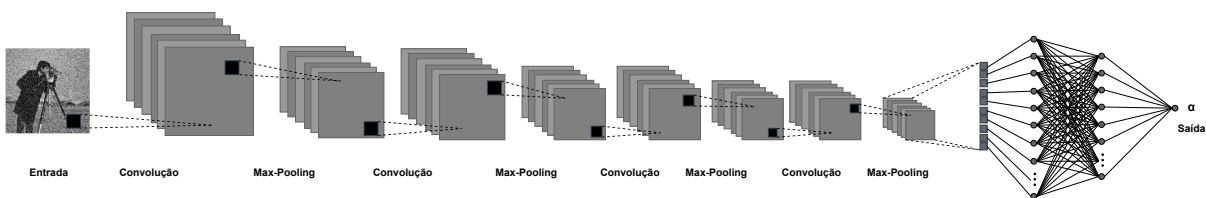
A função escolhida leva em sua definição duas importantes métricas de imagens bastante utilizadas, o MSE, sendo o erro quadrático médio entre os pixels das imagens, a outra, é o SSIM,

utilizado para medir a similaridade entre duas imagens, onde se considera algumas características-chaves de uma imagem, a luminância, contraste e estrutura. Portanto, todos os testes realizados abaixo consideram dois conjuntos de dados distintos, um formado pela função ℓ_1 e, outro por ℓ_2 .

O próximo passo é resolver 5.1 e obter o parâmetro de regularização para cada imagem no conjunto. Para isso utilizamos dois métodos. O primeiro, é o LSMR (*Least Squares Minimum Residual*) (FONG; SAUNDERS, 2011), utilizado para resolver problemas de quadrados mínimos com regularização. Este método está implementado, validado e disponível na biblioteca CuPyx (NISHINO; LOOMIS, 2017). O segundo, é o *golden-section*. Como este tipo de problema requer uma abundante quantidade de memória e esforço computacional, foram utilizadas bibliotecas com suporte a GPU para otimizar o tempo computacional.

Com o conjunto de dados gerado, utilizamos para este problema uma arquitetura composta por camadas convolucionais. Uma abordagem bastante utilizada quando o conjunto de dados é formado exclusivamente por imagens. A configuração da rede escolhida é formada por 4 camadas convolucionais, cada uma composta por 32 canais. Cada camada de convolução possui um *kernel* de tamanho 3×3 , com um *stride* e *padding* ambos definidos como 1. Além disso, adicionamos entre as camadas de convolução uma operação de *max-pooling* com dimensões 2×2 , *stride* igual a 2 e *padding* de 0. As camadas finais são compostas por duas camadas *multi-layer perceptron* com 8192 e 512 neurônios, respectivamente. Uma melhor representação da arquitetura proposta pode ser vista na Figura 9.

Figura 9 – Arquitetura proposta para o problema de estimar o parâmetro de regularização para o problema de *deblurring*.



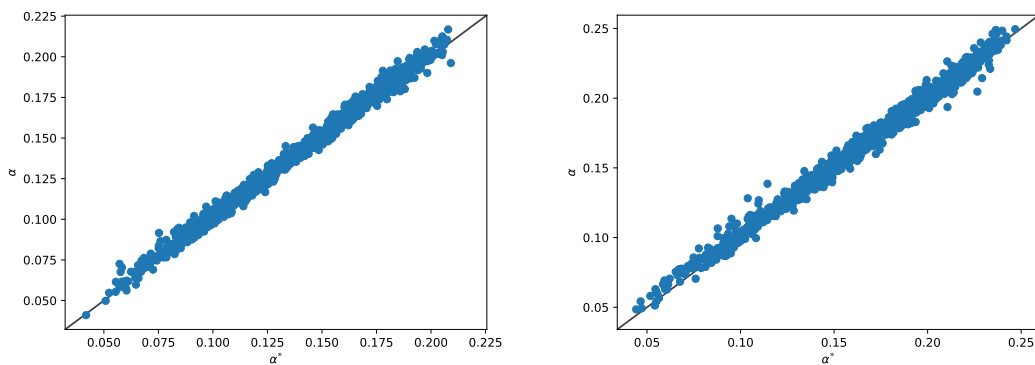
Fonte: Elaborada pelo autor.

Para realizar o treino da rede, é necessário a definição de alguns hiperparâmetros para o pleno funcionamento. Na Tabela 2 estão estabelecidos os hiperparâmetros utilizados para este problema. Após o treinamento, o desempenho realizado pela rede escolhida nos dois conjuntos de dados podem ser vistos na Figura 10, onde no lado esquerdo é possível visualizar a dispersão dos parâmetros estipulados para o conjunto formado pela função ℓ_1 e, do lado direito, formado pela função ℓ_2 .

Para avaliarmos o poder da generalização da rede, escolhemos três imagens de testes, nomeadamente *camera man*, *peppers* e *butterfly*, todas de tamanho 256×256 (ver Figura 11). As imagens de testes foram corrompidas pelo mesmo *kernel* e intervalo de ruído utilizado nos dados de treinamento. Foram gerados 1000 imagens com diferentes níveis de ruídos e,

Tabela 2 – Resumo dos hiperparâmetros para o problema de *deblurring*.

Hiperparâmetros	
Épocas	20
Batch size	32
Otimizador	Adam
Parâmetro de aprendizado	10^{-3}
Função de perda	MSE

Figura 10 – Dispersão dos valores estipulados pela rede para o problema de *deblurring* no conjunto de validação para a função ℓ_1 (esquerdo) e ℓ_2 (direito).

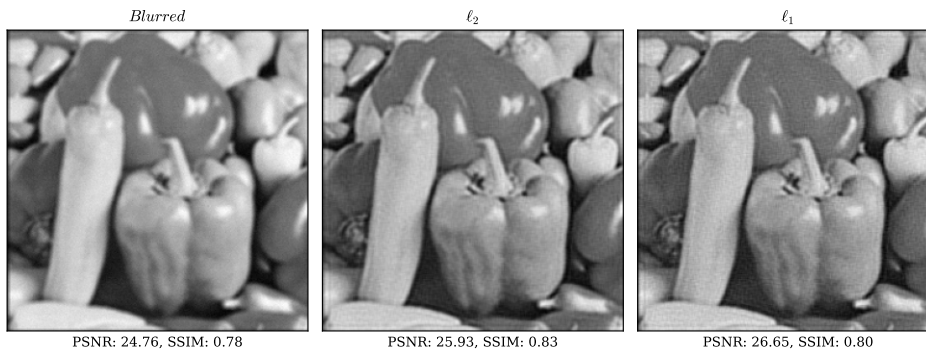
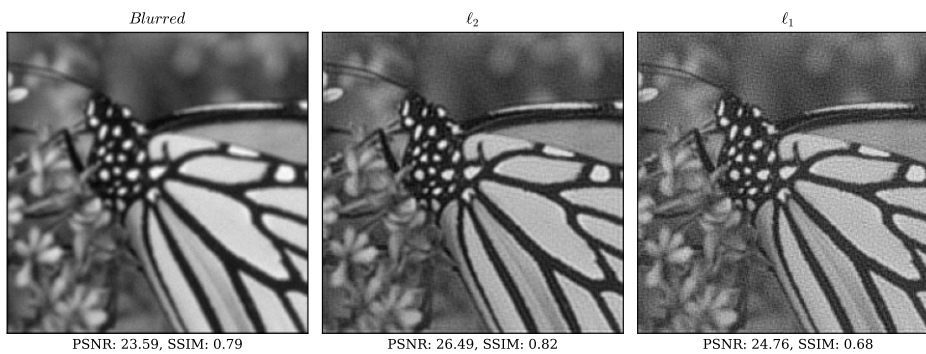
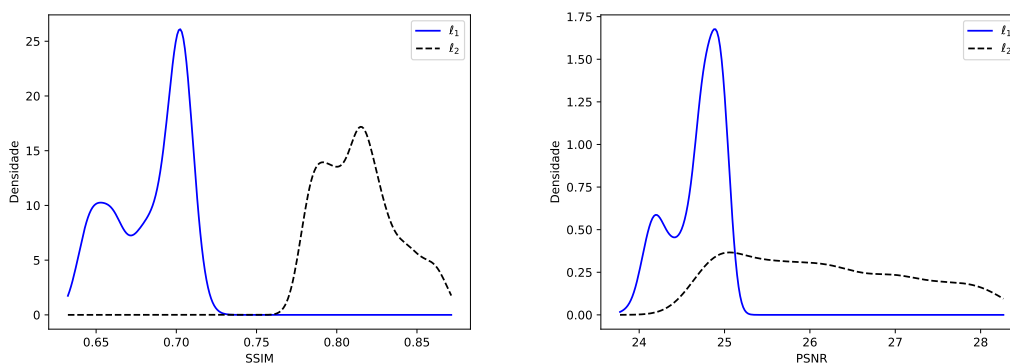
Fonte: Elaborada pelo autor.

para fins de comparação, consideramos duas métricas para avaliar a qualidade da imagem reconstruída, o PSNR (*peak signal-to-noise ratio*) e SSIM. Na Figura 12 e 13 é possível visualizar os resultados para um determinado erro. Para avaliar os resultados para todas as imagens de teste,

Figura 11 – Imagens de testes utilizadas para o problema de *deblurring*.

consideramos os gráficos de densidade, que podem ser vistas nas Figuras 14, 15 e 16. Com base na comparação qualitativa entre a função de erro utilizada neste trabalho e a apresentada por Afkham, Chung e Chung (2021), os resultados indicam uma melhoria na qualidade das soluções propostas.

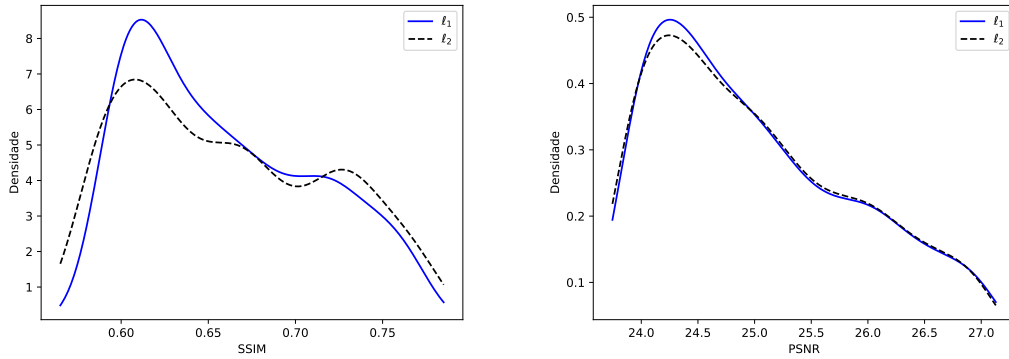
Ao analisar os resultados para a imagem *butterfly*, destacam-se vantagens substanciais em ambas as métricas, demonstrando a eficácia consistente da abordagem proposta. Essa vantagem é claramente verificada pelo gráfico de densidade, reforçando a robustez da proposta. No caso

Figura 12 – Resultados obtidos para a imagem *peppers* nas duas abordagens utilizando um erro de 2,5%.Figura 13 – Resultados obtidos para a imagem *butterfly* nas duas abordagens utilizando um erro de 2,5%.Figura 14 – Gráficos de densidade para a imagem *butterfly*. O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.

Fonte: Elaborada pelo autor.

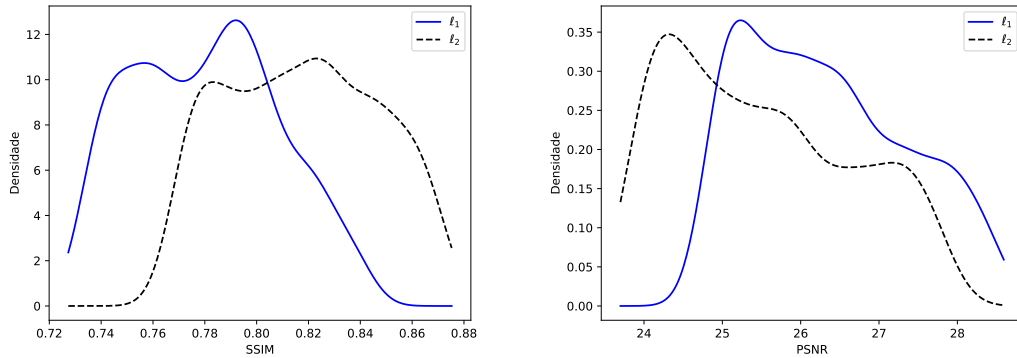
da imagem *camera man*, a métrica SSIM revelou áreas específicas onde a abordagem proposta apresentou melhor desempenho, enquanto a métrica PSNR demonstrou uma pequena vantagem. Na imagem *peppers*, a abordagem proposta exibiu um desempenho superior na métrica SSIM, embora tenha ficado em desvantagem em relação ao PSNR.

Figura 15 – Gráficos de densidade para a imagem *camera-man*. O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.



Fonte: Elaborada pelo autor.

Figura 16 – Gráficos de densidade para a imagem *peppers*. O lado esquerdo mostra a densidade para a métrica SSIM. No lado direito, mostra a densidade para a métrica PSNR.



Fonte: Elaborada pelo autor.

6.4 Tomografia Computadorizada

A tomografia computadorizada é uma técnica de processamento de imagens bastante utilizada na medicina para a visualização de estruturas internas de um corpo humano. O procedimento consiste em projetar um feixe em vários ângulos diferentes, medir a intensidade do feixe e com isto calcular a reconstrução da imagem. Neste experimento, tratamos de imagens em tons de cinza de tamanho 128×128 discretizada numa malha $[-1, 1]^2$. O intervalo dos ângulos escolhidos e as coordenadas dos raios são:

$$\Theta := \left\{ 0, \frac{\pi}{181}, 2\frac{\pi}{181}, \dots, 180\frac{\pi}{181} \right\},$$

$$L := \left\{ -1, -1 + \frac{2}{179}, -1 + 2\frac{2}{179}, \dots, -1 + 178\frac{2}{179}, 1 \right\}.$$

Com as projeções definidas e utilizando o método de Siddon, onde a intensidade do feixe num píxel da imagem é dado pelo comprimento da interseção entre os dois, é possível construir a matriz A .

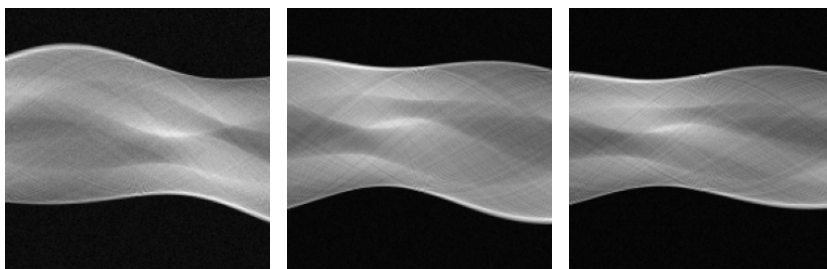
Para a geração do conjunto de dados, utilizamos 14.500 imagens *Shepp-Logan* aleatórias de tamanho 128×128 (ver Figura 17). A implementação para gerar estes exemplos está disponível em Chung (2018), porém o código-fonte está na linguagem de programação Matlab. Para se adequar as configurações deste trabalho, o código foi convertido para Python. Todas as configurações de ruído e métodos de soluções utilizados para o problema de *deblurring* visto na seção anterior foram considerados neste experimento. As mesmas funções de erro, (6.2) e (6.3), foram utilizadas também. Portanto, o conjunto de dados é composto pelos sinogramas ruidosos e o parâmetro de regularização calculado. Na Figura 18 é possível visualizar alguns sinogramas.

Figura 17 – Exemplos de *Shepp-Logan* contidas no conjunto de dados.



Fonte: Elaborada pelo autor.

Figura 18 – Exemplos de sinogramas contidas no conjunto de dados.

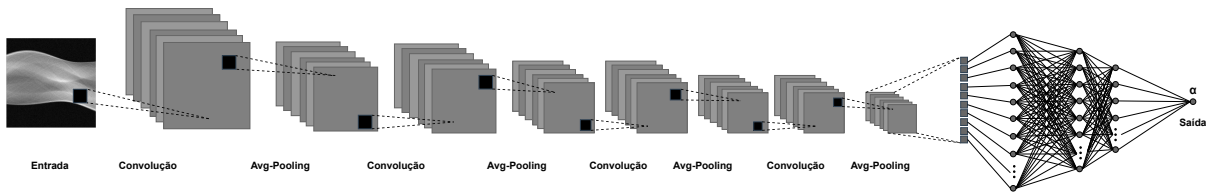


Fonte: Elaborada pelo autor.

A arquitetura da rede utilizada para este experimento é composta por 4 camadas convolucionais, com o mesmo número de canais, função de ativação e *kernel* utilizado para o problema de *deblurring*. A discrepância entre elas é a mudança do *padding* para 1 e a utilização de *average pooling* em vez de um *max-pooling*. As camadas finais para esta rede são compostas por 3 *multi-layer perceptron* com 15488, 256 e 128 neurônios, respectivamente. A Figura 19 apresenta uma representação da arquitetura.

Os hiperparâmetros utilizados para realizar o treinamento desta rede encontra-se na Tabela 3. O desempenho da rede nos dois conjuntos de dados depois do treinamento pode ser visualizado na Figura 20. A dispersão dos parâmetros estipulados para o conjunto de dados

Figura 19 – Arquitetura proposta para o problema de estimar o parâmetro de regularização para o problema de tomografia computadorizada.



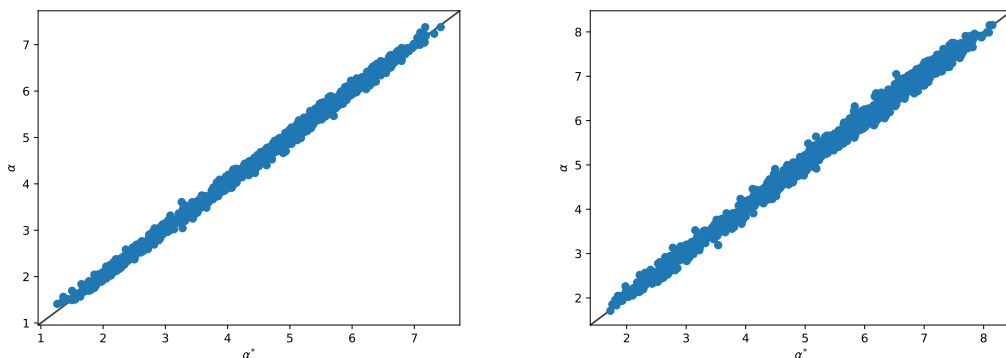
Fonte: Elaborada pelo autor.

formado pela função ℓ_1 pode ser observado no lado esquerdo, enquanto no lado direito a dispersão para a outra função.

Tabela 3 – Resumo dos hiperparâmetros para o problema de tomografia computadorizada.

Hiperparâmetros	
Épocas	30
Batch size	64
Otimizador	Adam
Parâmetro de aprendizado	10^{-3}
Função de perda	MSE

Figura 20 – Dispersão dos valores estipulados pela rede para o problema de tomografia computadorizada no conjunto de validação para a função ℓ_1 (esquerdo) e ℓ_2 (direito).



Fonte: Elaborada pelo autor.

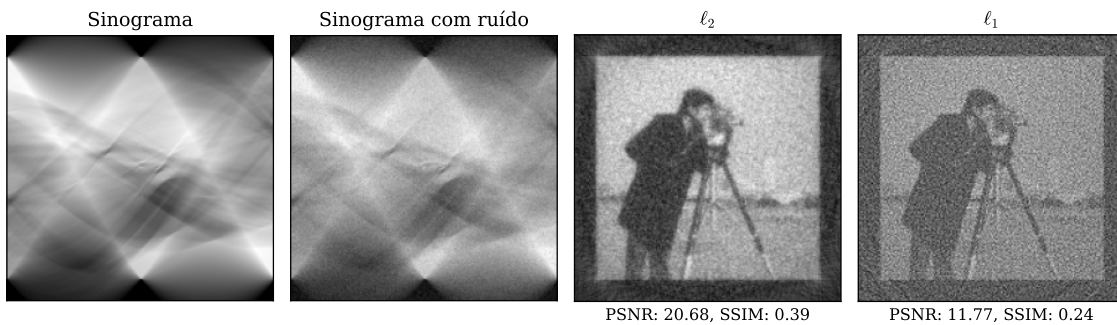
Para realizar os testes para verificar a generalização da rede, utilizaremos para isto a imagem *camera man* de tamanho 180×180 (ver Figura 21). A realização das projeções é feita utilizando as mesmas medidas utilizadas na criação do conjunto de dados. Com isso, foram gerados 1000 sinogramas corrompidas pelo mesmo intervalo de ruído utilizado no conjunto. Uma visualização da qualidade da reconstrução que as duas abordagens podem proporcionar pode ser visualizada na Figura 22.

Figura 21 – Imagem do *camera man* utilizado para o problema de tomografia computadorizada.



Fonte: Elaborada pelo autor.

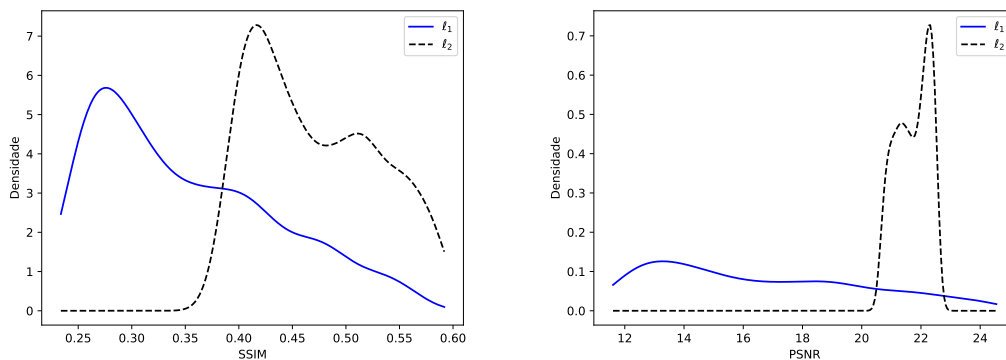
Figura 22 – Resultados obtidos na reconstrução do *camera man* para o sinograma com ruído de 5%.



Fonte: Elaborada pelo autor.

Como utilizado anteriormente no problema de *deblurring*, utilizaremos as métricas SSIM e PSNR para avaliar a qualidade das soluções. Os resultados adquiridos podem ser visualizados na Figura 23.

Figura 23 – Gráfico de densidade para a tomografia computadorizada do *camera man*. A densidade da métrica SSIM é exibida à esquerda, enquanto a da métrica PSNR é mostrada à direita.



Fonte: Elaborada pelo autor.

Analisando os resultados deste teste, destacam-se vantagens em ambas as métricas utilizadas para avaliar a qualidade da solução. Esta vantagem, é notória ao visualizar os gráficos de densidade gerados. Estes resultados, sustentam que a estratégia desenvolvida neste trabalho proporcionar uma melhoria na qualidade das soluções.

CONCLUSÕES

7.1 Considerações finais

Neste trabalho explorei três problemas inversos comumente conhecidos: calor, *deblurring* e tomografia computadorizada. Foram estudados fundamentos essenciais, como espaços de Hilbert, operadores lineares, operador inverso de Moore-Penrose e regularização. Essa fundamentação teórica foi crucial para o entendimento da teoria sobre problemas inversos lineares.

Ao longo deste trabalho, foi notada a importância da escolha do parâmetro de regularização em problemas inversos com dados ruidosos. Evidenciamos alguns métodos tradicionais utilizados para este propósito, porém foram notados algumas limitações destes métodos, como, estimativa precisa do ruído e custo computacional. Por conta disso, foram utilizadas duas abordagens, a primeira, utilizado por [Afkham, Chung e Chung \(2021\)](#) e, a segunda, proposta por este trabalho. Os resultados numéricos mostrados sustentam a conclusão de que a abordagem desenvolvida supera a metodologia proposta por [Afkham, Chung e Chung \(2021\)](#), proporcionando melhorias na qualidade das soluções.

7.2 Limitações e melhorias futuras

Com estudo deste trabalho, foi possível identificar limitações ao utilizar a abordagem. Como limitação, evidenciamos que a aplicação foi realizada para problemas específicos, por exemplo, a utilização do *kernel* gaussiano com uma variância definida para o problema de *deblurring*. Em problemas práticos, o *kernel* e a variância podem não serem conhecidos, sendo necessário avaliar como a abordagem se comportaria para novos contextos.

Outro ponto a ser considerado é a necessidade de obter um grande conjunto de dados, caso em que seria necessário expor o problema inverso tratado a uma ampla gama de situações diversas. Essa estratégia melhoraria a capacidade de generalização do modelo. Isto também

incluir expor o problema inverso a situações atípicas, de modo que o modelo seria testado em sua capacidade de lidar com estas situações.

O resultado deste trabalho motiva a continuidade da pesquisa em relação a técnicas de seleção do parâmetro de regularização utilizando aprendizado profundo. Como melhorias futuras a serem realizadas, destacam-se: utilização de regularizações mais sofisticadas; aplicação em novos problemas inversos e tratamento de problemas no espaço não linear.

REFERÊNCIAS

- AFKHAM, B. M.; CHUNG, J.; CHUNG, M. Learning regularization parameters of inverse problems via deep neural networks. **Inverse Problems**, IOP Publishing, v. 37, n. 10, p. 105017, 2021. Citado nas páginas 22, 49, 51, 55, 57 e 65.
- ALLEN, D. M. The relationship between variable selection and data augmentation and a method for prediction. **technometrics**, Taylor & Francis, v. 16, n. 1, p. 125–127, 1974. Citado na página 42.
- BAKUSHINSKII, A. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. **USSR Computational Mathematics and Mathematical Physics**, Elsevier, v. 24, n. 4, p. 181–182, 1984. Citado na página 38.
- BAUER, F.; LUKAS, M. A. Comparing parameter choice methods for regularization of ill-posed problems. **Mathematics and Computers in Simulation**, Elsevier, v. 81, n. 9, p. 1795–1841, 2011. Citado na página 41.
- BECK, J.; BLACKWELL, B.; CLAIR, C. **Inverse heat conduction: Ill-posed problems**. New York: Wiley, 1985. Citado na página 25.
- BERTERO, M.; BOCCACCI, P.; MOL, C. D. **Introduction to inverse problems in imaging**. Boca Raton, Florida: CRC press, 2021. Citado na página 25.
- CHANG, J.-R.; YEIH, W.; SHIEH, M.-H. On the modified tikhonov's regularization method for the cauchy problem of the laplace equation. **Journal of Marine Science and Technology**, v. 9, n. 2, p. 6, 2001. Citado na página 42.
- CHUNG, M. **Random-Shepp-Logan-phantom**. 2018. Disponível em: <<https://github.com/matthiaschung/Random-Shepp-Logan-Phantom>>. Acesso em: 12 Dez. 2023. Citado na página 60.
- CLINE, A. K.; DHILLON, I. S. Computation of the singular value decomposition. Citeseer, 2006. Citado na página 36.
- DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive subgradient methods for online learning and stochastic optimization. **Journal of machine learning research**, v. 12, n. 7, 2011. Citado na página 47.
- ENGL, H. W.; HANKE, M.; NEUBAUER, A. **Regularization of inverse problems**. Dordrecht, The Netherlands: Springer Science & Business Media, 1996. v. 375. Citado nas páginas 21, 30, 37, 38 e 41.
- FOLLAND, G. B. **Real analysis: modern techniques and their applications**. New York; Chichester: John Wiley & Sons, 1999. v. 40. Citado na página 76.
- FONG, D. C.-L.; SAUNDERS, M. LSMR: An iterative algorithm for sparse least-squares problems. **SIAM Journal on Scientific Computing**, SIAM, v. 33, n. 5, p. 2950–2971, 2011. Citado na página 56.

GOLUB, G. H.; HEATH, M.; WAHBA, G. Generalized cross-validation as a method for choosing a good ridge parameter. **Technometrics**, Taylor & Francis, v. 21, n. 2, p. 215–223, 1979. Citado nas páginas 22 e 42.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge, Massachusetts: MIT press, 2016. Citado na página 47.

GOSZ, M. R. **Finite element method: applications in solids, structures, and heat transfer**. Boca Raton, Florida: CRC Press, 2017. Citado na página 25.

GROETSCH, C. W. **Inverse problems in the mathematical sciences**. Braunschweig, Germany: Springer, 1993. v. 52. Citado na página 37.

GROSSBERG, S. Neural pattern discrimination. **Journal of Theoretical Biology**, Elsevier, v. 27, n. 2, p. 291–337, 1970. Citado na página 45.

HADAMARD, J. **Lectures on Cauchy's problem in linear partial differential equations**. New York: Dover, 2003. Citado nas páginas 21 e 23.

HANSEN, P. C. Analysis of discrete ill-posed problems by means of the l-curve. **SIAM Review**, SIAM, v. 34, n. 4, p. 561–580, 1992. Citado nas páginas 21 e 43.

_____. **Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion**. Philadelphia: SIAM, 1998. Citado na página 42.

HANSEN, P. C.; O'LEARY, D. P. The use of the l-curve in the regularization of discrete ill-posed problems. **SIAM Journal on Scientific Computing**, SIAM, v. 14, n. 6, p. 1487–1503, 1993. Citado nas páginas 22 e 43.

HASHEMI, R.; BRADLEY, W.; LISANTI, C. **MRI: The Basics**. [S.l.]: Lippincott Williams & Wilkins, 2010. (The Basics Series). ISBN 9781608311156. Citado na página 21.

HEBB, D. O. **The organization of behavior : a neuropsychological theory**. New edition. New York: Routledge, 2012. Citado na página 45.

JOSEPH, P. M. An improved algorithm for reprojecting rays through pixel images. **IEEE transactions on medical imaging**, IEEE, v. 1, n. 3, p. 192–196, 1982. Citado na página 27.

KAK, A. C.; SLANEY, M. **Principles of computerized tomographic imaging**. Philadelphia: SIAM, 2001. Citado na página 21.

KELLER, J. B. Inverse problems. **The American Mathematical Monthly**, Mathematical Association of America, v. 83, n. 2, p. 107–118, 1976. Citado na página 21.

KIEFER, J. Sequential minimax search for a maximum. **Proceedings of the American mathematical society**, v. 4, n. 3, p. 502–506, 1953. Citado na página 52.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014. Citado na página 47.

KOHONEN, T. Correlation matrix memories. **IEEE transactions on computers**, IEEE, v. 100, n. 4, p. 353–359, 1972. Citado na página 45.

KREYSZIG, E. **Introductory functional analysis with applications**. New York: John Wiley & Sons, 1991. v. 17. Citado nas páginas 35, 37, 72, 73, 77 e 79.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado na página [47](#).

LIU, S.; ZHANG, J. Machine-learning-based prediction of regularization parameters for seismic inverse problems. **Acta Geophysica**, Springer, v. 69, n. 3, p. 809–820, 2021. Citado na página [22](#).

LUKAS, M. A. On the discrepancy principle and generalised maximum likelihood for regularisation. **Bulletin of the Australian Mathematical Society**, Cambridge University Press, v. 52, n. 3, p. 399–424, 1995. Citado na página [42](#).

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943. Citado na página [45](#).

MEISE, R.; VOGT, D. **Introduction to functional analysis**. Oxford: Clarendon Press, 1997. Citado na página [81](#).

MOROZOV, V. A. On the solution of functional equations by the method of regularization. In: RUSSIAN ACADEMY OF SCIENCES. **Doklady Akademii Nauk**. USSR, 1966. v. 167, n. 3, p. 510–512. Citado na página [41](#).

NATTERER, F. **The mathematics of computerized tomograph**. Stuttgart, New York: SIAM, 2001. Citado na página [27](#).

NATTERER, F.; WÜBBELING, F. **Mathematical methods in image reconstruction**. Philadelphia: SIAM, 2001. Citado na página [21](#).

NETO, E. S. H. Algoritmos incrementais com aplicações em tomografia computadorizada. **Doutorado, Universidade Estadual de Campinas**, 2009. Citado na página [27](#).

NISHINO, R.; LOOMIS, S. H. C. Cupy: A numpy-compatible library for nvidia gpu calculations. **31st conference on neural information processing systems**, v. 151, n. 7, 2017. Citado na página [56](#).

OSHER, S.; BURGER, M.; GOLDFARB, D.; XU, J.; YIN, W. An iterative regularization method for total variation-based image restoration. **Multiscale Modeling & Simulation**, SIAM, v. 4, n. 2, p. 460–489, 2005. Citado na página [21](#).

PHILLIPS, D. L. A technique for the numerical solution of certain integral equations of the first kind. **Journal of the ACM (JACM)**, ACM New York, NY, USA, v. 9, n. 1, p. 84–97, 1962. Citado nas páginas [22](#), [41](#) e [51](#).

RADON, J. On the determination of functions from their integrals along certain manifolds. **Mathematisch-Physische Klasse**, v. 69, p. 262–277, 1917. Citado na página [26](#).

RECKTENWALD, G. W. Finite-difference approximations to the heat equation. **Mechanical Engineering**, v. 10, n. 01, 2004. Citado na página [25](#).

REGIŃSKA, T. A regularization parameter in discrete ill-posed problems. **SIAM Journal on Scientific Computing**, SIAM, v. 17, n. 3, p. 740–749, 1996. Citado na página [43](#).

- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958. Citado nas páginas 45 e 46.
- ROTH, S.; BLACK, M. J. Fields of experts. **International Journal of Computer Vision**, Springer, v. 82, p. 205–229, 2009. Citado na página 55.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986. Citado na página 46.
- SARKER, I. H. Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective. **SN Computer Science**, Springer, v. 2, n. 3, p. 154, 2021. Citado na página 47.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural Networks**, Elsevier, v. 61, p. 85–117, 2015. Citado na página 45.
- SEIDMAN, T. . I. Nonconvergence results for the application of least-squares estimation to ill-posed problems. **Journal of Optimization Theory and Applications**, Springer, v. 30, n. 4, p. 535–547, 1980. Citado na página 37.
- SIDDON, R. L. Fast calculation of the exact radiological path for a three-dimensional ct array. **Medical physics**, Wiley Online Library, v. 12, n. 2, p. 252–255, 1985. Citado na página 27.
- TIKHONOV, A. N. Solution of incorrectly formulated problems and the regularization method. **Soviet Math.**, v. 4, p. 1035–1038, 1963. Citado na página 39.
- TIKHONOV, A. N.; ARSENIN, V. Y. **Solutions of Ill-Posed Problems**. Washington; New York: Winston, 1977. Citado nas páginas 21 e 39.
- VIRTANEN, P.; GOMMERS, R.; OLIPHANT, T. E.; HABERLAND, M.; REDDY, T.; COURNAPEAU, D.; BUROVSKI, E.; PETERSON, P.; WECKESSER, W.; BRIGHT, J. *et al.* Scipy 1.0: fundamental algorithms for scientific computing in python. **Nature methods**, Nature Publishing Group, v. 17, n. 3, p. 261–272, 2020. Citado na página 52.
- VOGEL, C. R. **Computational methods for inverse problems**. Philadelphia: SIAM, 2002. Citado nas páginas 26 e 42.
- WANG, Z.; BOVIK, A. C.; SHEIKH, H. R.; SIMONCELLI, E. P. Image quality assessment: from error visibility to structural similarity. **IEEE transactions on image processing**, IEEE, v. 13, n. 4, p. 600–612, 2004. Citado na página 55.
- WIDROW, B.; HOFF, M. E. Associative storage and retrieval of digital information in networks of adaptive “neurons”. In: **Biological Prototypes and Synthetic Systems**. Ithaca, New York: Springer, 1962. p. 160–160. Citado na página 45.

FUNDAMENTOS TEÓRICOS

Neste apêndice, são desenvolvidas todas as definições, fatos e notações que foram utilizados durante todo este trabalho.

A.1 Espaços vetoriais normados

Denotaremos por X e Y dois espaços vetoriais sobre um mesmo corpo \mathbb{K} , que pode ser tanto \mathbb{R} ou \mathbb{C} .

Definição 4 (Norma). Uma função $\|\cdot\|_X : X \rightarrow \mathbb{R}^+$ é chamada norma se para todo $\mathbf{x} \in X$ vale:

1. $\|\lambda \mathbf{x}\|_X = |\lambda| \|\mathbf{x}\|_X, \quad \forall \lambda \in \mathbb{K};$
2. $\|\mathbf{x} + \mathbf{y}\|_X \leq \|\mathbf{x}\|_X + \|\mathbf{y}\|_X, \quad \forall \mathbf{y}, \mathbf{x} \in X;$
3. $\|\mathbf{x}\|_X = 0 \Leftrightarrow \mathbf{x} = 0.$

Exemplo 1. Uma função $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}^+$, onde para qualquer vetor $\mathbf{x} \in \mathbb{R}^n$

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}},$$

para qualquer $p \in [1, \infty)$ é uma norma. Esta norma é conhecido por norma- p .

Demonstração. Mostraremos a validade de cada uma das condições.

1. Seja qualquer $\mathbf{x} \in \mathbb{R}^n$, para $\lambda \in \mathbb{K}$ e $p \in [1, \infty)$ temos que

$$\begin{aligned} \|\lambda \mathbf{x}\|_p &= \left(\sum_{k=1}^n |\lambda x_k|^p \right)^{\frac{1}{p}} \\ &= \left(\sum_{k=1}^n |\lambda|^p |x_k|^p \right)^{\frac{1}{p}} \\ &= \left(|\lambda|^p \sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \\ &= |\lambda| \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \\ &= |\lambda| \|\mathbf{x}\|_p \end{aligned}$$

2. Sejam quaisquer \mathbf{x} e $\mathbf{y} \in \mathbb{R}^n$, assim

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &= \sum_{k=1}^n |x_k + y_k|^p \\ &= \sum_{k=1}^n |x_k + y_k| |x_k + y_k|^{p-1} \\ &\leq \sum_{k=1}^n (|x_k| + |y_k|) |x_k + y_k|^{p-1} \\ &= \sum_{k=1}^n |x_k| |x_k + y_k|^{p-1} + \sum_{k=1}^n |y_k| |x_k + y_k|^{p-1} \\ &= \sum_{k=1}^n |x_k (x_k + y_k)^{p-1}| + \sum_{k=1}^n |y_k (x_k + y_k)^{p-1}|. \end{aligned}$$

Para o passo seguinte utilizaremos a desigualdade de Hölder (KREYSZIG, 1991). Sejam quaisquer $q > 1$ e $r > 1$, tais que $\frac{1}{r} + \frac{1}{q} = 1$, então para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, vale que

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^r \right)^{\frac{1}{r}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}.$$

Aplicando a desigualdade de Hölder e tomando em particular $r = p$ temos que

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &\leq \sum_{k=1}^n |x_k (x_k + y_k)^{p-1}| + \sum_{k=1}^n |y_k (x_k + y_k)^{p-1}| \\ &\leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n (|x_k + y_k|^{p-1})^q \right)^{\frac{1}{q}} \\ &\quad + \left(\sum_{k=1}^n |y_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n (|x_k + y_k|^{p-1})^q \right)^{\frac{1}{q}} \\ &= \left(\sum_{k=1}^n (|x_k + y_k|^{p-1})^q \right)^{\frac{1}{q}} \left(\left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p \right)^{\frac{1}{p}} \right). \end{aligned}$$

Sabendo que $\frac{1}{p} + \frac{1}{q} = 1$, isso implica que $pq - q = p$. Substituindo na primeira parcela da última igualdade acima temos

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &\leq \left(\sum_{k=1}^n |x_k + y_k|^p \right)^{\frac{1}{q}} \left(\left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p \right)^{\frac{1}{p}} \right) \\ &= \|\mathbf{x} + \mathbf{y}\|_p^{\frac{p}{q}} (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p). \end{aligned}$$

Seguindo, temos

$$\|\mathbf{x} + \mathbf{y}\|_p^{p - \frac{p}{q}} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p,$$

por outro lado $p - \frac{p}{q} = p \left(1 - \frac{1}{q} \right) = 1$. Portanto,

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

3. Seja qualquer $\mathbf{x} \in \mathbb{R}^n$, se temos que

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} = 0,$$

como a soma é finita e os valores são todos não negativos, a única possibilidade de a soma ser nula é se $\mathbf{x} = \mathbf{0}$.

Com todas as exigências foram cumpridas, a função definida acima é de fato uma norma em \mathbb{R}^n . \square

Dizemos que duas normas $\|\cdot\|_A$ e $\|\cdot\|_B$ sobre o mesmo espaço X são equivalentes se existirem $c_1, c_2 > 0$ tais que

$$c_1 \|\mathbf{x}\|_B \leq \|\mathbf{x}\|_A \leq c_2 \|\mathbf{x}\|_B, \quad \forall \mathbf{x} \in X.$$

Se o espaço vetorial X for de dimensão finita, todas as normas são equivalentes (KREYSZIG, 1991). Outro resultado importante é que duas normas equivalentes definem uma mesma topologia para X .

Definição 5 (Espaço normado). Um espaço normado X é um espaço vetorial munido de uma norma.

Exemplo 2. O espaço vetorial \mathbb{R}^n suprido da norma- p para qualquer $p \in [1, \infty)$ é um espaço normado.

A noção de norma nos leva diretamente à noção de convergência. Dizemos que uma sequência $\{\mathbf{x}_n\}_{n \in \mathbb{N}} \subset X$ converge para um $\mathbf{x} \in X$ se $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}$ tal que

$$\|\mathbf{x}_n - \mathbf{x}\|_X < \varepsilon, \quad \text{para } n > n_0.$$

Outra maneira de denotar isto é escrever $\mathbf{x}_n \rightarrow \mathbf{x}$, quando $n \rightarrow \infty$. Além disso, dizemos que esta tal sequência é de Cauchy se $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}$ tal que

$$\|\mathbf{x}_m - \mathbf{x}_n\|_X < \varepsilon, \quad \text{para } m, n > n_0.$$

Exemplo 3. A sequência $\{\frac{1}{n}\}_{n \in \mathbb{N}^*}$ é de Cauchy.

Demonstração. Seja $\varepsilon > 0$, tomamos $N > \frac{2}{\varepsilon}$ e $m, n > N$, assim

$$\begin{aligned} \left| \frac{1}{m} - \frac{1}{n} \right| &\leq \left| \frac{1}{m} \right| + \left| \frac{1}{n} \right| = \frac{1}{m} + \frac{1}{n} < \frac{1}{N} + \frac{1}{N} \\ &= \frac{2}{N} \\ &< \varepsilon. \end{aligned}$$

Portanto, a sequência $\{\frac{1}{n}\}_{n \in \mathbb{N}^*}$ é de Cauchy □

Definição 6 (Espaço de Banach). Um espaço vetorial normado X é chamado espaço de Banach quando toda sequência de Cauchy em X é convergente com respeito à norma $\|\cdot\|_X$. O espaço de Banach também é conhecido como espaço vetorial normado completo.

Seja $U \subset X$, este conjunto é chamado de:

- **fechado** quando o limite de toda sequência convergente $\{\mathbf{x}_n\}_{n \in \mathbb{N}} \subset U$ está em U .
- **compacto** quando toda sequência convergente $\{\mathbf{x}_n\}_{n \in \mathbb{N}}$ contém uma subsequência convergente $\{\mathbf{x}_{n_k}\}_{k \in \mathbb{N}}$ que converge para um elemento em U .
- **denso** em X se para todo $\mathbf{x} \in X$ existe uma sequência $\{\mathbf{x}_n\}_{n \in \mathbb{N}} \subset U$ com $\mathbf{x}_n \rightarrow \mathbf{x}$.

A união de U com o conjunto de todos os limites de sequências convergentes em U é chamada fecho de U , que denotamos por \bar{U} .

Definição 7 (Produto interno). Uma função $\langle \cdot, \cdot \rangle_X : X \times X \rightarrow \mathbb{C}$ é chamada produto interno se vale:

1. $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle_X = \bar{\alpha} \langle \mathbf{x}, \mathbf{z} \rangle_X + \bar{\beta} \langle \mathbf{y}, \mathbf{z} \rangle_X, \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in X \text{ e } \alpha, \beta \in \mathbb{C};$
2. $\langle \mathbf{x}, \mathbf{y} \rangle_X = \overline{\langle \mathbf{y}, \mathbf{x} \rangle_X}, \quad \forall \mathbf{x}, \mathbf{y} \in X;$
3. $\langle \mathbf{x}, \mathbf{x} \rangle_X \geq 0, \quad \forall \mathbf{x} \in X;$
4. $\langle \mathbf{x}, \mathbf{x} \rangle_X = 0 \Leftrightarrow \mathbf{x} = 0,$

onde \bar{u} denota o complexo conjugado de $u \in \mathbb{C}$.

O produto interno pode ser entendido como uma forma de ver uma similaridade entre dois elementos do mesmo espaço X . Uma importante desigualdade proveniente do produto interno é a desigualdade de Cauchy-Schwarz. Explicitamente, sejam quaisquer dois vetores $\mathbf{x}, \mathbf{y} \in X$, a desigualdade de Cauchy-Schwarz afirma que

$$|\langle \mathbf{x}, \mathbf{y} \rangle_X| \leq \|\mathbf{x}\|_X \|\mathbf{y}\|_X.$$

Outra importante afirmação é quando $\langle \mathbf{x}, \mathbf{y} \rangle_X = 0$, onde dizemos que \mathbf{x} e \mathbf{y} são ortogonais entre si. Com isto, seja um conjunto qualquer $U \subset X$, definimos como o **complemento ortogonal** de U em X o conjunto:

$$U^\perp := \{\mathbf{x} \in X \mid \langle \mathbf{x}, \mathbf{u} \rangle_X = 0, \forall \mathbf{u} \in U\}.$$

Temos que U^\perp é um subespaço de X , além de que a própria definição implica que o mesmo é fechado. Em particular, o complemento ortogonal de X só irá conter o vetor nulo do espaço. Um fato interessante é que desde que temos um produto interno de um espaço X , podemos definir uma norma a partir dele.

Exemplo 4. Uma função $\|\cdot\|_X : X \rightarrow \mathbb{K}$ definida por

$$\|\mathbf{x}\|_X := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_X}$$

é uma norma.

Demonstração. Sejam quaisquer $\mathbf{x}, \mathbf{y} \in X$ e $\alpha \in \mathbb{K}$. Então,

1.

$$\|\alpha \mathbf{x}\|_X^2 = \langle \alpha \mathbf{x}, \alpha \mathbf{x} \rangle_X = \overline{\alpha} \overline{\langle \alpha \mathbf{x}, \mathbf{x} \rangle_X} = \overline{\alpha} \alpha \overline{\langle \mathbf{x}, \mathbf{x} \rangle_X} = \alpha^2 \langle \mathbf{x}, \mathbf{x} \rangle_X = \alpha^2 \|\mathbf{x}\|_X^2;$$

2.

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_X^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle_X \\ &= \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle_X + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle_X \\ &= \overline{\langle \mathbf{x} + \mathbf{y}, \mathbf{x} \rangle_X} + \overline{\langle \mathbf{x} + \mathbf{y}, \mathbf{y} \rangle_X} \\ &= \overline{\langle \mathbf{x}, \mathbf{x} \rangle_X} + \overline{\langle \mathbf{y}, \mathbf{x} \rangle_X} + \overline{\langle \mathbf{x}, \mathbf{y} \rangle_X} + \overline{\langle \mathbf{y}, \mathbf{y} \rangle_X} \\ &= \overline{\langle \mathbf{x}, \mathbf{x} \rangle_X} + 2\operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle_X) + \overline{\langle \mathbf{y}, \mathbf{y} \rangle_X} \\ &\leq \overline{\langle \mathbf{x}, \mathbf{x} \rangle_X} + 2|\langle \mathbf{x}, \mathbf{y} \rangle_X| + \overline{\langle \mathbf{y}, \mathbf{y} \rangle_X} \\ &\leq \overline{\langle \mathbf{x}, \mathbf{x} \rangle_X} + 2\|\mathbf{x}\|_X \|\mathbf{y}\|_X + \overline{\langle \mathbf{y}, \mathbf{y} \rangle_X} \\ &= \|\mathbf{x}\|_X^2 + 2\|\mathbf{x}\|_X \|\mathbf{y}\|_X + \|\mathbf{y}\|_X^2 \\ &= (\|\mathbf{x}\|_X + \|\mathbf{y}\|_X)^2; \end{aligned}$$

3.

$$\|\mathbf{x}\|_X^2 = \langle \mathbf{x}, \mathbf{x} \rangle_X = 0 \Rightarrow \mathbf{x} = 0.$$

□

Definição 8 (Espaço de Hilbert). Um espaço de Banach X é chamado espaço de Hilbert quando a norma $\|\cdot\|_X$ é definida pelo produto interno, ou seja, $\|\cdot\|_X^2 = \langle \cdot, \cdot \rangle_X$.

De agora em diante, quando nos referirmos a um espaço de Hilbert, denotaremos por uma letra caligráfica, por exemplo, o espaço de Hilbert \mathcal{X} . Com a definição de espaços de Hilbert, conseguimos tratar de espaços que antes não se podia pensar instantaneamente como um espaço vetorial. Por exemplo, $L^2[0, 1]$, o espaço das funções quadrado integráveis em $[0, 1]$, o qual é um espaço de Hilbert. Com isto, todos os teoremas provados para espaços de Hilbert abstratos podem ser aplicado a qualquer espaço de Hilbert concreto que se esteja trabalhando.

Teorema 8. Sejam \mathcal{X} um espaço de Hilbert e um subespaço $U \subset \mathcal{X}$ fechado. Então $\mathcal{X} = U \oplus U^\perp$, isto é, todo elemento $x \in \mathcal{X}$ pode ser escrito de maneira única como

$$x = u + u', \quad \text{com } u \in U, u' \in U^\perp.$$

Demonstração. Ver (FOLLAND, 1999, Teorema 5.24). □

A.2 Operadores Lineares

Definição 9. Sejam X e Y espaços vetoriais sobre um mesmo corpo \mathbb{K} e um conjunto $U \subset X$. Uma função $T : U \subset X \rightarrow Y$ é um operador linear se para todo $x, z \in U$ e $\alpha \in \mathbb{K}$, temos

$$T(\alpha x + z) = \alpha T x + T z.$$

Com a definição acima, usaremos a seguinte notação:

- $D(T) := U$ é o domínio de T ;
- $N(T) := \{x \in D(T) \mid T x = \mathbf{0} \text{ (vetor nulo do espaço)}\}$ é o núcleo de T ;
- $R(T) := \{T x \in Y \mid x \in D(T)\}$ é o conjunto imagem de T .

Sendo assim, um operador linear T é dito

1. Injetivo se $N(T) = \{\mathbf{0}\}$;
2. Sobrejetivo se $R(T) = Y$;
3. Bijetivo se T é injetivo e sobrejetivo.

No caso que o operador T é bijetivo, existe um operador $T^{-1} : R(T) \rightarrow D(T)$, onde para cada $\mathbf{y} \in R(T)$ conseguimos um $\mathbf{x} \in D(T)$ tal que $T^{-1}\mathbf{y} = \mathbf{x}$. Este operador é conhecido como o operador inverso de T . Com isto, temos que

$$\begin{aligned} T^{-1}T\mathbf{x} &= \mathbf{x}, \quad \forall \mathbf{x} \in D(T); \\ TT^{-1}\mathbf{y} &= \mathbf{y}, \quad \forall \mathbf{y} \in R(T). \end{aligned}$$

Se para um operador linear T qualquer, o operador inverso existir, o mesmo é linear (KREYSZIG, 1991). Caso T seja definida entre dois espaços vetoriais X e Y , ambos de dimensão finita, é possível representar T como uma matriz em relação às bases de X e Y . Suponha que $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ seja uma base de X e $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ uma base de Y . Podemos escrever $T\mathbf{x}_i$ como uma combinação linear dos vetores da base de Y , ou seja,

$$T\mathbf{x}_j = a_{1j}\mathbf{y}_1 + a_{2j}\mathbf{y}_2 + \dots + a_{mj}\mathbf{y}_m,$$

onde a_{ij} é o i -ésimo componente do vetor $T\mathbf{x}_j$ em relação à base de Y . Portanto, a matriz que representa o operador T em relação às bases escolhidas é uma matriz de dimensão $m \times n$ onde o elemento (i, j) da matriz é a_{ij} .

Exemplo 5. O operador identidade $I : X \rightarrow X$, definido como $I\mathbf{x} = \mathbf{x}$, para todo $\mathbf{x} \in X$ é linear.

Definição 10. Seja um operador linear $T : D(T) \subset X \rightarrow Y$. A restrição do operador T a um subconjunto $U \subset D(T)$ é denotada por $T|_U$, onde $T|_U : U \rightarrow Y$ é tal que $T|_U\mathbf{x} = T\mathbf{x}$, para todo $\mathbf{x} \in U$.

Definição 11. Seja um operador linear $T : D(T) \subset X \rightarrow Y$. Uma extensão do operador T a um subconjunto $U \supset D(T)$ é denotada por \tilde{T} , onde $\tilde{T} : U \rightarrow Y$ é tal que $\tilde{T}|_{D(T)} = T$, ou seja, $\tilde{T}\mathbf{x} = T\mathbf{x}$ para qualquer $\mathbf{x} \in D(T)$.

Definição 12. Sejam X e Y espaços normados e um operador linear $T : X \rightarrow Y$. Dizemos que T é limitado se existir um $c > 0$ tal que

$$\|T\mathbf{x}\|_Y \leq c\|\mathbf{x}\|_X, \quad \forall \mathbf{x} \in D(T).$$

Definição 13. O espaço $\mathcal{L}(X, Y)$ é o espaço de todos os operadores lineares limitados de X para Y .

Definição 14. Sejam X e Y espaços normados e um operador linear $T : X \rightarrow Y$. Dizemos que T é contínuo em $\mathbf{x}_0 \in X$ quando, para todo $\varepsilon > 0$, pode-se obter um $\delta > 0$ tal que

$$\|T\mathbf{x}_0 - T\mathbf{z}\|_Y \leq \varepsilon, \quad \forall \mathbf{z} \in X, \text{ com } \|\mathbf{x}_0 - \mathbf{z}\|_X \leq \delta.$$

Quanto T é contínuo em todo $\mathbf{x}_0 \in X$, dizemos que T é contínuo.

Proposição 2. Sejam X e Y espaços normados e um operador linear $T : X \rightarrow Y$, as seguintes propriedades são equivalentes:

- i) T é limitado;
- ii) T é contínuo;
- iii) T é contínuo em 0.

Demonstração. i) \Rightarrow ii)

Seja T limitado, assim existe $c > 0$ tal que $\forall \mathbf{x} \in X$,

$$\|T\mathbf{x}\|_Y \leq c\|\mathbf{x}\|_X.$$

Pela linearidade de T , temos que para quaisquer $\mathbf{x}_1, \mathbf{x}_2 \in X$,

$$\|T\mathbf{x}_1 - T\mathbf{x}_2\|_Y = \|T(\mathbf{x}_1 - \mathbf{x}_2)\|_Y \leq c\|\mathbf{x}_1 - \mathbf{x}_2\|_X.$$

Dado um $\varepsilon > 0$ qualquer, fazemos $\delta = \frac{\varepsilon}{c}$, assim $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ tais que $\|\mathbf{x}_1 - \mathbf{x}_2\|_X < \delta$, teremos

$$\|T\mathbf{x}_1 - T\mathbf{x}_2\|_Y \leq c\|\mathbf{x}_1 - \mathbf{x}_2\|_X < c\delta = \varepsilon.$$

Portanto, T é uniformemente contínuo.

ii) \Rightarrow iii)

Trivial.

iii) \Rightarrow i)

Temos que T é contínuo em 0. Então, tomemos $\varepsilon = 1$, assim $\exists \delta > 0$, tal que

$$\|T\mathbf{x}\|_Y \leq 1, \quad \forall \mathbf{x} \in X \text{ com } \|\mathbf{x}\|_X \leq \delta.$$

Tome $\mathbf{z} \in X$, com $\|\mathbf{z}\|_X \leq 1$. Com isto, $\|\frac{\delta}{2}\mathbf{z}\|_X \leq \delta$, assim

$$\left\| T\left(\frac{\delta}{2}\mathbf{z}\right) \right\|_Y = \frac{\delta}{2} \|T\mathbf{z}\|_Y \leq 1 \Rightarrow \|T\mathbf{z}\|_Y \leq \frac{2}{\delta}.$$

Façamos $\mathbf{z} = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|_X}$, para qualquer $\mathbf{x}_1 \in X$, $\mathbf{x}_1 \neq 0$, assim $\|\mathbf{z}\|_X = 1$. Para o caso em que $\mathbf{x}_1 = 0$, tomamos $\mathbf{z} = 0$ e assim, $\|\mathbf{z}\|_X = 0$. Dado que $\|\mathbf{z}\|_X \leq 1$, podemos concluir, pela linearidade de T que

$$\|T\mathbf{z}\|_Y = \left\| T\left(\frac{\mathbf{x}_1}{\|\mathbf{x}_1\|_X}\right) \right\|_Y = \left\| \frac{1}{\|\mathbf{x}_1\|_X} T\mathbf{x}_1 \right\|_Y = \frac{1}{\|\mathbf{x}_1\|_X} \|T\mathbf{x}_1\|_Y \leq \frac{2}{\delta}$$

Tomando $c = \frac{2}{\delta}$, temos que $\|T\mathbf{x}_1\|_Y \leq c\|\mathbf{x}_1\|_X$, para todo $\mathbf{x}_1 \in X$. □

Definição 15. Sejam X e Y espaços normados e um operador linear $T : D(T) \rightarrow Y$. Dizemos que T é fechado se para cada sequência $\{\mathbf{x}_n\}_{n=1}^{\infty} \subset D(T)$ tal que $\mathbf{x}_n \rightarrow \mathbf{x}$ e $T\mathbf{x}_n \rightarrow \mathbf{y}$, então $\mathbf{x} \in D(T)$ e $T\mathbf{x} = \mathbf{y}$.

Definição 16. Sejam X e Y espaços normados e um operador linear $T : D(T) \rightarrow Y$. Dizemos que T é compacto se para cada sequência limitada $\{x_n\}_{n=1}^{\infty}$ existe uma subsequência $\{x_{n_k}\}_{k=1}^{\infty}$ tal que Tx_{n_k} é convergente.

Definição 17. O espaço $\mathcal{K}(X, Y)$ é o espaço de todos os operadores lineares compactos de X para Y .

Teorema 9. Sejam X e Y espaços de Banach. Se $T \in \mathcal{L}(X, Y)$ é bijetora, então T^{-1} é contínua.

Demonstração. Ver [Kreyszig \(1991\)](#), Teorema 4.12-2 □

Teorema 10 (Gráfico fechado). Sejam X e Y espaços de Banach e um operador linear $T : D(T) \rightarrow Y$. Se o gráfico de T

$$gr(T) := \{(x, y) \mid x \in D(T), y = Tx\}$$

é fechado em $X \times Y$, então T é contínua.

Demonstração. Ver ([KREYSZIG, 1991](#), Teorema 4.13-2) □

Definição 18. Sejam \mathcal{X} , \mathcal{Y} espaços de Hilbert e $T : \mathcal{X} \rightarrow \mathcal{Y}$ um operador linear. Temos que um operador $T^* : \mathcal{Y} \rightarrow \mathcal{X}$ é dito adjunto de T se para todo $x \in \mathcal{X}$ e $y \in \mathcal{Y}$ temos que

$$\langle y, Tx \rangle_{\mathcal{Y}} = \langle T^*y, x \rangle_{\mathcal{X}}.$$

Se $T^* = T$, dizemos que T é auto-adjunto.

Seguindo a notação do Teorema 8, podemos definir de acordo com [Kreyszig \(1991\)](#) o operador projetor ortogonal P sobre um subespaço U fechado.

Definição 19. Seja $\mathcal{X} = U \oplus U^{\perp}$, onde $x = u + u'$ com $u \in U$ e $u' \in U^{\perp}$. Com isto definimos o mapeamento

$$P_U : \mathcal{X} \rightarrow U$$

$$x \mapsto u.$$

P_U é chamado de projetor ortogonal sobre o subespaço fechado U .

Proposição 3. O projetor ortogonal P_U satisfaz as seguintes propriedades:

1. P_U é linear;
2. P_U é auto-adjunto;
3. $\|P_U\|_{\mathcal{L}(\mathcal{X}, \mathcal{X})} = 1$;
4. $P_U \circ P_U = P_U$.

Demonstração. 1. Sejam $\mathbf{x}_1 = \mathbf{u}_1 + \mathbf{u}'_1$ e $\mathbf{x}_2 = \mathbf{u}_2 + \mathbf{u}'_2 \in \mathcal{X}$, com $\mathbf{u}_1, \mathbf{u}_2 \in U$ e $\mathbf{u}'_1, \mathbf{u}'_2 \in U^\perp$.
Dados $\alpha_1, \alpha_2 \in \mathbb{K}$ temos que,

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 = \alpha_1(\mathbf{u}_1 + \mathbf{u}'_1) + \alpha_2(\mathbf{u}_2 + \mathbf{u}'_2) = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_1 \mathbf{u}'_1 + \alpha_2 \mathbf{u}'_2.$$

Deste modo,

$$P_U(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 = \alpha_1 P_U \mathbf{x}_1 + \alpha_2 P_U \mathbf{x}_2.$$

Portanto, P_U é linear.

2. Sejam $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ e suas respectivas projeções $P_U \mathbf{x}_1 = \mathbf{u}_1$ e $P_U \mathbf{x}_2 = \mathbf{u}_2$. Temos que,

$$\langle P_U \mathbf{x}_1, \mathbf{x}_2 \rangle_{\mathcal{X}} = \langle \mathbf{u}_1, \mathbf{x}_2 \rangle_{\mathcal{X}} = \langle \mathbf{u}_1, \mathbf{u}_2 + \mathbf{u}'_2 \rangle_{\mathcal{X}} = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle_{\mathcal{X}}.$$

Por outro lado,

$$\langle \mathbf{x}_1, P_U \mathbf{x}_2 \rangle_{\mathcal{X}} = \langle \mathbf{x}_1, \mathbf{u}_2 \rangle_{\mathcal{X}} = \langle \mathbf{u}_1 + \mathbf{u}'_1, \mathbf{u}_2 \rangle_{\mathcal{X}} = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle_{\mathcal{X}}.$$

Portanto,

$$\langle P_U \mathbf{x}_1, \mathbf{x}_2 \rangle_{\mathcal{X}} = \langle \mathbf{x}_1, P_U \mathbf{x}_2 \rangle_{\mathcal{X}}$$

e, assim P_U é auto-adjunto.

3. Seja $\mathbf{x} = \mathbf{u} + \mathbf{u}' \in \mathcal{X}$ com $\mathbf{u} \in U$ e $\mathbf{u}' \in U^\perp$, assim sendo

$$\begin{aligned} \|\mathbf{x}\|_{\mathcal{X}}^2 &= \|\mathbf{u} + \mathbf{u}'\|_{\mathcal{X}}^2 = \langle \mathbf{u} + \mathbf{u}', \mathbf{u} + \mathbf{u}' \rangle_{\mathcal{X}} = \langle \mathbf{u}, \mathbf{u} + \mathbf{u}' \rangle_{\mathcal{X}} + \langle \mathbf{u}', \mathbf{u} + \mathbf{u}' \rangle_{\mathcal{X}} \\ &= \overline{\langle \mathbf{u} + \mathbf{u}', \mathbf{u} \rangle_{\mathcal{X}}} + \overline{\langle \mathbf{u} + \mathbf{u}', \mathbf{u}' \rangle_{\mathcal{X}}} \\ &= \overline{\langle \mathbf{u}, \mathbf{u} \rangle_{\mathcal{X}}} + \overline{\langle \mathbf{u}', \mathbf{u} \rangle_{\mathcal{X}}} + \overline{\langle \mathbf{u}, \mathbf{u}' \rangle_{\mathcal{X}}} + \overline{\langle \mathbf{u}', \mathbf{u}' \rangle_{\mathcal{X}}} \\ &= \|\mathbf{u}\|_{\mathcal{X}}^2 + \|\mathbf{u}'\|_{\mathcal{X}}^2. \end{aligned}$$

Com isto, temos que $\|\mathbf{u}\|_{\mathcal{X}}^2 \leq \|\mathbf{x}\|_{\mathcal{X}}^2$, logo $\|P_U \mathbf{x}\|_{\mathcal{X}}^2 \leq \|\mathbf{x}\|_{\mathcal{X}}^2$ para todo $\mathbf{x} \in X$, então $\|P_U\|_{\mathcal{L}(\mathcal{X}, \mathcal{X})} \leq 1$. Como $\forall \mathbf{x} \in U, P_U \mathbf{x} = \mathbf{x}$, assim $\|P_U \mathbf{x}\|_{\mathcal{X}} = \|\mathbf{x}\|_{\mathcal{X}}$, isto implica que $\|P_U\|_{\mathcal{L}(\mathcal{X}, \mathcal{X})} = 1$.

4. Para todo $\mathbf{x} \in \mathcal{X}$, seja $P_U \mathbf{x} \in U$. Assim,

$$(P_U \circ P_U) \mathbf{x} = P_U(P_U \mathbf{x}) = P_U \mathbf{x}, \quad \forall \mathbf{x} \in \mathcal{X}.$$

Portanto, $P_U \circ P_U = P_U$

□

Caso o subespaço U não seja fechado, somente temos que $(U^\perp)^\perp = \overline{U}$. Sendo assim, para qualquer $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ temos que

Proposição 4. Para qualquer operador $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, temos as seguintes propriedades:

1. $R(T)^\perp = N(T^*)$;
2. $R(T^*)^\perp = N(T)$.

Demonstração. 1.

$$\begin{aligned}
\mathbf{y} \in N(T^*) &\Leftrightarrow T^*\mathbf{y} = 0 \\
&\Leftrightarrow \langle T^*\mathbf{y}, \mathbf{x} \rangle_{\mathcal{X}} = 0, \quad \forall \mathbf{x} \in \mathcal{X} \\
&\Leftrightarrow \langle \mathbf{y}, T\mathbf{x} \rangle_{\mathcal{Y}} = 0, \quad \forall \mathbf{x} \in \mathcal{X} \\
&\Leftrightarrow \mathbf{y} \in R(T)^\perp.
\end{aligned}$$

Portanto, $R(T)^\perp = N(T^*)$ e por consequência, $(R(T)^\perp)^\perp = N(T^*)^\perp \Rightarrow N(T^*)^\perp = \overline{R(T)}$.

2.

$$\begin{aligned}
\mathbf{x} \in N(T) &\Leftrightarrow T\mathbf{x} = 0 \\
&\Leftrightarrow \langle T\mathbf{x}, \mathbf{y} \rangle_{\mathcal{Y}} = 0, \quad \forall \mathbf{y} \in \mathcal{Y} \\
&\Leftrightarrow \langle \mathbf{x}, T^*\mathbf{y} \rangle_{\mathcal{X}} = 0, \quad \forall \mathbf{y} \in \mathcal{Y} \\
&\Leftrightarrow \mathbf{x} \in R(T^*)^\perp
\end{aligned}$$

Portanto, $R(T^*)^\perp = N(T)$ e por consequência, $(R(T^*)^\perp)^\perp = N(T)^\perp \Rightarrow \overline{R(T^*)} = N(T)^\perp$.

□

Teorema 11. Seja \mathcal{X} um espaço de Hilbert e um operador linear compacto autoadjunto $T \in \mathcal{L}(\mathcal{X}, \mathcal{X})$. Então, existe uma sequência $\{\lambda_n\}_{n \in \mathbb{N}}$ de autovalores e uma sequência $\{\mathbf{v}_n\}_{n \in \mathbb{N}}$ ortonormal de autovetores, tais que

$$T\mathbf{x} = \sum_{n \in \mathbb{N}} \lambda_n \langle \mathbf{x}, \mathbf{v}_n \rangle_{\mathcal{X}} \mathbf{v}_n,$$

para todo $\mathbf{x} \in \mathcal{X}$. Se a sequência $\{\lambda_n\}$ for infinita, então $\lambda_n \rightarrow 0$.

Demonstração. Ver (MEISE; VOGT, 1997, proposição 16.2)

□

