

Volume 59, 2012

Editores

Cassio Machiaveli Oishi

Universidade Estadual Paulista - UNESP
Presidente Prudente, SP, Brasil

Fernando Rodrigo Rafaeli

Universidade Estadual Paulista - UNESP
São José do Rio Preto, SP, Brasil

Rosana Sueli da Motta Jafelice (Editor Chefe)

Universidade Federal de Uberlândia - UFU
Uberlândia, MG, Brasil

Rubens de Figueiredo Camargo

Universidade Estadual Paulista - UNESP
Bauru, SP, Brasil

Sezimária de Fátima P. Saramago

Universidade Federal de Uberlândia - UFU
Uberlândia, MG, Brasil

Vanessa Avansini Botta Pirani (Editor Adjunto)

Universidade Estadual Paulista - UNESP
Presidente Prudente, SP, Brasil



A Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC publica, desde as primeiras edições do evento, monografias dos cursos que são ministrados nos CNMAC.

Para a comemoração dos 25 anos da SBMAC, que ocorreu durante o XXVI CNMAC em 2003, foi criada a série **Notas em Matemática Aplicada** para publicar as monografias dos minicursos ministrados nos CNMAC, o que permaneceu até o XXXIII CNMAC em 2010.

A partir de 2011, a série passa a publicar, também, livros nas áreas de interesse da SBMAC. Os autores que submeterem textos à série Notas em Matemática Aplicada devem estar cientes de que poderão ser convidados a ministrarem minicursos nos eventos patrocinados pela SBMAC, em especial nos CNMAC, sobre assunto a que se refere o texto.

O livro deve ser preparado em **Latex (compatível com o Miktex versão 2.7)**, as **figuras em eps** e deve ter entre **80 e 150 páginas**. O texto deve ser redigido de forma clara, acompanhado de uma excelente revisão bibliográfica e de **exercícios de verificação de aprendizagem** ao final de cada capítulo.

Veja todos os títulos publicados nesta série na página
<http://www.sbmac.org.br/notas.php>

PROBLEMAS INVERSOS EM TOMOGRAFIAS

Adriano De Cezaro
adrianocezaro@furg.br
Fabiana Travessini De Cezaro
fabianacezaro@furg.br

Instituto de Matemática Estatística e Física
Universidade Federal do Rio Grande - FURG



Sociedade Brasileira de Matemática Aplicada e Computacional

São Carlos - SP, Brasil
2012

Coordenação Editorial: Elbert Einstein Nehrer Macau

Coordenação Editorial da Série: Rosana Sueli da Motta Jafelice

Editora: SBMAC

Capa: Matheus Botossi Trindade

Patrocínio: SBMAC

Copyright ©2012 by Adriano De Cezaro e Fabiana Travessini De Cezaro.

Direitos reservados, 2012 pela SBMAC. A publicação nesta série não impede o autor de publicar parte ou a totalidade da obra por outra editora, em qualquer meio, desde que faça citação à edição original.

**Catálogo elaborado pela Biblioteca do IBILCE/UNESP
Bibliotecária: Maria Luiza Fernandes Jardim Froner**

De Cezaro, Adriano

Problemas Inversos em Tomografias - São Carlos, SP :
SBMAC, 2012, 132 p., 20.5 cm - (Notas em Matemática
Aplicada; v. 59)

e-ISBN 978-85-8215-010-8

1. Problemas Inversos 2. Tomografias 3. Regularização
I. De Cezaro, Adriano II. De Cezaro, Fabiana T.
III. Título. IV. Série

CDD - 51

Conteúdo

Prefácio	9
1 Problemas Inversos: Uma Breve Introdução por meio de Exemplos	11
1.1 Introdução	11
1.2 O que são Problemas Inversos?	12
1.3 O Problema Inverso da Diferenciação	15
1.4 Transformada de Fourier	17
1.5 Problemas Inversos em Tomografias	19
1.5.1 Tomografia Computadorizada	19
1.5.2 Tomografia Computadorizada: Caso Discreto . .	25
1.5.3 Tomografia por Emissão	27
1.5.4 Tomografia por Fluorescência Óptica	30
1.5.5 Tomografia por Difração	31
1.6 Exercícios	34
2 Transformada de Radon	37
2.1 Propriedades da Transformada de Radon	37
2.2 Fórmula de Inversão para Tomografia Computadorizada	40
2.3 Exercícios	42
3 Transformada de Radon Atenuada	45
3.1 O Problema de Riemann-Hilbert	45
3.2 Fórmula de Inversão para Tomografia por Emissão	46

3.3	Exercícios	55
4	Tomografias por Difração e Aproximações de Born	57
4.1	Espalhamento Inversos em Uma Dimensão	57
4.1.1	Aproximação de Born	58
4.2	Espalhamento Inverso em Dimensão Dois ou Três . . .	59
4.2.1	Aproximação de Born	59
4.2.2	Fórmula de Reconstrução	60
4.3	Exercícios	63
5	Tomografia por Fluorescência Óptica	65
5.1	Uma Fórmula de Inversão Utilizando Séries de Newmann	65
5.2	Exercícios	67
6	Sistemas de Equações Lineares	69
6.1	Pseudo-Inversa de Operadores Lineares	70
6.1.1	Definições e Propriedades Básicas	70
6.2	A Decomposição em Valores Singulares	75
6.2.1	Funções de Operadores: O Teorema da Aplicação Espectral	78
6.2.2	Relação entre Má-colocação e Valores Espectrais	80
6.3	Exercícios	83
7	Regularização para Problemas Inversos	87
7.1	O Conceito de Regularização	88
7.2	Resultados de Convergência	89
7.2.1	Escolha <i>a priori</i> do Parâmetro de Regularização	92
7.2.2	Escolha <i>a posteriori</i> do Parâmetro de Regula- rização	94
7.3	Regularização por Truncamento	95
7.4	Exercícios	97
8	Regularização de Tikhonov	99
8.1	Convêrgencia	99
8.2	Semi-Convergência	102
8.2.1	Taxas de convergência	104

8.3	Regularização de Tikhonov e a Inversa Generalizada . .	104
8.4	Regularização de Tikhonov para Tomografia Computa- dORIZADA	105
8.5	Exercícios	106
9	Regularização por Métodos Iterativos	109
9.1	Método de Landweber	110
9.1.1	Convergência	111
9.2	Método de Máxima Descida e a Inversa Generalizada .	116
9.3	Métodos tipo Kaczmarz	120
9.3.1	Método ART	120
9.3.2	Método de Landweber-Kaczmarz	122
9.4	Aplicação: Tomografia Computadorizada	123
9.5	Exercícios	125
	Bibliografia	127

Prefácio

Este manuscrito tem o intuito de apresentar, de forma introdutória, diversos aspectos relacionados à teoria e aplicações dos problemas inversos e despertar o interesse pela pesquisa nesta área. Uma característica do texto é o tratamento formal do ponto de vista matemático, dos resultados apresentados. Por outro lado, são discutidos aspectos práticos e aplicações, com ênfase a problemas associados a algumas formas de tomografia. A escolha e distribuição dos temas abordados busca dar um equilíbrio entre as aplicações em tomografia e a teoria matemática necessária para resolver estes problemas.

Por se tratar de um texto introdutório, vários resultados relevantes foram omitidos, mas estes podem ser encontrados na bibliografia sugerida. Os autores acreditam que o assunto da forma em que está apresentado no manuscrito coloca o leitor a par do desenvolvimento da teoria e aplicações dos problemas inversos, além de fornecer diversas informações sobre temas de pesquisa atuais nesta área.

O texto é, em sua quase totalidade, autocontido. O leitor familiarizado com conceitos básicos de álgebra linear e análise não deve ter maiores dificuldades para acompanhar boa parte do manuscrito. Em algumas partes do texto é necessário que o leitor tenha intimidade com conceitos básicos de equações diferenciais parciais e de análise funcional, para os quais sugerimos boas referências.

O manuscrito é dividido basicamente em três partes: No Capítulo 1 são introduzidos alguns problemas inversos que servirão de motivação para os demais capítulos. A segunda parte compreende o Capítulo 2 ao Capítulo 5, onde são tratados os aspectos teóricos dos problemas de tomografia propostos no Capítulo 1. O Capítulo 6 ao Capítulo 9 são

dedicados a teoria de regularização para problemas inversos e como esta teoria pode ser aplicada aos problemas tratados anteriormente.

Rio Grande, 30 de Março de 2012.

Adriano De Cezaro
Fabiana Travessini De Cezaro

Capítulo 1

Problemas Inversos: Uma Breve Introdução por meio de Exemplos

1.1 Introdução

O estudo de problemas inversos não é muito recente. Há cerca de dois milênios atrás, no livro VII do diálogo “*República*”, Platão (427 – 347 a.c.) propôs o filosófico problema de reconstruir a “realidade” através da observação da imagem de objetos, cujas sombras eram projetadas na penumbra de uma caverna. Com a ideia de discutir aspectos filosóficos das fontes de conhecimento humano, Platão, também acabou introduzindo o primeiro exemplo de problemas inversos que se tem relatos, [32].

Um dos primeiros problemas inversos na ciência aplicada foi proposto por Eratóstenes (284-202 a.c.). O problema consistia em determinar o diâmetro da terra através de medições feitas em duas cidades distintas. Para tal, eram conhecidas a distância entre as cidades, as suas latitudes e o ângulo que a sombra de um marco vertical (em cada uma destas cidades) fazia com a direção do sol, [16]. O problema inverso de determinar a órbita de um cometa a partir de dados de sua órbita anterior foi estudada por Gauss em 1800. Ele fez uso do

método dos mínimos quadrados para reconstruir a órbita do cometa. Radon em 1917 [33], respondeu a seguinte pergunta: É possível recuperar uma função f , definida em um domínio limitado do plano, a partir do conhecimento dos valores de todas as suas integrais de linha que passam por este domínio? A solução deste problema inverso foi o precursor dos métodos de diagnósticos médicos não-invasivos por Tomografia Computadorizada, [6].

Nos últimos anos, os problemas inversos têm conquistado uma grande quantidade de pesquisadores adeptos. Esta área de pesquisa trata de problemas como os formulados por Platão, Eratóstenes, entre outros, mas cuja abordagem exige o desenvolvimento de métodos matemáticos como os apresentados por Gauss e Radon. O súbito crescimento é devido, certamente, ao grande número de aplicações nas outras ciências e o aparato de novas técnicas e teorias matemáticas envolvidas no ataque a tais problemas. Por exemplo, problemas inversos aparecem naturalmente na geofísica, nas explorações sísmicas para detecção de depósito de petróleo [28, 39] e nas ciências médicas e tomografias com ênfase na reconstrução de imagens e diagnósticos não-invasivos, ultrassonografia [30, 31], na engenharia com testes não-destrutivos em componentes de semi-condutores e nanotecnologia, [5, 10, 4]. Além da relevância das aplicações, a formulação e solução de tais problemas envolvem o conhecimento de vários campos da matemática, das ciências aplicadas e o envolvimento de profissionais dessas áreas.

1.2 O que são Problemas Inversos?

Alguns problemas inversos já foram introduzidos neste manuscrito, mas ainda devemos responder a pergunta: O que são os tais ”**problemas inversos**?” de forma mais clara. Ou ainda, a pergunta pode ser: ”**inversos do quê**?” Para J.B. Keller [22], dois problemas são o inverso um do outro, se a formulação de um envolve o conhecimento (mesmo que parcial) do outro. A grosso modo, *Problemas Inversos* estão relacionados com a determinação de **causas** através da observação (ou medida) de **efeitos**. O problema contrário é dito ser o ”**Problema**

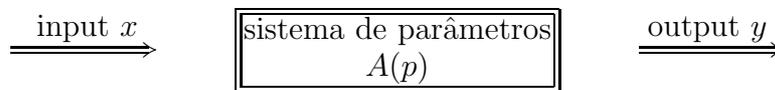
Direto”.

Do ponto de vista das aplicações, existem pelo menos duas motivações distintas para estudar ”Problemas Inversos”. A primeira é movida pela curiosidade humana de conhecer estados físicos passados ou parâmetros em um sistema físico que descreve certos modelos. A segunda é prever os fenômenos futuros influenciados pelos estados atuais. Ambas motivações são modeladas por equações matemáticas.

Associado ao estudo e solução de problemas inversos estão fatores relevantes no desenvolvimento de uma nação. Por exemplo, problemas inversos em imagens médicas influenciam em

- **fatores sociais:** técnicas de detecção de tumores implicam em prolongar a vida das pessoas.
- **fatores econômicos:** detecção de tumores implica em tratamentos mais eficazes contra o câncer, diminuindo os custos dos mesmos. Ainda, prolonga a vida ativa das pessoas que, consequentemente, geram mais riquezas.
- **desenvolvimento tecnológico:** desenvolvimento de novos métodos e máquinas de tomografia para a obtenção de imagens médicas.

Sintetizando as ideias acima, podemos assumir que o fenômeno físico (biológico e etc) a ser estudado é modelado por um processo que envolve três quantidades principais do modelo:



O problema direto: Dados o input (**causa**) e o sistema de parâmetros, determinar o output do modelo (**efeito**).

O problema inverso: Esse pode aparecer de duas formas.

1. *O problema de reconstrução:* Dado o sistema de parâmetros e observado o output (**efeito**), encontrar que input (**causa**) que corresponde ao output.
2. *O problema de identificação.* Dados input (**causa**) e output (**efeito**), determinar o sistema de parâmetros que relaciona input/output.

De um modo geral, representamos um problema inverso por equações do tipo

$$A(p)x = y, \quad (1.2.1)$$

para um dado efeito y ; uma causa a ser determinada x (que pertence a um determinado conjunto de parâmetros admissíveis) e $A(p)$ representa o modelo que manda a causa no determinado efeito.

Em termos práticos, o efeito y dificilmente é obtido de forma precisa, pois sua obtenção é feita através de medidas. Assim, costumamos denotar as medições obtidas por y^δ , das quais assumimos conhecer o nível de **ruídos** δ , de forma a satisfazer

$$\|y - y^\delta\| \leq \delta. \quad (1.2.2)$$

Na formulação matemática, $A(p)$ é um operador, por exemplo uma matriz, definido entre espaços vetoriais que, para nossos objetivos, consideraremos espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 , [24].

Denotaremos por \mathcal{U} o conjunto de parâmetros admissíveis, obtemos as seguintes caracterizações:

O problema direto: Dado $x \in \mathcal{H}_1$ e $p \in \mathcal{U}$, encontrar $y := A(p)x$.

O problema inverso: Esse aparece, pelo menos de duas formas:

1. *O problema de reconstrução:* Observado $y \in \mathcal{H}_2$ e conhecido sistema de parâmetros $A(p)$ para $p \in \mathcal{U}$, encontrar $x \in \mathcal{H}_1$ tal que $A(p)x = y$.
2. *O problema de identificação.* Dados $x \in \mathcal{H}_1$ e $y \in \mathcal{H}_2$, encontrar $p \in \mathcal{U}$ tal que $A(p)x = y$.

Uma característica importante que diferencia um **problema direto** de um **problema inverso** é que o segundo é *mal-posto* no sentido de Hadamard, [18]. Um problema é dito *bem-posto* no sentido de Hadamard se satisfaz as condições de existência, unicidade e dependência contínua dos dados iniciais. Caso um dos requerimentos acima não seja satisfeito, o problema é dito mal-posto.

Matematicamente, a questão de existência de soluções pode ser enfraquecida, bastando para tanto aumentar o conjunto de soluções, por exemplo, admitindo soluções generalizadas para o problema. Em relação a unicidade, sua ausência significa que estão faltando informações no modelo. Podemos adicionar propriedades, impor mais condições ao operador, ou ainda, podemos escolher uma entre as várias soluções, a qual melhor corresponde ao modelo. A violação da estabilidade é a de maior importância. Se um problema não possui a propriedade de dependência contínua dos dados, surgem sérias complicações para obtermos uma solução aproximada. Tais complicações não podem ser compensadas por "truques matemáticos". Como os dados não são obtidos exatamente (veja a equação (1.2.2)), a falta de dependência contínua implica, em geral, em desastres na obtenção de soluções. Para evitarmos estes desastres, precisamos de métodos adequados para a obtenção de soluções. Estes métodos são chamados de métodos de regularização e serão estudados no Capítulo 7.

Problemas inversos em dimensão finita aparecem naturalmente na solução de sistemas de equações ou na discretização (aproximação numérica) de problemas em dimensão infinita [10, 5, 3, 19]. Do ponto de vista computacional [38, 19], sempre tratamos de problemas inversos em dimensão finita, assim, precisamos estudá-los com o devido interesse. Ainda, os problemas de mal-condicionamento estão intimamente ligados aos autovalores da matriz que representa o sistema. Se os autovalores são próximos de zero ou os dados não pertencem ao espaço solução do problema, podemos enfrentar sérias complicações numéricas, [29, 13].

1.3 O Problema Inverso da Diferenciação

Nesta seção apresentaremos o problema inverso da diferenciação, o qual é uma forma de motivar a necessidade de estudos detalhados da teoria de problemas inversos.

Sejam $y : [0, 1] \rightarrow \mathbb{R}$ e $y^\delta : [0, 1] \rightarrow \mathbb{R}$ funções contínuas com y^δ contaminada por ruídos de forma que

$$\|y(t) - y^\delta(t)\|_\infty \leq \delta, \quad \forall t \in [0, 1].$$

Gostaríamos de reconstruir a derivada $x = y'$ de y . Uma das estratégias (pelo menos do ponto de vista numérico) é considerar aproximações por diferenças simétricas, isto é, para qualquer $\tau \in (0, 1)$,

$$x^{\delta,h}(\tau) := \frac{y^\delta(\tau + h) - y^\delta(\tau - h)}{2h}.$$

Gostaríamos que $x^{\delta,h}(\tau)$ se aproximasse de x quando h tende a zero. No entanto, um simples argumento com a desigualdade triangular fornece

$$\begin{aligned} \|x^{\delta,h}(\tau) - x(\tau)\|_\infty &\leq \left\| \frac{y(\tau + h) - y(\tau - h)}{2h} - x(\tau) \right\|_\infty \\ &\quad + \left\| \frac{(y^\delta - y)(\tau + h) - (y^\delta - y)(\tau - h)}{2h} \right\|_\infty. \end{aligned}$$

Suponhamos que

$$\|x'(t)\|_\infty \leq E, \quad \forall t \in [0, 1].$$

Substituímos o limitante na desigualdade acima, obtemos a estimativa de erro

$$\|x^{\delta,h}(\tau) - x(\tau)\|_\infty \leq hE + \frac{\delta}{h}. \quad (1.3.3)$$

A equação (1.3.3) é típica em problemas inversos e reaparecerá novamente em nossas notas. O que é importante, por agora, é entender o que ela quer nos dizer. Temos dois termos nessa estimativa de erro: um devido a aproximação da aplicação inversa (hE) e o outro devido ao erro de medida ($\frac{\delta}{h}$). Observe que, quanto mais refinarmos a aproximação (quanto mais próximo de zero tomarmos h) mais precisamente estamos calculando a derivada y' . Por outro lado, como os dados estão corrompidos por erros, (1.3.3) nos ensina que, se h for tomado muito pequeno, então $x^{\delta,h}$ pode estar longe da solução verdadeira.

O melhor que podemos fazer é escolher h de forma a balancear o lado direito de (1.3.3). A Figura 1.1 ilustra a escolha de $h(\delta)$. Ou seja, tomamos

$$h(\delta) := E^{-\frac{1}{2}} \delta^{\frac{1}{2}}.$$

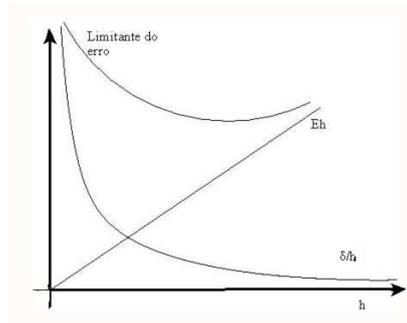


Figura 1.1: Estimativa fundamental.

1.4 Transformada de Fourier

Nesta seção veremos alguns fatos importantes da Transformada de Fourier e a relação desta importante ferramenta com os problemas inversos. Para que este manuscrito seja o mais autocontido, iniciaremos esta seção introduzindo alguns espaços de funções fundamentais para o entendimento do que segue.

Uma função mensurável a valores reais f é dita pertencer a $L^p(\mathbb{R}^n)$, com $1 \leq p < \infty$ se

$$\|f\|_{L^p}^p = \int_{\mathbb{R}^n} |f(x)|^p dx < \infty. \quad (1.4.4)$$

Neste caso, (1.4.4) define a norma no espaço de funções $L^p(\mathbb{R}^n)$, [12].

Em particular, o espaço $L^2(\mathbb{R}^n)$ é um espaço de Hilbert com o produto interno dado por

$$\langle f, g \rangle = \int_{\mathbb{R}^n} f(x) g(x) dx. \quad (1.4.5)$$

Para as funções $f \in L^2(\mathbb{R}^n)$, definimos a **Transformada de Fourier de f** por

$$\widehat{f}(k) = (\mathcal{F}f)(k) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-ikx} f(x) dx, \quad (1.4.6)$$

cuja **inversa** é dada por

$$(\mathcal{F}^{-1}f)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ikx} \hat{f}(k) dk. \quad (1.4.7)$$

Teorema 1.1. *A Transformada de Fourier $\mathcal{F} : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$ é uma isometria.*

Demonstração: É fácil ver que a Transformada de Fourier e sua inversa são operadores lineares. A identidade de Parseval, [12], implica que

$$\langle \hat{f}, \hat{g} \rangle = (2\pi)^n \langle f, g \rangle, \quad \forall f, g \in L^2(\mathbb{R}^n). \quad (1.4.8)$$

De (1.4.8) segue que

$$\|\hat{f}\|_{L^2(\mathbb{R}^n)} = (2\pi)^{n/2} \|f\|_{L^2},$$

o que conclui a demonstração. ■

O teorema abaixo diz que os problemas inversos que têm na sua formulação a dependência da Transformada de Fourier são bem-postos no sentido de Hadamard.

Teorema 1.2. *A Transformada de Fourier $\mathcal{F} : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$ é um operador bem posto.*

Demonstração: **Existência e unicidade:** Seguem diretamente do Teorema 1.1 que garante que a Transformada de Fourier é uma isometria.

Estabilidade: Sejam $f, f^\delta \in L^2(\mathbb{R}^n)$ tais que $\hat{f}^\delta(k) = \hat{f}(k) + \hat{g}(k)$ com $\|\hat{g}\|_{L^2} \leq (2\pi)^{n/2} \delta$. Da identidade de Parseval (1.4.8) e da linearidade da Transformada de Fourier segue que

$$\|f - f^\delta\|_{L^2} = \|\mathcal{F}^{-1}(\hat{f} - \hat{f}^\delta)\|_{L^2} = (2\pi)^{-n/2} \|\hat{g}\|_{L^2} \leq \delta.$$

■

1.5 Problemas Inversos em Tomografias

Entre os problemas que têm chamado a atenção da comunidade científica estão os problemas em tomografias e suas aplicações [11, 27, 26, 34, 31, 30]. Em outras palavras, tais problemas consistem em recuperar a forma ou a densidade e a localização de um objeto imerso em uma região a partir de medidas (parciais) sobre a fronteira da região. As medidas, na maioria dos casos, são adquiridas por um número reduzido de experimentos, [17]. Assim, uma característica comum destes tipos de problemas é a falta de informações e, conseqüentemente, ruídos nos dados.

Existem vários tipos de tomografias. Neste manuscrito, estudaremos o problema inverso associado a Tomografia Computadorizada, Tomografia por Emissão e por Difração, [2, 5, 6, 31, 31].

1.5.1 Tomografia Computadorizada

A Tomografia Computadorizada (TC) é um método de obtenção de diagnósticos baseados em imagens internas de uma região, de forma não-invasiva. As informações são obtidas somente na parte exterior deste objeto através da medição das intensidades de feixes de raios-X que passam por esse objeto. Depois, estas informações são processadas por um algoritmo computacional que as transforma em uma imagem bi ou tridimensional.

Em outras palavras, TC é baseada nos mesmos princípios da radiografia convencional, isto é, os tecidos com diferentes composições absorvem a radiação emitida pelo raio-X de forma diferente. Ao serem atravessados por raios-X, tecidos mais densos, por exemplo o fígado, ou com elementos mais pesados, como o cálcio presente nos ossos, absorvem mais radiação que tecidos menos densos, como o pulmão, que está cheio de ar. Assim, um processo de TC indica a quantidade de radiação absorvida por cada parte do corpo analisada (radiodensidade) e traduz essas variações numa escala de cinza produzindo uma imagem. Cada pixel da imagem corresponde à média da absorção dos tecidos na região escaneada pelo processo.

Descrição do Problema: Uma maneira intuitiva de compreender o processo de reconstrução de imagens usando TC é considerar o caso em que possuímos apenas um feixe de raio-X em forma de uma reta. Fisicamente, qualquer raio que passe por um obstáculo é atenuado. A atenuação da intensidade de radiação, medida pelo detector que está do lado oposto à fonte de emissão com relação a um obstáculo homogêneo, pode ser modelada por um único coeficiente de absorção μ . Seja L a reta pela qual está direcionado o feixe de fótons de raio-X. Chamamos de I_0 a intensidade da aplicação (input) do feixe de fótons de raio-X e por I a intensidade depois de sair de Ω , veja Figuras 1.2 e 1.3.

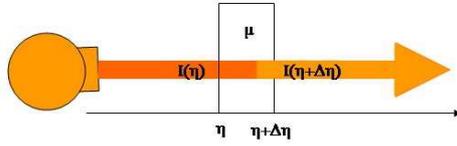


Figura 1.2: Atenuação do raio-X.

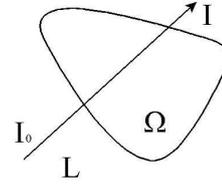


Figura 1.3: Raio-X.

Com esse modelo simples, a atenuação total de um feixe de raio-X monocromático pode ser calculada da seguinte forma: a intensidade de radiação I , após percorrer a distância $\Delta\eta$ através de um objeto, é dada por

$$I(\eta + \Delta\eta) = I(\eta) - \mu(\eta)I(\eta)\Delta\eta. \quad (1.5.9)$$

Se reescrevermos a equação (1.5.9), temos que

$$\frac{I(\eta + \Delta\eta) - I(\eta)}{\Delta\eta} = -\mu(\eta)I(\eta). \quad (1.5.10)$$

Suponhamos que $\Delta\eta$ seja infinitesimal, isto é, se fizermos $\Delta\eta \rightarrow 0$ em (1.5.10), obtemos a seguinte equação diferencial ordinária

$$\frac{dI}{d\eta} = -\mu(\eta)I(\eta). \quad (1.5.11)$$

Para determinarmos I , integramos (1.5.11) ao longo da reta L (método de separação de variáveis) e obtemos que

$$I = I_0 \exp \left(- \int_L \mu(\eta) d\eta \right), \quad (1.5.12)$$

onde L é o comprimento do feixe. O coeficiente de absorção μ compreende os efeitos de absorção e de espalhamento do raio-X que é emitido.

A equação (1.5.12) é conhecida como **lei de Beer** e nos ensina que a intensidade de raio-X é atenuada exponencialmente ao longo de L , [6], onde consideramos os efeitos de espalhamento desprezíveis. A quantidade

$$p(L) = - \ln \left(\frac{I}{I_0} \right) = \int_L \mu(\eta) d\eta, \quad (1.5.13)$$

mede a variação da intensidade de raio-X, ou seja, a atenuação do raio-X durante o processo. Denominaremos $p(L)$ em (1.5.13) de projeção ou projeção integral.

Na prática, as projeções $p(L)$, dadas pela equação (1.5.13), podem ser medidas somente sobre uma quantidade finita de retas L . Dependendo de como tais medidas são feitas, basicamente, duas formas de geometrias são obtidas. A primeira, em um modelo de escaneamento paralelo, um conjunto de integrais de linhas como (1.5.13) é obtido sobre um conjunto de retas L_j paralelas e igualmente espaçadas. Esse modelo requer uma única fonte de emissão e um único detector, os quais se movem e rodam (com um ângulo pré-determinado) durante o processo de escaneamento, veja a Figura 1.4. A outra forma de escaneamento é dada por um feixe na forma de cone de raios-X. Nesta, a fonte circula o objeto a ser escaneado e as intensidades são medidas por uma linha de detectores, simultaneamente, a cada posição da fonte, veja Figura 1.4.

Para obter informações suficientes para reconstruir uma imagem bidimensional, o feixe de raio-X é movimentado de forma paralela a distribuição linear dos detectores. De forma sequencial, o conjunto de detectores é rotacionado por um ângulo $\theta \in [0, 2\pi)$, de forma que todo

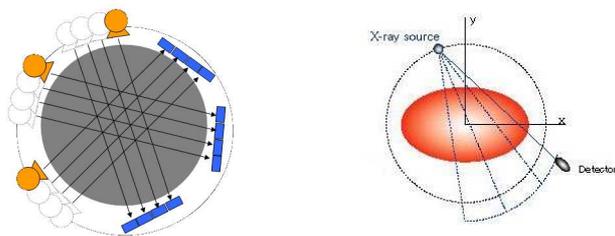


Figura 1.4: Tomografia por raios-X paralela (E), do tipo cone (D).

o objeto de interesse seja escaneado. O processo é representado pela Figura 1.4.

Assim, para um ângulo θ fixo e para uma posição em particular da fonte de raio-X, o processo é descrito pela projeção integral (1.5.13). Desta maneira, é razoável reescrevermos a equação (1.5.13) em função da posição da fonte de raio-X - ou da localização do correspondente detector ξ - e do ângulo de projeção θ . Assim, consideraremos o sistema de coordenada (ξ, η) , o qual gira juntamente com o raio-X e a fonte de detectores, ao invés do sistema de coordenadas fixo (x, y) . A Figura 1.5 mostra a correspondência entre os dois sistemas¹.

No novo sistema de variáveis, a equação (1.5.13) fica reescrita como

$$p_{\theta}(\xi) = \int_L \mu(\xi, \eta) d\eta. \quad (1.5.14)$$

A equação (1.5.14) é uma integral de linha ao longo do segmento de reta L . Ela descreve a posição ξ da fonte de raio-X e do respectivo detector com um ângulo de projeção θ em relação ao plano de coordenadas (x, y) .

Na prática, os valores do coeficiente de atenuação devem ser dados como uma função das coordenadas fixas $\mathbf{x} = (x, y)$. Para descrever a relação entre os sistemas de coordenadas (ξ, η) e (x, y) , definimos² o vetor $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta)) \in \mathbb{S}^1$ e seja $\boldsymbol{\theta}^{\perp}$ o vetor unitário ortogonal a

¹A Figura 1.5 foi retirada de [6]

² \mathbb{S}^{n-1} denota a esfera unitária em \mathbb{R}^n , para $n \geq 2$.

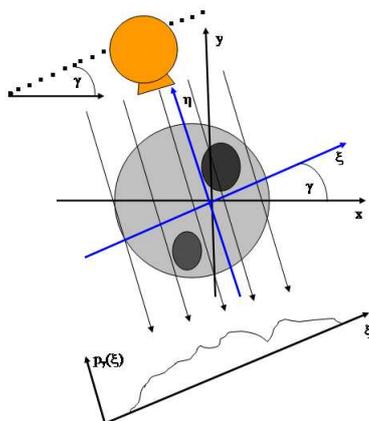


Figura 1.5: Mudança de variável.

θ . Com essa notação, temos que

$$\xi = x \cos(\theta) + y \sin(\theta), \quad \eta = -x \sin(\theta) + y \cos(\theta). \quad (1.5.15)$$

Substituímos (1.5.15) em $\mu(\xi, \eta)$, obtemos o coeficiente de atenuação nas variáveis $\mathbf{x} = (x, y)$ definido por

$$f(\mathbf{x}) = \mu(x \cos(\theta) + y \sin(\theta), -x \sin(\theta) + y \cos(\theta)). \quad (1.5.16)$$

Assim, do ponto de vista físico, os coeficientes de atenuação $f(x, y)$ e $\mu(\xi, \eta)$ são os mesmos.

Exercício 1.1. *Mostre que qualquer ponto $\mathbf{x} = (x, y)$ do plano cartesiano que está a uma distância $s = \sqrt{x^2 + y^2}$ da origem pode ser parametrizado como $\mathbf{x} \cdot \boldsymbol{\theta}^\perp = s$, onde $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta))$ e $\theta \in [0, 2\pi)$. Veja a Figura 1.6.*

Se considerarmos $I_0 = 1$ (normalizado) na equação (1.5.13) e parametrizando a reta L de acordo com o Exercício 1.1, obtemos

$$p_\theta(\xi) = \mathcal{R}f(\theta, s) = \int_{\mathbf{x} \cdot \boldsymbol{\theta}^\perp = s} f dl = \int_{\mathbb{R}} f(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}) dt. \quad (1.5.17)$$

O operador \mathcal{R} é chamada de **Transformada de Radon** bi-dimensional de f . Esta é uma aplicação que leva funções de $(x, y) \in \mathbb{R}^2$ em funções de $(\theta, s) \in \mathbb{S}^1 \times \mathbb{R}$.

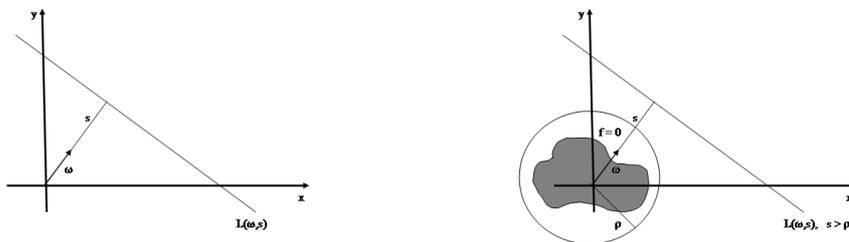


Figura 1.6: Parametrização da reta L que é perpendicular a w a uma distância s da origem (D) e Se $f(x, y) = 0$, quando $|(x, y)| > \rho$, então $\mathcal{R}f(w, s) = 0$ quando $s > \rho$.

O **problema inverso** associado a tomografia é: Encontrar uma aproximação apropriada para a distribuição de densidade f através da medida dos outputs de várias seções transversais em diferentes ângulos da região Ω . Ou seja, a solução do problema inverso de reconstruir a densidade f consiste em inverter o operador \mathcal{R} (em um espaço de funções adequado).

Segue uma pergunta importante: É possível determinar o coeficiente de absorção f a partir de sua integral de linha (1.5.17) ao longo da reta L ? Esta pergunta é exatamente a mesma feita por *Johann Radon* em seu famoso trabalho de 1917, [33]. A tentativa de respondê-la, nos leva a várias outras perguntas:

1. \mathcal{R} é injetiva?
2. Qual é a imagem de \mathcal{R} ?
3. É possível encontrar uma fórmula para a inversa de \mathcal{R} ?
4. Se $\mathcal{R}f(\theta, s) = 0$ para $|s| > \rho$ e para todas as direções $\theta = (\cos(\theta), \sin(\theta)) \in \mathbb{S}^1$, é verdade que $f(x, y) = 0$ para todo $x \in \mathbb{R}^2$ com $|x| > \rho$? Veja Figura 1.6.

Do ponto de vista prático, é impossível obtermos medidas sobre todas as retas que passam por Ω . Deste fato, surge uma outra pergunta:

Quantas medidas são necessárias para obtermos uma boa aproximação? Esta, talvez, seja a mais difícil. Abaixo apresentaremos uma abordagem discreta da TC onde aparecerá, naturalmente, tal questão. Para mais detalhes consulte [30, 31].

1.5.2 Tomografia Computadorizada: Caso Discreto

Nesta seção, faremos uma breve discussão do aspecto prático da TC e da Transformada de Radon. Como comentamos anteriormente, do ponto de vista prático, é impossível obtermos a integral de linha de f em todas as direções. Na verdade, o conjunto de informações obtidas num processo de tomografia são valores da Transformada de Radon medidos por N detectores.

As limitações físicas do processo de medição implicam em uma discretização da imagem tomográfica. O tamanho e o número N de pixels, dentro do campo de visão que devem ser reconstruídos, consistem de um vetor de variáveis desconhecidas f_j para $j = \{1, \dots, N\}$. Os f_j são os coeficientes de atenuação. A Figura 1.7 mostra, esquematicamente, uma imagem tomográfica a ser reconstituída³.

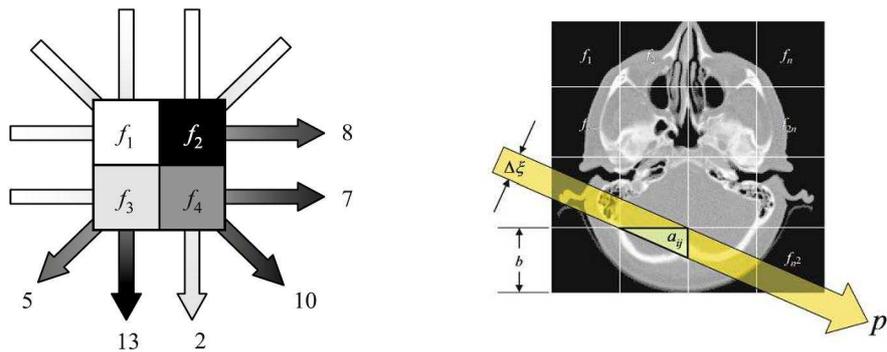


Figura 1.7: Tomografia discreta (D) e feixe de raio-X (E).

Na Figura 1.7 os valores p_i representam as medições realizadas nos N detectores. Em outras palavras, os valores p_i são valores da Trans-

³A Figura 1.7 foi retirada de [6].

formada de Radon, para cada um dos raios que atravessam o objeto a ser imageado. Na Figura 1.7, a imagem que está a direita traz alguns valores das medidas p_i , para uma quantidade de 6 detectores.

Fisicamente, cada feixe de raio-X possui uma espessura. Quando o feixe de raio-X passa pela região Ω , temos que levar em conta quanto do pixel a ser reconstruído é afetado pelo feixe. Para este propósito, são introduzidos pesos que refletem a relação entre a área iluminada pelo feixe de raio-X com relação a área total do pixel. A Figura 1.7 ilustra a situação.

Para um feixe de espessura $\Delta\xi$, o peso a_{ij} é determinado pela relação

$$a_{ij} = \frac{\text{área iluminada do pixel } j \text{ pelo raio } i}{\text{área total do pixel } j}. \quad (1.5.18)$$

Assim, para um conjunto de f_j , $j = \{1, \dots, N\}$, densidades a serem determinadas e dado um conjunto de $i = \{1, \dots, M\}$ raios-X medidos, com intensidade p_i , obtemos um sistema de M equações lineares com N coeficientes

$$\sum_{j=1}^N a_{ij} f_j = p_i \quad i = \{1, \dots, M\}. \quad (1.5.19)$$

Escrevemos na forma matricial:

$$Af = p, \quad (1.5.20)$$

onde $A = (a_{ij})_{M \times N}$ pode ser considerada como a **caixa preta da máquina de tomografia**.

Fazemos uma comparação direta entre (1.5.17) e (1.5.20):

$$\begin{array}{ccc} A & f & = & p \\ \updownarrow & \updownarrow & = & \updownarrow \\ \mathcal{R} & f(\theta, s) & = & \int_{L(\theta, s)} f dl \end{array}$$

As das dificuldades de reconstrução no modelo discreto são:

- O sistema (1.5.20) possui solução exata somente com condições ideais. Para dados reais, a presença de ruídos implica em obter apenas soluções aproximadas do sistema (1.5.20), mesmo quando $M = N$. No caso em que $M > N$, isto é, que temos mais informações (medidas) que o número de densidades a serem determinadas, possivelmente, obtemos reconstruções melhores da densidade.
- Tipicamente, a matriz A é singular e, em geral, não-quadrada. Isto indica que o problema é mal-posto.
- A não possui uma estrutura simples. Assim, mesmo que A seja não singular, é difícil determinar uma maneira de resolver o sistema (1.5.20) de forma eficaz e com pouco custo computacional.
- Nos problemas práticos, a dimensão de A é muito grande, assim, métodos diretos de inversão são inapropriados, pois são computacionalmente muito intensos e custosos.

1.5.3 Tomografia por Emissão

A operacionalização das tomografias por emissão, que temos interesse nestas notas, basicamente consiste na injeção de um radio-isótopo no ser vivo (em geral, na circulação sanguínea). O marcador é quimicamente incorporado a uma molécula biologicamente ativa. Há um período de espera para que as moléculas se concentrem nos tecidos de interesse. O raio-isótopo sofre decaimento por emissão de pósitrons (um anti-elétron). Esta radiação se propaga através do meio. Parte desta radiação é absorvida neste processo. O que ultrapassa a fronteira do domínio é medida numa máquina de tomografia.

Da mesma forma que em Tomografia Computadorizada, o problema inverso em Tomografia por Emissão tem por objetivo reconstruir a densidade f do meio através das medições das radiações no exterior do domínio.

Exemplos de Tomografia por Emissão são SPECT (Single Photon Emission Computer Tomography) e PET (Positron Emission Tomography). A diferença entre estas tomografias está, essencialmente,

ligada ao tipo de emissão. No processo de SPECT é emitido um único raio gama que é medido diretamente pela máquina de tomografia, enquanto em PET, o marcador emite um pósitron que se aniquila com um elétron a poucos milímetros da fonte de emissão, produzindo um par de raios gama a serem emitidos em direções opostas. O escaner em PET detecta a emissão "coincidente" de ambos os raios gama. Isto produz mais informações que o processo de SPECT. Por outro lado, os aparelhos de SPECT são muito mais baratos que os de PET.

Tomografia por Emissão de um Fóton Simples - SPECT

Para o modelo de SPECT, denotaremos por $u(\mathbf{x}, \theta)$ a densidade do fóton irradiado e por $f(\mathbf{x})$ a densidade da fonte de radiação dos fótons. Suponhamos que a absorção de fótons pelo meio (pelos tecidos do corpo humano no caso de aplicações médicas) sejam modeladas pelo coeficiente $a(\mathbf{x})$. O coeficiente $a(\mathbf{x})$ pode ser obtido, por exemplo, através de um método de TC como na Subseção 1.5.1. Assim, neste manuscrito, assumiremos que $a(\mathbf{x})$ é conhecido.

A densidade $u(\mathbf{x}, \theta)$ satisfaz a seguinte equação do transporte

$$\boldsymbol{\theta} \cdot \nabla u(\mathbf{x}, \theta) + a(\mathbf{x})u(\mathbf{x}, \theta) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^2, \boldsymbol{\theta} \in \mathbb{S}^1, \quad (1.5.21)$$

onde $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta))$ determina a orientação, para $\theta \in [0, 2\pi[$. Assumiremos que $f(\mathbf{x})$ possui suporte compacto e que não existe fontes de radiação que venham do infinito, o que dá uma condição de fronteira para a EDP (1.5.21)

$$\lim_{s \rightarrow \infty} u(\mathbf{x} - s\boldsymbol{\theta}, \theta) = 0. \quad (1.5.22)$$

Definimos a seguinte simetrização

$$\mathcal{A}a(\mathbf{x}) = \frac{1}{2} \int_0^\infty [a(\mathbf{x} - t\boldsymbol{\theta}) - a(\mathbf{x} + t\boldsymbol{\theta})] dt.$$

Disto segue que

$$\boldsymbol{\theta} \cdot \nabla \mathcal{A}a(\mathbf{x}) = a(\mathbf{x}). \quad (1.5.23)$$

Assim, $e^{\mathcal{A}a(\mathbf{x})}$ é um fator integrante para a EDP (1.5.21), no sentido que

$$\boldsymbol{\theta} \cdot \nabla (e^{\mathcal{A}a(\mathbf{x})} u(\mathbf{x}, \theta)) = e^{\mathcal{A}a(\mathbf{x})} f(\mathbf{x}, \theta). \quad (1.5.24)$$

Portanto,

$$e^{\mathcal{A}a(\mathbf{x})} u(\mathbf{x}, \theta) = \int_0^\infty e^{\mathcal{A}a(\mathbf{x})} f(\mathbf{x} - t\boldsymbol{\theta}, \theta) dt \quad (1.5.25)$$

é a única solução da EDP (1.5.21), [12].

Como $\mathbf{x} = s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}$ e denotando $\mathcal{A}a(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}) = \mathcal{A}a$ segue da equação (1.5.25) que

$$\lim_{t \rightarrow \infty} e^{\mathcal{A}a} u(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta) = \int_{-\infty}^\infty (e^{\mathcal{A}a} f)(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta) dt. \quad (1.5.26)$$

Notemos que $u(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta)$ é a radiação fora do domínio e, portanto, mensurável. Ainda, $e^{\mathcal{A}a}$ envolve o coeficiente de atenuação $a(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta})$ que assumimos conhecido. Assim, o lado esquerdo da equação (1.5.26) é conhecido.

O **problema inverso** no modelo de SPECT é recuperar f no lado direito da equação (1.5.26).

Definimos a **Transformada de Radon Atenuada**

$$(\mathcal{R}_a f)(\theta, s) = (\mathcal{R}(e^{\mathcal{A}a} f))(\theta, s) = \int_{-\infty}^\infty (e^{\mathcal{A}a} f)(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta) dt, \quad (1.5.27)$$

onde \mathcal{R} é a Transformada de Radon definida em (1.5.17).

Observação 1.1. *Notemos que, se $a(\mathbf{x}) = 0$, a equação (1.5.27) é a Transformada de Radon de $f(\mathbf{x})$, como definido em (1.5.17). Se $a(\mathbf{x})$ é constante, os mesmos argumentos para obter uma fórmula de inversão são bem conhecidos [2].*

Existem duas técnicas independentes para a obtenção de uma fórmula de inversão da Transformada de Radon (1.5.27). Estas são o método de funções A -analíticas, [2], e o método baseado na extensão da equação do transporte no plano complexo e na solução do problema de Riemann-Hilbert [2]. No Capítulo 3 apresentaremos a última das técnicas.

Tomografia por Emissão de Pósitrons - PET

Na Tomografia de Emissão de Pósitrons (PET) a fonte de radiação emite partículas em direções opostas que são medidas em detectores opostos. Essa é a principal diferença entre PET e SPECT. Assim, a PET é modelada pela equação do transporte (1.5.21) com condições de fronteira dadas por (1.5.22), com o limite sendo tomado em $\pm\infty$.

Portanto, para derivar o modelo matemático por trás da PET é similar ao modelo de SPECT bastando para tal definir a simetriação

$$\tilde{\mathcal{A}}a(\mathbf{x}) = \frac{1}{2} \int_{-\infty}^{\infty} [a(\mathbf{x} - t\boldsymbol{\theta}) - a(\mathbf{x} + t\boldsymbol{\theta})] dt$$

e proceder como feito no caso da tomografia SPECT. Os detalhes são deixados para o leitor.

O **problema inverso** no modelo de PET é recuperar f no lado direito da equação (1.5.26) onde $\mathcal{A}a(\mathbf{x})$ deve ser substituído por $\tilde{\mathcal{A}}a(\mathbf{x})$.

Como já dissemos, em PET as fontes ejetam um par de partículas em direções opostas. Nesse processo, as medidas só são realizadas quando ambas as partículas chegam ao mesmo tempo. Assim, muita informação contida no processo de PET é perdida, ou ainda, não é levada em consideração. Devido ao pequeno número de eventos mensurados os processos de PET podem e devem, por natureza, ser considerados estocásticos. Deste modo, recentemente, muitos modelos estocásticos tem sido usados em tomografias por emissão em substituição aos métodos que usam transformadas integrais como a Transformada de Radon. Tais métodos são puramente discretos. Para mais detalhes veja [30, 31].

1.5.4 Tomografia por Fluorescência Óptica

Tomografia óptica é uma forma de tomografia computadorizada que procura reconstruir a imagem de um objeto a partir de luz transmitida e espalhada através de um objeto. Partículas fluorescentes são injetadas em um corpo e emitem (perto de infra-vermelhos) luz sob certas condições. A luz emitida é medida na fronteira do domínio.

O modelo de tomografia por fluorescência óptica que estudaremos é um modelo simplificado, no qual vamos supor que a propagação do fóton é bidimensional. Assim, o **problema direto** no modelo de tomografia de fluorescência óptica é obter $u(\mathbf{x}, \theta)$ solução da equação de transporte bidimensional

$$\boldsymbol{\theta} \cdot \nabla u(\mathbf{x}, \theta) + a(x)u(\mathbf{x}, \theta) = Ku(\mathbf{x}, \theta) + f(\mathbf{x}), \quad (1.5.28)$$

com fonte dependendo do coeficiente de espalhamento $\sigma_s(\mathbf{x}, \boldsymbol{\theta} \cdot \boldsymbol{\theta}^\perp) = \sigma_s(\mathbf{x}, \theta - \theta')$ e condição de fronteira dada por

$$\lim_{t \rightarrow \infty} u(\mathbf{x} - t\boldsymbol{\theta}, \theta) = 0, \quad (1.5.29)$$

onde $a(\mathbf{x})$ e $f(\mathbf{x})$ são conhecidos e

$$Ku(\mathbf{x}, \theta) = \int_{\mathbb{S}^1} \sigma_s(\mathbf{x}, \boldsymbol{\theta} \cdot \boldsymbol{\theta}^\perp, \theta')u(\mathbf{x}, \theta')d\theta', \quad (1.5.30)$$

é o operador de espalhamento.

O que é medido no modelo de tomografia por fluorescência óptica é a quantidade

$$g(\theta, s) = \lim_{t \rightarrow \infty} [e^{Aa}u](s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}). \quad (1.5.31)$$

O **problema inverso** no modelo de tomografia por fluorescência óptica é reconstruir a densidade $f(\mathbf{x})$ de partículas fluorescentes a partir das medidas como em (1.5.31). Como as aplicações de moléculas fluorescentes em humanos ainda são limitadas, a maior parte do trabalho em tomografia de fluorescência tem sido no campo da pesquisa pré-clínica do câncer, ou ainda aplicação na indústria como um sensor de espessura e da estrutura interna de semicondutores.

Obteremos uma fórmula de inversão para o modelo de tomografia por fluorescência óptica no Capítulo 5.

1.5.5 Tomografia por Difração

Tomografias por Difração são baseadas em ondas acústicas ou micro-ondas. Estas ondas são caracterizadas por ter um comprimento de

onda suficientemente pequeno de forma que possa capturar as formas de interesse dos objetos, mas também grande o suficiente tal que o limite das grandes frequências ópticas (onde ondas são substituídas por partículas) não são válidas. Exemplos de tais tomografias são ultrassom e tomografias por microondas. Estas ondas não seguem ao longo de raios e dependem da estrutura interna dos objetos a serem imageados.

Na tomografia por difração o modelo é caracterizado por problemas de espalhamento inverso para a equação da onda. Este problema é extremamente não-linear. No entanto, nestas notas utilizaremos métodos de aproximação para a obtenção de soluções o que torna o problema mais simples. Tais aproximações são conhecidas como aproximações de Born ou de Rytov.

Formulação do Problema

Consideremos a equação da onda

$$\frac{1}{v(\mathbf{x})} \frac{\partial^2 U}{\partial t^2} - \Delta U = 0, \quad \mathbf{x} \in \mathbb{R}^n, \quad t > 0, \quad (1.5.32)$$

para $n = 2, 3$ e com condições iniciais e de contorno apropriadas. As condições iniciais e de contorno serão enunciadas mais tarde. Em (1.5.32), $v(\mathbf{x})$ é a velocidade que é desconhecida. Neste manuscrito, por simplicidade, assumiremos que $v(\mathbf{x}) = c$ fora do objeto a ser imageado, onde c é uma constante positiva. Ainda, assumiremos que $U = 0$ para $t < 0$. Passamos para o domínio da frequência, ou seja, definimos

$$u(\mathbf{x}, \omega) = \int_{\mathbb{R}} U(\mathbf{x}, t) e^{-i\omega t} dt \quad (1.5.33)$$

como a Transformada de Fourier no tempo de U , temos que u satisfaz a seguinte equação de Helmholtz

$$\left(\Delta + \frac{\omega^2}{v(\mathbf{x})^2} \right) u(\mathbf{x}, \omega) = 0, \quad \mathbf{x} \in \mathbb{R}^n, \quad \omega \in \mathbb{R}, \quad (1.5.34)$$

$$\frac{\mathbf{x}}{|\mathbf{x}|} \cdot \nabla u(\mathbf{x}, \omega) - i \frac{\omega}{c} u(\mathbf{x}, \omega) = o(|\mathbf{x}|^{-(n-1)/2}).$$

A segunda equação em (1.5.34) é a condição de radiação. Esta condição garante que não existe energia vindo do infinito (existe somente ondas de radiação viajando para o infinito). A notação $o(t)$ significa que $o(t)/t \rightarrow 0$ quando $t \rightarrow \infty$. Deste modo, temos de (1.5.34) que o decaimento é maior que $|\mathbf{x}|^{-1/2}$ em dimensão dois e maior que $|\mathbf{x}|^{-1}$ em dimensão três.

Para simplificar o problema faremos a seguinte hipótese sobre v na equação (1.5.34)

$$\frac{1}{v^2(\mathbf{x})} = \frac{1}{c^2}(1 + a(\mathbf{x})), \quad k = \frac{\omega}{c}. \quad (1.5.35)$$

Substituimos na equação de Helmholtz e obtemos que

$$\begin{aligned} (\Delta + k^2) u(\mathbf{x}, \omega) &= -a(\mathbf{x})k^2 u(\mathbf{x}, \omega), \quad \mathbf{x} \in \mathbb{R}^n, \quad \omega \in \mathbb{R}, \\ \frac{\mathbf{x}}{|\mathbf{x}|} \cdot \nabla u(\mathbf{x}, \omega) - iku(\mathbf{x}, \omega) &= o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned} \quad (1.5.36)$$

Consideremos a onda plana

$$u_{in}(\mathbf{x}, \omega; \boldsymbol{\theta}) = e^{ik\boldsymbol{\theta} \cdot \mathbf{x}}, \quad (1.5.37)$$

onde $\boldsymbol{\theta}$ pertence a esfera unitária \mathbb{S}^{n-1} .

Notemos que u_{in} satisfaz a equação de Helmholtz homogênea

$$(\Delta + k^2) u_{in}(\mathbf{x}, \omega; \boldsymbol{\theta}) = 0. \quad (1.5.38)$$

No entanto, é fácil vermos que u_{in} em (1.5.37) não satisfaz a condição de radiação na equação (1.5.36). Na verdade, u_{in} satisfaz a condição de radiação na direção $\boldsymbol{\theta}$ mas não na direção $-\boldsymbol{\theta}$.

O **problema direto** consiste em: dada a onda plana vinda do infinito, encontrar a solução $u_s(\mathbf{x}, \omega)$ que modela a resposta do sistema. Assim, u_s não deve carregar informação do infinito e, portanto, deve satisfazer a condição de radiação na equação (1.5.36), que é ilustrada na Figura 1.8⁴.

⁴ A Figura 1.8 foi retirada de Wikipédia

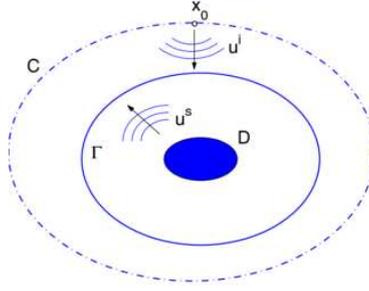


Figura 1.8: Ilustração de Tomografia por Difração.

O campo total $u = u_{in} + u_s$ deve satisfazer a equação de Helmholtz como no modelo físico. Em outras palavras, queremos resolver o seguinte problema de espalhamento

$$\begin{aligned} (\Delta + k^2) u_s(\mathbf{x}, \omega) &= -a(\mathbf{x})k^2(u_s + u_{in})(\mathbf{x}, \omega) & (1.5.39) \\ \frac{\mathbf{x}}{|\mathbf{x}|} \cdot \nabla u_s(\mathbf{x}, \omega) - ik u_s(\mathbf{x}, \omega) &= o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned}$$

Sob hipóteses adequadas sobre $a(x)$, a equação (1.5.39) possui uma única solução, [12].

O **problema inverso** consiste em reconstruir $a(x)$ por meio de medidas de u_s no infinito e em todas as possíveis direções x e para todas as possíveis ondas $\theta \in \mathbb{S}^{n-1}$. O problema inverso é também conhecido como *problema de espalhamento inverso*. Trataremos da chamada aproximação de Born como um método para obter soluções aproximadas deste problema inverso no Capítulo 4.

1.6 Exercícios

Exercício 1.2. Com base na figura 1.7 determine a matriz A associada. Para o vetor de intensidades p , determine uma solução f para o sistema (1.5.20), onde A é a matriz obtida analisando a Figura 1.7.

Exercício 1.3. Com base no exercício acima, compare a solução $f = (f_1, f_2, f_3, f_4)$ e a Figura 1.7. O que você tem a dizer?

Exercício 1.4. Use o método das características para mostrar que existe uma única solução para a EDP (1.5.21) com condição de fronteira (1.5.22).

Exercício 1.5. Verifique a equação (1.5.23).

Exercício 1.6. Mostre que $e^{Aa(\mathbf{x})}$ é um fator integrante para a EDP (1.5.22), no sentido que satisfaz (1.5.24).

Exercício 1.7. Mostre que a equação (1.5.33) satisfaz a equação de Helmholtz (1.5.34).

Exercício 1.8. Mostre que a mudança de variável (1.5.35) satisfaz a equação (1.5.36).

Exercício 1.9. Mostre que u_{in} satisfaz a equação de Helmholtz (1.5.36), mas não satisfaz a condição de radiação.

Exercício 1.10. Use (1.5.38) para verificar que $u = u_{in} + u_s$ satisfaz a equação (1.5.39).

Capítulo 2

Transformada de Radon

Neste capítulo estudaremos as propriedades da Transformada de Radon. Estas propriedades implicam na existência de uma fórmula inversa para o problema de Tomografia Computadorizada em casos especiais.

2.1 Propriedades da Transformada de Radon

Nesta seção provaremos alguns resultados interessantes sobre a Transformada de Radon. O primeiro resultado é a linearidade da Transformada de Radon.

Lema 2.1. *A Transformada de Radon \mathcal{R} é uma aplicação linear.*

Demonstração: Segue diretamente da linearidade da integral de linha.

■

O próximo teorema nos ensina que, para cada direção $\boldsymbol{\theta}$ dada, a Transformada de Radon fornece a integral de f sobre todas as retas paralelas a $\boldsymbol{\theta}$. Assim, obviamente, se as oscilações na direção $\boldsymbol{\theta}$ são perdidas, o mesmo não acontece nas direções ortogonais $\boldsymbol{\theta}^\perp$. A saber, a oscilação de f na direção $\boldsymbol{\theta}^\perp$ é da forma $\hat{f}(k\boldsymbol{\theta}^\perp)$ para $k \in \mathbb{R}$. Por isso, não é surpresa que este teorema seja utilizado nos métodos de reconstrução.

Teorema 2.1 (Teorema das Fatias de Fourier). *Seja f uma função suave. Então, para todo $\boldsymbol{\theta} \in \mathbb{S}^1$ temos que*

$$(\widehat{\mathcal{R}f}(\boldsymbol{\theta}, s))(k) = \hat{f}(k\boldsymbol{\theta}^\perp), \quad k \in \mathbb{R}. \quad (2.1.1)$$

Demonstração: Pelas definições da Transformada de Fourier em (1.4.7), a Transformada de Radon em (1.5.17) e o Teorema de Fubini [12], temos que

$$\widehat{\mathcal{R}f}(\boldsymbol{\theta}, s)(k) = \int_{\mathbb{R}} e^{-isk} \int_{\mathbb{R}} f(\mathbf{x} \cdot \boldsymbol{\theta}^\perp - s) dx ds = \hat{f}(k\boldsymbol{\theta}^\perp). \quad \blacksquare$$

A pergunta sobre injetividade da Transformada de Radon que fizemos na Subseção (1.5.1) é respondida no corolário:

Corolário 2.1. *Seja $f \in L^2(\mathbb{R}^n)$. Então, $\mathcal{R}f(\boldsymbol{\theta}, s) = 0$ se e só se $f(\boldsymbol{\theta}, s) = 0$.*

Demonstração: Como a Transformada de Fourier é um isomorfismo, temos que $\mathcal{R}f(\boldsymbol{\theta}, s) = 0$ se e só se $\widehat{\mathcal{R}f}(\boldsymbol{\theta}, s) = 0$. Pelo Teorema 2.1, se e só se $\hat{f}(k\boldsymbol{\theta}^\perp) = 0$. Novamente pelo isomorfismo da Transformada de Fourier, se e só se $f = 0$. Agora a injetividade segue do Lema 2.1. \blacksquare

Exercício 2.1. *Mostre que*

$$\mathcal{R}f(\boldsymbol{\theta}, s) = \mathcal{R}f(-\boldsymbol{\theta}, -s). \quad (2.1.2)$$

Notemos que, dado $m \in \mathbb{N}$,

$$\int_{\mathbb{R}} s^m \mathcal{R}f(\boldsymbol{\theta}, s) ds = \int_{\mathbb{R}} \int_{L(\boldsymbol{\theta}, s)} (\mathbf{x} \cdot \boldsymbol{\theta})^m f(\mathbf{x}) dS ds. \quad (2.1.3)$$

Porém, quando calculamos $(\mathbf{x} \cdot \boldsymbol{\theta})^m$ obtemos um polinômio homogêneo de grau m em $\boldsymbol{\theta}$, com coeficientes que também são polinômios homogêneos de grau m na variável \mathbf{x} . Já, o resultado da integral (2.1.3) é um polinômio de grau m em $\boldsymbol{\theta}$.

Com isso obtemos que: $F(\boldsymbol{\theta}, s)$ está na imagem de \mathcal{R} se satisfizer:

$$F(\boldsymbol{\theta}, s) = F(-\boldsymbol{\theta}, -s) \quad (2.1.4)$$

$$\int_{\mathbb{R}} s^m F(\boldsymbol{\theta}, s) ds = P_m(\boldsymbol{\theta}) \quad (2.1.5)$$

onde P_m é um polinômio homogêneo de grau m em $\boldsymbol{\theta}$.

Exercício 2.2. *Seja*

$$\begin{aligned} P : \mathbb{S}^{n-1} \times (0, \infty) &\longrightarrow \mathbb{R}^n \\ (\boldsymbol{\theta}, s) &\longmapsto s\boldsymbol{\theta} = \mathbf{x}. \end{aligned}$$

i) *Mostre que P é um isomorfismo e $P^{-1}(\mathbf{x}) = (\mathbf{x}/\|\mathbf{x}\|, \|\mathbf{x}\|)$ é a inversa de P .*

ii) *Demonstre que se $f \in C^\infty(\mathbb{R}^n)$, então $f \circ P \in C^\infty(\mathbb{S}^{n-1} \times [0, \infty))$.*

iii) *Seja $F \in C^\infty(\mathbb{S}^{n-1} \times \mathbb{R})$. Mostre que*

$$F \circ P^{-1} = F(\mathbf{x}/\|\mathbf{x}\|, \|\mathbf{x}\|) \in C^\infty(\mathbb{R}^n - \{0\}).$$

iv) *Seja $F \in C^\infty(\mathbb{S}^{n-1} \times \mathbb{R})$. Demonstre que $f = F \circ P^{-1} \in C^\infty(\mathbb{R}^n)$.*

Ainda:

$$\frac{\partial^m}{\partial s^m} F(\boldsymbol{\theta}, s) \Big|_{s=0} = P_m(\boldsymbol{\theta}),$$

é um polinômio homogêneo de grau m , para todo $m \in \mathbb{N}$.

Definimos o conjunto de funções suaves

$$\mathcal{S}(\mathbb{S}^{n-1} \times \mathbb{R}) := \left\{ F \in C^\infty(\mathbb{S}^{n-1} \times \mathbb{R}) : |s|^k \left| \frac{\partial^l}{\partial s^l} \frac{\partial^{m_1}}{\partial w_1^{m_1}} \cdots \frac{\partial^{m_n}}{\partial w_n^{m_n}} F(w, s) \right| \leq C(k, l, m_1, \dots, m_n) \right\}.$$

O conjunto \mathcal{S} é chamado de espaço de Schwartz.

Exercício 2.3. *Mostre que $\mathcal{S}(\mathbb{R}^n)$ é um subespaço vetorial de $L^2(\mathbb{R}^n)$.*

Com os resultados acima, estamos prontos para mostrar a sobrejetividade da Transformada de Radon.

Teorema 2.2. *A Transformada de Radon é uma aplicação bijetora*

$$\mathcal{R} : \mathcal{S}(\mathbb{R}^n) \longrightarrow M := \{ \mathcal{S}(\mathbb{S}^{n-1} \times \mathbb{R}) : F \text{ satisfaz (2.1.4) e (2.1.5)} \}$$

Demonstração: A injetividade foi provada no Lema 2.1. Seja $F(\boldsymbol{\theta}, s) \in M$. Como

$$\int_{\mathbb{R}} s^m F(\boldsymbol{\theta}, s) ds = \widehat{s^m F}(\boldsymbol{\theta}, 0) = i^m \frac{\partial^m}{\partial s^m} \hat{F}(\boldsymbol{\theta}, 0)$$

é, por hipótese, um polinômio homogêneo de grau m . Seja $\xi = \lambda\boldsymbol{\theta}$. Pelo Exercício 2.2, $f(\xi) = \widehat{F}(\xi/\|\xi\|, \|\xi\|) \in C^\infty(\mathbb{R}^n)$. Como $F \in M$ e a Transformada de Fourier é um isomorfismo em \mathcal{S} obtemos que $f \in \mathcal{S}(\mathbb{R}^n)$. De (2.1.1), deduzimos que $F = \mathcal{R}\mathcal{F}^{-1}f$. ■

2.2 Fórmula de Inversão para Tomografia Computadorizada

Como a Transformada de Radon é uma bijeção, cabe a pergunta: Será que é possível encontrar uma forma analítica para \mathcal{R}^{-1} ? A idéia é utilizar a relação entre a Transformada de Radon e a Transformada de Fourier provada no Teorema 2.1.

Exercício 2.4. *Mostre que*

$$\mathcal{F}^{-1}(\lambda^{n-1}\widehat{\mathcal{R}f}(\boldsymbol{\theta}, \lambda)) = 2\pi \frac{1}{i^{n-1}} \frac{\partial^{n-1}}{\partial s^{n-1}} \mathcal{R}f(\boldsymbol{\theta}, s).$$

Para encontrarmos uma expressão analítica para \mathcal{R}^{-1} temos que considerar o caso n par e n ímpar. A demonstração de cada caso difere muito (veja [35]).

Teorema 2.3. *Seja $n > 1$ ímpar. Então, para $f \in \mathcal{S}(\mathbb{R}^n)$*

$$f(\mathbf{x}) = 2^{-n}(i\pi)^{1-n} \int_{\mathbb{S}^{n-1}} \frac{\partial^{n-1}}{\partial s^{n-1}} \mathcal{R}f(\boldsymbol{\theta}, \mathbf{x} \cdot \boldsymbol{\theta}) dS, \quad (2.2.6)$$

onde $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta)) \in \mathbb{S}^1$ e $\theta \in [0, 2\pi)$.

Demonstração: Pela Transformada de Fourier Inversa, coordenadas polares e o fato que $(n-1)$ é par, temos que

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^n} \int_{\mathbb{S}^{n-1}} \int_0^\infty e^{i\mathbf{x} \cdot \lambda\boldsymbol{\theta}} \widehat{f}(\lambda\boldsymbol{\theta}^\perp) \lambda^{n-1} d\lambda dS_\theta \\ &= \frac{1}{2(2\pi)^n} \int_{\mathbb{S}^{n-1}} \int_{-\infty}^\infty e^{i\mathbf{x} \cdot \lambda\boldsymbol{\theta}} \widehat{f}(\lambda\boldsymbol{\theta}^\perp) \lambda^{n-1} d\lambda dS_\theta. \end{aligned}$$

Por (2.1.1) e pelo exercício 2.4 concluímos a demonstração. ■

A fórmula (2.2.6) possui propriedades interessantes que devem ser destacadas. Para obtermos $f(\mathbf{x}_0)$ é necessário conhecer os valores de $\mathcal{R}f(\theta, s)$ para $s = \boldsymbol{\theta} \cdot \mathbf{x}_0$. Ou seja, não precisamos conhecer as integrais de f ao longo de todas as retas $L(\theta, s)$. Basta obtermos informações sobre os que distam $\boldsymbol{\theta} \cdot \mathbf{x}_0$ da origem.

Notemos que a demonstração do Teorema 2.3 não é verdadeira para o caso de n ser par. Neste caso, é preciso introduzir a Transformada de Hilbert, veja [31, 35].

Definição 2.1. [Transformada de Hilbert] Sejam $f \in \mathcal{S}(\mathbb{R})$. Então

$$\mathcal{H}(f(s)) = (\pi)^{-1} v.p. \int_{\mathbb{R}} \frac{f(t)}{s-t} dt, \quad (2.2.7)$$

onde *v.p.* significa o valor principal de Cauchy, que neste contexto é dado por

$$v.p. \int_{\mathbb{R}} \frac{f(t)}{s-t} dt = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{s-\varepsilon} \frac{f(t)}{s-t} dt + \int_{s+\varepsilon}^{\infty} \frac{f(t)}{s-t} dt.$$

Teorema 2.4. Seja $n > 1$ par. Então

$$f(\mathbf{x}) = 2^{-n} (i\pi)^{1-n} \int_{\mathbb{S}^{n-1}} \left(\frac{\partial^{n-1}}{\partial s^{n-1}} \mathcal{H}\mathcal{R}f \right) (\boldsymbol{\theta}, \mathbf{x} \cdot \boldsymbol{\theta}) dS. \quad (2.2.8)$$

Demonstração: Veja [31, 35] para detalhes. ■

O que podemos aprender de (2.2.6) e de (2.2.8) é que ambas as fórmulas envolvem derivadas de $\mathcal{R}f$. Pelo que vimos na Seção 1.3, isto é uma indicação de um problema mal-posto. Claro que, por exemplo, na equação (2.2.8), depois de derivarmos, fazemos uma operação suavizante novamente, a saber, integramos. Por outro lado, o núcleo em (2.2.8) é singular e, portanto, não anula totalmente a instabilidade introduzida pela diferenciação. Assim, métodos mais adequados que a invertibilidade direta da Transformada de Radon devem ser considerados.

O próximo resultado é importante pois relaciona o suporte da Transformada de Radon com o suporte da função f . Também é útil para deduzir uma fórmula explícita para \mathcal{R}^{-1} para o caso especial $n = 2$.

Teorema 2.5. *Seja $f \in C^\infty(\mathbb{R}^n)$ tal que $f(\mathbf{x}) \rightarrow 0$ mais rapidamente que qualquer polinômio. Se $\mathcal{R}f(\theta, s) = 0$ para $|s| > \rho$, então $f(\mathbf{x}) = 0$ para $\|\mathbf{x}\| > \rho$.*

Demonstração: A demonstração foge ao escopo destas notas. Veja [35] para detalhes. ■

2.3 Exercícios

Recomendamos ao leitor [31, 35] como referências para ajudar a resolver os exercícios abaixo.

Exercício 2.5. *Mostre que $\mathcal{H}(f(s)) = (\pi s)^{-1} * f(s)$, onde $*$ é o operador de convolução.*

Exercício 2.6. *Seja $\text{sign}(\lambda) = 1$ para $\lambda > 0$, $\text{sign}(\lambda) = -1$ para $\lambda < 0$ e $\text{sign}(0) = 0$. Mostre que $\widehat{\mathcal{H}(f(s))}(\lambda) = -i \text{sign}(\lambda) \hat{f}(\lambda)$.*

Exercício 2.7. *Mostre que $\mathcal{H}(\mathcal{H}f(s)) = -f(s)$.*

Exercício 2.8. *Demonstre que*

$$\mathcal{R}f(\theta, s) = \int_{\Sigma} \int_0^{\infty} f(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}) t^{n-2} dt d\boldsymbol{\theta},$$

onde $\Sigma = \{\boldsymbol{\theta} \in \mathbb{S}^{n-1} : \mathbf{x} \cdot \boldsymbol{\theta} = s\}$.

Exercício 2.9. *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita uma função radial, se existir uma função $g : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(\mathbf{x}) = g(\|\mathbf{x}\|)$, $\forall \mathbf{x} \in \mathbb{R}^n$. Mostre que se f é radial, então*

$$\mathcal{R}f(\theta, s) = 2 \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2)} \int_0^{\infty} g(\sqrt{s^2 + t^2}) t^{n-2} dt.$$

Agora, consideramos Ω como um círculo de raio ρ . Ainda suponhamos que f é simétrica com respeito a origem. Assim, toda a informação que necessitamos saber está na direção $\boldsymbol{\theta}_0 = (0, \pm 1)$.

Deste modo, podemos assumir que

$$f(\boldsymbol{\theta}, s) = f(s), \quad 0 < s \leq \rho \quad \boldsymbol{\theta} \in \mathbb{S}^1.$$

Exercício 2.10. *Sejam as hipóteses do Teorema 2.5 satisfeitas e $0 < s \leq \rho$. Use o Exercício 2.9 para mostrar que $\mathcal{R}f(\boldsymbol{\theta}_0, s)$ satisfaz uma equação integral de Abel de primeira espécie*

$$\mathcal{R}f(\boldsymbol{\theta}_0, s) = \pi \int_s^\rho \frac{rf(r)}{\sqrt{r^2 - s^2}} dr. \quad (2.3.9)$$

Exercício 2.11. *Supondo que $\mathcal{R}f(\boldsymbol{\theta}_0, \rho) = 0$, prove que*

$$f(s) = -\pi^{-1} \int_r^\rho \frac{d/ds(\mathcal{R}f(\boldsymbol{\theta}_0, s))}{\sqrt{s^2 - r^2}} ds. \quad (2.3.10)$$

Capítulo 3

Transformada de Radon Atenuada

Nestas notas usaremos o famoso problema de Riemann-Hilbert para obtermos uma fórmula inversa para a Transformada de Radon Atenuada obtida como resultado do modelo de SPECT na Subseção 1.5.

3.1 O Problema de Riemann-Hilbert

O problema de Riemann-Hilbert possui uma quantidade muito grande de aplicações em várias áreas, em particular, em variáveis complexas. Neste manuscrito necessitaremos de uma versão simplificada do teorema de Riemann-Hilbert. Iniciaremos esta seção definindo o nosso material de estudo.

Seja γ a curva no plano que denota a circunferência de raio um, que aqui parametrizaremos por $\lambda \in \mathbb{C} : |\lambda| = 1$ e sejam $D^+ := \{\lambda \in \mathbb{C} : |\lambda| < 1\}$ e $D^- := \{\lambda \in \mathbb{C} : |\lambda| > 1\}$. Suponhamos que a orientação de γ foi escolhida de forma que a região D^+ fique a esquerda ao percorrermos γ .

Seja $\phi(\lambda)$ uma função suave definida em $D^+ \cup D^-$. Denotaremos por $\phi^+(t)$ e $\phi^-(t)$ o traço de ϕ em γ em D^+ e D^- , respectivamente.

Em outras palavras,

$$\phi^+(t) = \lim_{\varepsilon \rightarrow 0^+} \phi((1 - \varepsilon)t) \quad \phi^-(t) = \lim_{\varepsilon \rightarrow 0^+} \phi((1 + \varepsilon)t). \quad (3.1.1)$$

Definimos o salto de função ϕ sobre a curva γ por

$$\psi(t) = \phi^+(t) - \phi^-(t). \quad (3.1.2)$$

Suponhamos que $\psi(t)$ seja uma função suave em γ , **o problema de Riemann-Hilbert** é: Encontrar uma função $\phi(\lambda)$ definida em $D^+ \cup D^-$ tal que

1. $\phi(\lambda)$ seja analítica em D^+ e em D^- .
2. $\lambda\phi(\lambda)$ seja limitada para $|\lambda| \rightarrow \infty$ em D^- .
3. O salto de ϕ em γ seja dado pela função $\psi(t)$ em (3.1.2).

O problema de Riemann-Hilbert possui uma única solução que é dada pela fórmula de Cauchy, [1],

$$\phi(\lambda) = \frac{1}{2i\pi} \int_{\gamma} \frac{\psi(t)}{t - \lambda} dt, \quad \lambda \in D^+ \cup D^-. \quad (3.1.3)$$

Na próxima seção demonstraremos que a Transformada de Radon Atenuada está intimamente relacionada com o problema de Riemann-Hilbert e que podemos usar a fórmula de Cauchy para obter uma inversa para (1.5.27).

3.2 Fórmula de Inversão para Tomografia por Emissão

Para obtermos uma fórmula para a inversa da Transformada de Radon Atenuada, o primeiro passo é estender a equação de transporte (1.5.21) para o corpo dos complexos. Para tal, seja $\lambda = e^{i\theta}$, $\theta \in (0, 2\pi)$. Notemos que λ desta forma toma valores sobre o círculo unitário γ

para $\theta \in (0, 2\pi)$. Por simplicidade, denotaremos $\mathbf{x} = (x, y) \in \mathbb{R}^2$, então a equação do transporte (1.5.21) pode ser reescrita como

$$\left(\frac{(\lambda + \lambda^{-1})}{2} \frac{\partial}{\partial x} + \frac{(\lambda - \lambda^{-1})}{2i} \frac{\partial}{\partial y} + a(\mathbf{x}) \right) u(\mathbf{x}, \lambda) = f(\mathbf{x}). \quad (3.2.4)$$

Identificamos $\mathbf{x} \in \mathbb{R}^2$ por $z = x + iy \in \mathbb{C}$ e fazemos

$$2 \frac{\partial(\cdot)}{\partial z} = \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) (\cdot) \quad e \quad 2 \frac{\partial(\cdot)}{\partial \bar{z}} = \overline{\frac{\partial(\cdot)}{\partial z}},$$

onde \bar{z} é o conjugado de $z \in \mathbb{C}$, e podemos reescrever a equação (3.2.4) como

$$\left(\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} + a(z) \right) u(z, \lambda) = f(z), \quad z \in \mathbb{C}. \quad (3.2.5)$$

É claro que a mesma condição de que nenhuma informação vem do infinito como em (1.5.22) deve ser considerada.

Obteremos a fórmula de inversão em três etapas:

1. **Passo 1:** Verificamos que $u(z, \lambda)$ é analítica em $D^+ \cup D^-$ e que $\lambda u(z, \lambda)$ é limitada quando $|\lambda| \rightarrow \infty$.
2. **Passo 2:** Verificamos que os saltos de u solução de (3.2.4) sob a curva $\lambda = e^{i\theta}$, que denotaremos por $\psi(\mathbf{x}, \theta) = u(\mathbf{x}, \theta)^+ - u(\mathbf{x}, \theta)^-$, para cada \mathbf{x} , podem ser obtidos a partir dos dados $\mathcal{R}_a f(s, \theta)$.
3. **Passo 3:** Resolvemos o problema de Riemann-Hilbert usando (3.1.3) e avaliamos a equação (3.2.5) em $\lambda = 0$ para obtermos $f(x, y) = f(z)$.

Passo 1: Também é conhecido como o **problema $\bar{\partial}$** .

Observemos que, se $a = 0$ na equação (3.2.5), então ela satisfaz

$$\left(\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} \right) G(z, \lambda) = \delta(z), \quad \lambda \in \mathbb{C} - \gamma \cup \{0\}. \quad (3.2.6)$$

onde $\delta(z)$ representa a distribuição de Dirac e $G(z, \lambda)$ é a solução fundamental.

Lema 3.1. *A equação (3.2.6) possui uma única solução dada pela função de Green*

$$G(z, \lambda) = \frac{\text{sing}(|\lambda| - 1)}{\pi(\lambda\bar{z} - \lambda^{-1}z)}, \quad \lambda \notin \gamma \cup \{0\}. \quad (3.2.7)$$

Demonstração: Deixamos como exercício verificar que G em (3.2.7) satisfaz a equação (3.2.6). ■

Definimos a seguinte mudança de variável $\zeta = \lambda\bar{z} - \lambda^{-1}z$. Por enquanto, assumimos que $|\lambda| > 1$.

Exercício 3.1. *O Jacobiano da transformação é $J(\lambda) := |\lambda|^2 - |\lambda|^{-2}$. Assim,*

$$\left(\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} \right) = -J(\lambda) \frac{\partial}{\partial \zeta}.$$

Denotamos por $\tilde{G}(\zeta) = G(z)$, então temos que

$$\frac{\partial}{\partial \zeta} \tilde{G}(\zeta) = -\frac{1}{J(\zeta)} \delta(z(\zeta)) = -\delta(\zeta).$$

Portanto, $-\tilde{G}(\zeta)$ é a solução do operador $\bar{\partial} := \frac{\partial}{\partial \zeta}$.

Afirmção: $\frac{\partial}{\partial \zeta} \frac{1}{\zeta} = \pi \delta(\zeta)$. De fato, seja $\varphi(z) \in C_0^\infty(\mathbb{C})$ e denotemos por $d\mu(\zeta)$ a medida de Lebesgue $dx dy$ de $\mathbb{C} \sim \mathbb{R}^2$. Então, pela a fórmula de Green para variáveis complexas

$$\begin{aligned} \int_{\mathbb{C}} \varphi(\zeta) \frac{\partial}{\partial \zeta} \frac{1}{\zeta} d\mu(\zeta) &= - \int_{\mathbb{C}} \frac{\partial}{\partial \bar{\zeta}} \frac{\varphi(\zeta)}{\zeta} d\mu(\zeta) = - \lim_{\varepsilon \rightarrow 0} \int_{\tilde{\mathbb{C}}} \frac{\partial}{\partial \bar{\zeta}} \frac{\varphi(\zeta)}{\zeta} d\mu(\zeta) \\ &= - \lim_{\varepsilon \rightarrow 0} \int_{\tilde{\mathbb{C}}} \frac{\partial \zeta^{-1}}{\partial \bar{\zeta}} \varphi(\zeta) d\mu(\zeta) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2i} \int_{|\zeta|=\varepsilon} \frac{\varphi}{\zeta} d\mu(\zeta) \\ &= \frac{1}{2i} 2i\pi \varphi(0) = \int_{\mathbb{R}^2} \varphi(\zeta) \pi \delta(\zeta) d\mu(\zeta), \end{aligned}$$

onde $\tilde{\mathbb{C}} := \mathbb{C} - \{|\zeta| < \varepsilon\}$. Isto demonstra a afirmação.

Da afirmação segue que $\tilde{G}(\zeta) = (-\pi\zeta)^{-1}$. Portanto, $G(z) = (\pi(\lambda\bar{z} - \lambda^{-1}z))^{-1}$. E assim, segue (3.2.7) para $|\lambda| > 1$.

Consideramos agora $|\lambda| < 1$. Pelo exercícios 3.7 temos que o Jacobiano é dado por $-J(\zeta)$. Portanto,

$$\frac{\partial}{\partial \bar{\zeta}} \tilde{G}(\zeta) = \frac{1}{J(\zeta)} \delta(z(\zeta)) = \delta(\zeta).$$

Agora, pelos mesmos passos da afirmação acima obtemos (3.2.7) para $|\lambda| < 1$. ■

Lema 3.2. *Consideramos a equação*

$$\left(\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} \right) h(z, \lambda) = a(z), \quad \lambda \in \mathbb{C} - (\gamma \cup \{0\}).$$

com $|h(z, \lambda)| \rightarrow 0$ quando $z \rightarrow \infty$.

A solução é dada por

$$h(z, \lambda) = \int_{\mathbb{R}^2} G(z - \zeta, \lambda) a(\zeta) d\mu(\zeta). \quad (3.2.8)$$

onde $G(z, \lambda)$ é a solução fundamental de (3.2.6).

Demonstração: Segue das propriedades da função de Green, [12].

Lema 3.3. *Seja $e^{h(z, \lambda)}$ onde h é dado por (3.2.8). Então,*

$$\left(\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} \right) (e^{h(z, \lambda)} u(z, \lambda)) = e^{h(z, \lambda)} f(z), \quad \lambda \in \mathbb{C} - (\gamma \cup \{0\}).$$

Demonstração: A demonstração é feita por verificação e assim deixamos para o leitor. ■

Teorema 3.1. *A solução de (3.2.4) é dada por*

$$u(z, \lambda) = e^{-h(z, \lambda)} \int_{\mathbb{R}^2} G(z - \zeta, \lambda) e^{h(\zeta, \lambda)} f(\zeta) d\mu(\zeta), \quad (3.2.9)$$

para toda f suficientemente suave e $G(z, \lambda)$ dada por (3.2.7).

Demonstração: Segue do Lema 3.3 e por integração. ■

Lema 3.4. *A função de Green $G(z, \lambda)$ é analítica em D^+ , em D^- e em $z = 0$.*

Demonstração: Veja Exercício 3.11. ■

Lema 3.5. *Suponhamos que $a(z)$ e $f(z)$ sejam suficientemente suaves. Então, a função $h(z, \lambda)$ em (3.2.8) é analítica em D^+ , em D^- e em $z = 0$.*

Demonstração: Veja Exercício 3.12. ■

Teorema 3.2. *A solução $u(z, \lambda)$ de (3.2.9) satisfaz o **Passo 1**.*

Demonstração: Que u é analítica segue diretamente da analiticidade de G e h verificadas nos lemas acima. Usando (3.2.9), é fácil de verificar que $\lambda u(z, \lambda)$ é limitado em D^- . ■

O teorema acima implica que a solução da equação do transporte $u(z, \lambda)$, estendida ao campo complexo, é um bom candidato a satisfazer o problema de Riemann-Hilbert.

Passo 2: O que queremos verificar agora é o que acontece com $u(\lambda)$ a medida que λ se aproxima de γ por cima (em D^-) e por baixo (em D^+). Esta condição é conhecida como a **condição de salto**.

Denotemos por $\lambda = re^{i\theta}$. Quando $r \rightarrow \pm 1$, λ se aproxima de D^\pm . No sentido das distribuições, [12], temos o seguinte limite

Lema 3.6. *Seja $r - 1 \rightarrow \pm 0$, a função de Green $G(x, \lambda)$ satisfaz o seguinte limite*

$$G_\pm(\mathbf{x}, \theta) = \frac{\pm 1}{-2\pi i(-\boldsymbol{\theta}^\perp \cdot \mathbf{x} \pm i0 \operatorname{sign}(\boldsymbol{\theta} \cdot \mathbf{x}))},$$

onde ± 0 denota o limite $\pm \varepsilon \rightarrow 0$.

Demonstração: Seja $|\lambda| > 1$, isto é, $r = 1 + \varepsilon$ com $\varepsilon > 0$. Então

$$\begin{aligned} \pi G^-(z, r e^{i\theta}) &= \frac{1}{r e^{i\theta} \bar{z} - e^{-i\theta} z / r} = \frac{e^{-i\theta}}{(1 + \varepsilon) e^{2i\theta} \bar{z} - z(1 - \varepsilon) + o(\varepsilon)} \\ &= \frac{1}{-i \operatorname{Im}(e^{-i\theta} z) + \varepsilon \operatorname{Re}(e^{-i\theta} z) + o(\varepsilon)} = \frac{-i/2}{\mathbf{x} \cdot (\boldsymbol{\theta}^\perp + i\varepsilon \boldsymbol{\theta}) + o(\varepsilon)}. \end{aligned}$$

Fazemos $\varepsilon \rightarrow 0$ e obtemos o resultado. Idem para G^+ . ■

A derivação do resultado passa pelo seguinte exercício, [12], conhecido como fórmula de Plemelj:

Exercício 3.2. *Mostre que para qualquer $f \in C_0^\infty(\mathbb{R})$ temos que*

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}} \frac{f(x)}{ix + \varepsilon} dx = -i \operatorname{v.p.} \int_{\mathbb{R}} \frac{f(x)}{x} dx + \operatorname{sign}(\varepsilon) \pi f(0),$$

onde *v.p.* representa o valor principal de Cauchy.

Lema 3.7. *Seja $\lambda = e^{i\theta}$. Para toda função suave ψ temos que*

$$\int_{\mathbb{R}^2} G^\pm(\mathbf{x} - \mathbf{y}, \theta) \psi(\mathbf{y}) d\mathbf{y} = \pm \frac{1}{2i} (\mathcal{H}R\psi)(\mathbf{x} \cdot \boldsymbol{\theta}^\perp) + (\mathcal{A}\psi)(\mathbf{x}),$$

onde \mathcal{H} é a Transformada de Hilbert como na Definição 2.1 e $G^\pm(z, \theta)$ é a função de Green como no Lema 3.6.

Demonstração: Denotemos por $\mathbf{x} = \sigma \boldsymbol{\theta}^\perp + \tau \boldsymbol{\theta}$ e $\mathbf{y} = s \boldsymbol{\theta}^\perp + t \boldsymbol{\theta}$ para $\sigma, \tau, s, t \in \mathbb{R}$. Então,

$$\begin{aligned} \int_{\mathbb{R}^2} G^+(\mathbf{y}) \psi(\mathbf{x} - \mathbf{y}) d\mathbf{y} &= \frac{1}{2\pi} \int_{\mathbb{R}^2} \frac{\psi((\sigma - s) \boldsymbol{\theta}^\perp + (\tau - t) \boldsymbol{\theta})}{is + i0 \operatorname{sign}(t)} ds dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \operatorname{v.p.} \left(\int_{\mathbb{R}} \frac{\psi((\sigma - s) \boldsymbol{\theta}^\perp + (\tau - t) \boldsymbol{\theta})}{is} ds \right) dt \\ &\quad + \frac{1}{2} \int_{\mathbb{R}} \operatorname{sign}(t) \psi(\mathbf{x} - t \boldsymbol{\theta}) dt = \frac{1}{2i} (\mathcal{H}R\psi)(\mathbf{x} \cdot \boldsymbol{\theta}^\perp) + (\mathcal{A}\psi)(\mathbf{x}). \end{aligned}$$

De forma análoga, provamos o limite G^- em D^- . ■

Segue disto o seguinte corolário

Corolário 3.1. *A função $h(z, \lambda)$ definida em (3.2.8) satisfaz os seguintes limites*

$$h^\pm(z, \theta) = \pm \frac{1}{2\pi} (\mathcal{H}\mathcal{R}a)(\mathbf{x} \cdot \boldsymbol{\theta}^\perp) + (\mathcal{A}a)(\mathbf{x}).$$

Notemos que \mathcal{R} e \mathcal{A} envolvem integração somente na direção $\boldsymbol{\theta}$ e, assim,

$$\begin{aligned} \mathcal{R}[u(\mathbf{x})v(\mathbf{x} \cdot \boldsymbol{\theta}^\perp)](s) &= v(s)\mathcal{R}[u](s), \\ \mathcal{A}[u(\mathbf{x})v(\mathbf{x} \cdot \boldsymbol{\theta}^\perp)](x) &= v(\mathbf{x} \cdot \boldsymbol{\theta}^\perp)\mathcal{A}[u](\mathbf{x}). \end{aligned}$$

Da observação acima e dos resultados anteriores temos que

Teorema 3.3. *Os limites da função $u(z, \lambda)$ são dados por*

$$\begin{aligned} u^\pm(\mathbf{x}, \theta) &= e^{(\pm \frac{i}{2}(\mathcal{H}\mathcal{R}a)(s) - \mathcal{A}a)} \left[\pm \frac{i}{2} \mathcal{H} \left(e^{\pm \frac{1}{2i}(\mathcal{H}\mathcal{R}a)(s)} \mathcal{R}(e^{\mathcal{A}a}(f)) \right) \right] (s) \\ &\quad + e^{-\mathcal{A}a} \mathcal{A}(e^{\mathcal{A}a}(f))(\mathbf{x}). \end{aligned} \quad (3.2.10)$$

Demonstração: Basta unirmos os limites no Lema 3.7, Corolário 3.1 a equação (3.2.9). ■

Como observamos anteriormente, a quantidade $\mathcal{R}(e^{\mathcal{A}a}f) = \mathcal{R}_a f(s)$ são as medidas (veja (1.5.27)). No entanto, como mostra o Teorema 3.3 u^+ e u^- não dependem somente dos dados. Estes dependem ainda de $\mathcal{A}(e^{\mathcal{A}a}(f))(\mathbf{x})$. Mas, $u^+ - u^-$ depende somente dos dados.

Definimos

$$\varphi(\mathbf{x}, \theta) = (u^+ - u^-)(\mathbf{x}, \boldsymbol{\theta}) \quad (3.2.11)$$

e os operadores

$$\begin{aligned} \mathcal{H}_a &= C_c \mathcal{H} C_c + C_s \mathcal{H} C_s \\ C_c g(s, \theta) &= g(s, \theta) \cos(\mathcal{H}\mathcal{R}a(s)/2), \\ C_s g(s, \theta) &= g(s, \theta) \sin(\mathcal{H}\mathcal{R}a(s)/2). \end{aligned}$$

Consideramos o operador adjunto de \mathcal{R}_a definido formalmente por

$$\mathcal{R}_{a,\theta}^* = e^{\mathcal{A}a(\mathbf{x})} g(\mathbf{x} \cdot \boldsymbol{\theta}^\perp).$$

Do Teorema 3.3 temos que

$$i\varphi(\mathbf{x}, \theta) = \mathcal{R}_{-a, \theta}^* \mathcal{H}_a \mathcal{R}_a(\theta). \quad (3.2.12)$$

Observamos que $i\varphi(\mathbf{x}, \theta)$ é uma função real da forma $e^{-Aa(x)} M(\mathbf{x} \cdot \boldsymbol{\theta}^\perp, \theta)$ para algum operador M . Ainda, φ satisfaz

$$\boldsymbol{\theta} \cdot \nabla \varphi(\mathbf{x}, \theta) + a\varphi(\mathbf{x}, \theta) = 0. \quad (3.2.13)$$

Isso refletirá na fórmula de inversão.

Passo 3: Agora estamos prontos para obtermos uma Fórmula de Inversão. Para tal, lembremos que os Passos 1 e 2 são satisfeitos por $u(z, \lambda)$. Portanto, u satisfaz a fórmula de Cauchy

$$u(z, \lambda) = \frac{1}{2i\pi} \int_\gamma \frac{\varphi(\mathbf{x}, t)}{t - \lambda} dt \quad \lambda \in D^+ \cup D^-, \quad (3.2.14)$$

onde identificamos $\varphi(\mathbf{x}, t)$ com $\varphi(\mathbf{x}, \theta)$ usando $t = e^{i\theta}$.

Exercício 3.3. *Mostre que, numa vizinhança de $\lambda = 0$ vale a identidade*

$$f(\mathbf{x}) = \lim_{\lambda \rightarrow 0} \lambda^{-1} \frac{\partial}{\partial \bar{z}} u(\mathbf{x}, \lambda).$$

Exercício 3.4. *Verifique que $u(\mathbf{x}, \lambda) = \mathcal{O}(\lambda)$ em D^+ . Verifique ainda que $a(\mathbf{x})u(\mathbf{x}, \lambda) \rightarrow 0$ quando $\lambda \rightarrow 0$.*

Exercício 3.5. *Mostre que*

$$u(\mathbf{x}, \lambda) = \frac{1}{2i\pi} \int_\gamma \frac{\varphi(\mathbf{x}, t)}{t} dt + \frac{\lambda}{2i\pi} \int_\gamma \frac{\varphi(z, t)}{t^2} dt + \mathcal{O}(\lambda^2).$$

Sugestão: *Utilize a equação (3.2.14) e a expansão de $1/(t - \lambda)$ em séries de potências.*

Dos exercícios acima segue a fórmula de reconstrução

Teorema 3.4. *Se φ depende dos dados $\mathcal{R}_a f(\theta, s)$, vale que*

$$f(\mathbf{x}) = \frac{1}{2i\pi} \int_{\gamma} \frac{\partial \varphi}{\partial \bar{z}}(\mathbf{x}, t) \frac{dt}{t^2}, \quad (3.2.15)$$

com a seguinte condição de compatibilidade

$$\frac{1}{2i\pi} \int_{\gamma} \frac{\varphi(\mathbf{x}, t)}{t} dt = 0.$$

Demonstração: Segue diretamente dos exercícios acima. ■

A fórmula de reconstrução acima não parece muito amistosa. O próximo resultado ajudará a melhorar o entendimento da mesma.

Corolário 3.2. *Seja $g_a(\theta, s) = \mathcal{R}_a f(\theta, s)$ as medidas no modelo de SPECT e φ satisfazendo (3.2.13). Então,*

$$f(\mathbf{x}) = \frac{1}{4\pi} \int_0^{2\pi} \boldsymbol{\theta}^\perp \cdot \nabla(\mathcal{R}_{\theta, -a}^* \mathcal{H}_a g)(\mathbf{x}, \theta) d\theta. \quad (3.2.16)$$

Demonstração: Denotemos por $t = e^{i\theta}$. Então, $dt = ie^{i\theta} d\theta$. Assim, é fácil verificarmos que

$$\begin{aligned} \frac{4\pi}{2i\pi} \int_{\gamma} \frac{\partial \varphi}{\partial \bar{z}}(\mathbf{x}, t) \frac{dt}{t^2} &= \int_0^{2\pi} (\boldsymbol{\theta}^\perp \cdot \nabla(i\varphi)(\mathbf{x}, \theta) + \boldsymbol{\theta} \cdot \nabla\varphi(\mathbf{x}, \theta)) d\theta \\ \frac{1}{2i\pi} \int_{\gamma} \frac{\varphi(\mathbf{x}, t)}{t} dt &= \frac{1}{2\pi} \int_0^{2\pi} \varphi(\mathbf{x}, \theta) d\theta \end{aligned}$$

De (3.2.13) e (3.2.15), obtemos que

$$f(\mathbf{x}) = \frac{1}{4\pi} \int_0^{2\pi} \boldsymbol{\theta} \cdot \nabla\varphi(\mathbf{x}, \theta) d\theta.$$

Finalmente, usando (3.2.12) o corolário está provado. ■

3.3 Exercícios

Uma boa fonte de referência para os exercícios deste capítulo é [2].

Exercício 3.6. *Mostre que $|G(z, \lambda)| \rightarrow 0$ quando $z \rightarrow \infty$.*

Exercício 3.7. *Mostre que, caso $|\lambda| < 1$ o Jacobiano da transformação é dado por $-J(\zeta)$, onde J foi definido no exercício 3.1.*

Exercício 3.8. *Faça os detalhes do Lema 3.2.*

Exercício 3.9. *Faça os detalhes do Lema 3.3.*

Exercício 3.10. *Faça os detalhes do Teorema 3.1.*

Exercício 3.11. *Faça os detalhes do Lema 3.4.*

Exercício 3.12. *Faça os detalhes do Lema 3.5.*

Exercício 3.13. *Mostre que a fórmula de reconstrução (3.2.16) é igual a inversa da Transformada de Radon, quando $a = 0$.*

Exercício 3.14. *Deduza a fórmula de reconstrução no caso em que $a(\mathbf{x})$ é constante.*

Capítulo 4

Tomografias por Difração e Aproximações de Born

Como discutimos na Seção 1.5, problemas relacionados com ultrassom e radar são modelados pela equação da onda, nos quais se conhece a onda refletida pelo objeto (espalhamento inverso) imerso no meio. O problema de recuperar propriedades do meio através de informações medidas em espalhamento inverso é muito difícil. Uma boa referência para entender quais as dificuldades envolvidas nestes modelos é [20].

4.1 Espalhamento Inversos em Uma Dimensão

Um modelo simples de entender e muito intuitivo é o modelo unidimensional. Vamos considerar a equação da onda (1.5.32) em uma dimensão, com termo forçante $\delta(t)\delta(x - x_s)$, $(t, x) \in \mathbb{R} \times \mathbb{R}$. Esta é a distribuição de Dirac em $t = 0$ e na posição x_0 , [12].

Como verificamos na Subseção 1.5.5, no espaço das frequências, a equação da onda corresponde a equação de Helmholtz (1.5.36), com velocidade dada pela equação (1.5.35).

No caso unidimensional, a solução u_{in} da equação de Helmholtz (1.5.38), com condição de radiação como em (1.5.36), é dada por

$$u_{in}(x, \omega; \theta) = g(x - \theta, \omega) = \frac{-e^{ik|x-\theta|}}{2i k\pi}, \quad (4.1.1)$$

onde $g(x, \omega)$ representa a função de Green para a equação de Helmholtz.

Pelo princípio de superposição, a solução da equação (1.5.39) é dada por

$$u_s(x, \omega; \theta) = \omega^2 \int_{\theta}^{\infty} \frac{a(x)}{c^2} (u_s - u_{in})(y, \omega) g(x - y, \omega) dy, \quad (4.1.2)$$

onde u_s representa os dados refletidos.

4.1.1 Aproximação de Born

A aproximação de Born para o problema de espalhamento inverso em uma dimensão consiste em considerar que $a(x)$ é pequeno com relação a $v(x)$. Isto corresponde a uma linearização em torno da velocidade constante c . Deste modo, a quantidade $a(x)u_s$ é da ordem de a^2 . Portanto muito pequena de forma que pode ser desconsiderada no lado direito de (4.1.2). Substituímos u_{in} em (4.1.2) por sua expressão em (4.1.1) e obtemos que

$$u_s(x_s, ck/2; \theta) = - \int_{\mathbb{R}} \frac{a(x)}{4} e^{ikx} dx \quad (4.1.3)$$

é uma boa aproximação para os dados refletidos u_s .

Teorema 4.1. *A quantidade $a(x)$ é unicamente determinada por u_s em (4.1.3). Ainda, a reconstrução é estável.*

Demonstração: De fato, por (4.1.3), $a(x)$ é dada pela Transformada de Fourier Inversa

$$a(x) = -\frac{2}{\pi} \int_{\mathbb{R}} e^{-ikx} u_s(x_s, ck/2; \theta) dk.$$

Pelo Teorema 1.2, $a(x)$ é unicamente determinado e a reconstrução é estável. ■

4.2 Espalhamento Inverso em Dimensão Dois ou Três

Nesta seção, faremos uma generalização da seção anterior, para dimensões dois e três. Como comentado anteriormente, o caso geral é muito complicado, assim utilizaremos uma aproximação de Born para obter uma fórmula de reconstrução.

4.2.1 Aproximação de Born

A aproximação de Born que estamos interessados nestas notas consiste em linearizar a equação (1.5.39) em torno da velocidade constante $v(\mathbf{x}) = c$. Em outras palavras, assumiremos que $a(\mathbf{x})$ é muito pequeno em norma. Como o termo $a(\mathbf{x})u_s$ é um termo de segunda ordem em (1.5.39) este pode ser desconsiderado. Desta forma a equação (1.5.39) fica

$$\begin{aligned} (\Delta + k^2) u_s(\mathbf{x}, \omega) &= -a(\mathbf{x})k^2 u_{in}(\mathbf{x}, \omega) & (4.2.4) \\ \frac{\mathbf{x}}{|\mathbf{x}|} \cdot \nabla u_s(\mathbf{x}, \omega) - iku_s(\mathbf{x}, \omega) &= o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned}$$

Lema 4.1. *A equação (4.2.4) possui uma única solução dada por*

$$u_s(\mathbf{x}, \omega) = k^2 \int_{\mathbb{R}^n} a(\mathbf{y}) u_{in}(\mathbf{y}, \omega) G_n(\mathbf{x} - \mathbf{y}, \omega) d\mathbf{y}, \quad (4.2.5)$$

onde G_n é a função de Green da equação

$$\begin{aligned} (\Delta + k^2) G_n(\mathbf{x}, \omega) &= \delta(\mathbf{x}) \\ \frac{\mathbf{x}}{|\mathbf{x}|} \cdot \nabla u_s(\mathbf{x}, \omega) - iku_s(\mathbf{x}, \omega) &= o(|\mathbf{x}|^{-(n-1)/2}) \end{aligned}$$

que é dada por

$$G_2(\mathbf{x}) = \frac{i}{4} \widetilde{\mathcal{H}}_0(k|\mathbf{x}|), \quad G_3(\mathbf{x}) = \frac{e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} \quad (4.2.6)$$

onde $\widetilde{\mathcal{H}}_0$ é a função de Hankel dada por

$$\widetilde{\mathcal{H}}_0(K|(x, y)|) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{\sqrt{K^2 - p^2}} e^{i(px + \sqrt{k^2 - p^2}y)} dy.$$

Demonstração: Segue facilmente do método de sobreposição. ■

4.2.2 Fórmula de Reconstrução

Lembramos que, matematicamente, conhecemos u_s no limite quando $|\mathbf{x}| \rightarrow \infty$. Assim, usando que

$$|\mathbf{x} - \mathbf{y}| = |\mathbf{x}| + \mathbf{x}_0 \cdot \mathbf{y} + \mathcal{O}((\mathbf{x}, \mathbf{y})),$$

onde $\mathcal{O}((\mathbf{x}, \mathbf{y}))$ representa termos de ordem menor, temos que o valor de u_s como no Lema 4.1 é dado por

$$u_s(\mathbf{x}, \omega) = \frac{k^2 e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} \int_{\mathbb{R}^n} a(\mathbf{y}) e^{ik(\boldsymbol{\theta} - \mathbf{x}_0)y} dy + \mathcal{O}((\mathbf{x}, \mathbf{y})). \quad (4.2.7)$$

Portanto, vemos de (4.2.7) que as medidas u_s podem ser escritas da forma

$$u_s(\mathbf{x}, \omega) = \frac{k^2 e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} \hat{a}(k(\mathbf{x}_0 - \boldsymbol{\theta})) + \mathcal{O}(1/|\mathbf{x}|). \quad (4.2.8)$$

Observação 4.1. Lembramos que $\omega = ck$. Assim, para uma onda plana a uma dada frequência ω , isto é, em um dado número de onda k , numa direção $\boldsymbol{\theta}$ a quantidade

$$\hat{a}(k(\mathbf{x}_0 - \boldsymbol{\theta})) = \frac{k^2 e^{-ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} u_s(\mathbf{x}, \omega), \quad (4.2.9)$$

para $|\mathbf{x}|$ suficientemente grande.

Portanto, cada nova medida nos fornece novas informações sobre a Transformada de Fourier de $a(\mathbf{x})$.

Características Especiais: No caso bidimensional, podemos observar os seguintes tipos de medidas distintas:

1. Medidas na direção \mathbf{x} tal que $\mathbf{x} \cdot \boldsymbol{\theta} > 0$. Estas medidas são conhecidas como **medidas de transmissão**, pois são as medidas de ondas que passaram pelo objeto a ser imageado.

2. Medidas na direção \mathbf{x} tal que $\mathbf{x} \cdot \boldsymbol{\theta} < 0$. Estas medidas são conhecidas como **medidas de reflexão**, pois são as medidas de ondas que são refletidas pelo objeto a ser imageado.

Em muitos casos só é possível obter acesso a um dos tipos de medidas acima.

Medidas de Transmissão

Vamos analisar primeiro o caso em que temos acesso somente medidas de transmissão como no Item (1). Ou seja, que temos acesso a $\hat{a}(k(\mathbf{x}_0 - \boldsymbol{\theta}))$ como em (4.2.9) para $\mathbf{x}_0 \cdot \boldsymbol{\theta} > 0$. Em particular, para $\mathbf{x}_0 = \boldsymbol{\theta}$, temos $\hat{a}(0)$. Este é a média dos valores de $a(\mathbf{x})$ no domínio.

Lema 4.2. *O problema inverso de Tomografia por Difração não possui única solução.*

Demonstração: De fato, como \mathbf{x}_0 varia sobre \mathbb{S}^1 tal que $\mathbf{x}_0 \cdot \boldsymbol{\theta} > 0$, obtemos o valor de $\hat{a}(\mathbf{y})$ sobre o semicírculo passando pela origem e simétrico com relação a $\boldsymbol{\theta}$. Como $\boldsymbol{\theta}$ varia sobre o disco unitário, temos que $\hat{a}(k(\mathbf{x}_0 - \boldsymbol{\theta}))$ permanece no disco de raio $\sqrt{2}k$. Portanto, para cada k fixo, tudo o que podemos obter é $\hat{a}(\mathbf{y})$ para \mathbf{y} tal que $|\mathbf{y}| \leq \sqrt{2}K$. ■

No caso tridimensional o resultado é similar.

Fórmula de Reconstrução para Medidas de Transmissão

Como as frequências maiores que $\sqrt{2}k$ não podem ser reconstruídas, vamos assumir que

$$a(\mathbf{x}) = (\mathcal{F}^{-1} \xi_{\sqrt{2}k} \hat{a})(\mathbf{x}). \quad (4.2.10)$$

onde $\xi_{\sqrt{2}k} = 1$ se $|\mathbf{y}| \leq \sqrt{2}k$ e zero caso contrário.

Corolário 4.1. *Sob a hipótese que $a(\mathbf{y})$ é tal que $|\mathbf{y}| \leq \sqrt{2}k$, então o problema inverso em Tomografia por Difusão possui uma única solução. Esta solução é estável.*

Demonstração: Segue de (4.2.10) e do Teorema 1.2. ■

O próximo passo é obtermos uma fórmula de reconstrução. Vamos analisar o caso bidimensional.

Queremos é recuperar $a(\mathbf{x})$, conhecendo $\hat{a}(\mathbf{x}_0 - \boldsymbol{\theta})$ com \mathbf{x}_0 e $\boldsymbol{\theta}$ em \mathbb{S}^1 . Em outras palavras, queremos recuperar uma função de duas variáveis a partir de dados que dependem de duas variáveis também, a saber \mathbf{x}_0 e $\boldsymbol{\theta}$.

Pela Transformada Inversa de Fourier, temos que

$$a(\mathbf{x}) = \frac{k^2}{(2\pi)^2} \int_0^{2\pi} \int_0^{\sqrt{2}} e^{ik\rho\mathbf{x}\cdot\boldsymbol{\theta}} \hat{a}(k\rho\boldsymbol{\theta}) \rho d\rho d\boldsymbol{\theta}.$$

Notemos que, como $\boldsymbol{\theta}$ varia sobre o círculo unitário, todos os pontos do círculo $|\mathbf{y}| \leq \sqrt{2}k$ são identificados duas vezes pela variação de \mathbf{x}_0 , uma vez quando $\mathbf{y} \cdot \boldsymbol{\theta}^\perp > 0$ e outra quando $\mathbf{y} \cdot \boldsymbol{\theta}^\perp \leq 0$. Portanto, a informação contida em $\mathbf{y} \cdot \boldsymbol{\theta}^\perp > 0$ é suficiente. Tal informação pode ser parametrizada da seguinte maneira: Para um dado $\boldsymbol{\theta}$, seja \mathbf{x}_0 parametrizado por

$$\mathbf{x}_0(\beta\boldsymbol{\theta}) = \text{sen}(\beta\boldsymbol{\theta}) + \cos(\beta\boldsymbol{\theta}^\perp), \quad 0 \leq \beta \leq \pi/2.$$

Desta parametrização obtemos que

$$\begin{aligned} \hat{a}(k(\mathbf{x}_0 - \boldsymbol{\theta})) &= \hat{a}(k\rho(\beta) (\text{sen}(\beta - 1)\boldsymbol{\theta} + \cos(\beta)\boldsymbol{\theta}^\perp) (\rho(\beta))^{-1}) \\ &= \hat{a}(k\rho(\beta)R(\beta)\boldsymbol{\theta}), \end{aligned}$$

onde $\rho(\beta) = \sqrt{2(1 - \cos(\beta))}$ e $R(\beta)$ é a matriz de rotação (que depende de β).

Teorema 4.2. *Para $x_0 = \mathbf{x}_0(\beta, \boldsymbol{\theta})$ parametrizado como acima, temos a seguinte fórmula de reconstrução*

$$a(\mathbf{x}) = \frac{k^2}{(2\pi)^2} \int_0^{2\pi} \int_0^{\sqrt{2}} e^{ik\rho(\beta)\mathbf{x}\cdot R(\beta)\boldsymbol{\theta}} \hat{a}(k|\mathbf{x}_0 - \boldsymbol{\theta}|) \text{sen}(\beta) d\beta d\boldsymbol{\theta}.$$

Demonstração: É simples e assim deixamos como exercício. ■

Medidas de Reflexão: Algumas Observações

Para finalizar faremos algumas observações com relação ao caso de termos acesso a medidas de reflexão como no Item (2).

Seguindo os mesmos passos acima, podemos mostrar que não é possível recuperar o número de onda de $a(\mathbf{x})$ para baixas frequências e nem para altas frequências. Em outras palavras, só é possível recuperar valores de \mathbf{y} que estejam no anel $\sqrt{2}k < |\mathbf{y}| < 2K$. Para valores que estão no anel acima, é possível obtermos uma fórmula de reconstrução como anteriormente. Este é um desafio interessante para o leitor. Para maiores detalhes consulte [2].

4.3 Exercícios

Exercício 4.1. *Mostre que (4.1.1) satisfaz a equação de Helmholtz (1.5.38) com a condição de radiação (1.5.36).*

Exercício 4.2. *Demonstre o Lema 4.1.*

Exercício 4.3. *Use o princípio da superposição para mostrar que a solução da equação (1.5.39) é dada por u_s como na equação (4.1.2), para o caso unidimensional.*

Exercício 4.4. *Faça os detalhes da demonstração do Corolário 4.1.*

Exercício 4.5. *Faça os cálculos na demonstração do Teorema 4.2.*

Capítulo 5

Tomografia por Fluorescência Óptica

Dedicaremos este capítulo para a obtenção de uma fórmula para recuperar f na equação (1.5.28) com uma fonte de espalhamento, a partir das medidas (1.5.31), em dimensão dois.

5.1 Uma Fórmula de Inversão Utilizando Séries de Neumann

Nossa abordagem é muito simplificada mas dá algumas ideias das técnicas a serem aplicadas. Estas simplificações dão origem a uma abordagem que permite utilizar a série de Neumann para garantir uma fórmula de inversão. Como sub-produto do resultado, temos um algoritmo iterativo para a inversão.

Iniciaremos esta seção definindo a transformada raio

$$Sg(\mathbf{x}, \theta) = \int_{-\infty}^0 g(\mathbf{x} + s\boldsymbol{\theta}) ds, \quad (5.1.1)$$

para $g \in L^1(\mathbb{R})$ e $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta)) \in \mathbb{S}^1$, para $\theta \in [0, 2\pi)$.

Lema 5.1. *A função $v(\mathbf{x}, \theta) = e^{Aa}(\mathbf{x})u(\mathbf{x}, \theta)$ satisfaz a seguinte equação*

integral

$$v(\mathbf{x}, \theta) = [Se^{Aa}Ke^{-Aa}v](\mathbf{x}, \theta) + [Se^{Aa}f](\mathbf{x}, \theta). \quad (5.1.2)$$

Demonstração: É fácil e, portanto, deixamos para o leitor. ■

Definimos $T = Se^{Aa}Ke^{-Aa}$, então temos que a equação (5.1.2) pode ser reescrita por

$$(I - T)v(\mathbf{x}, \theta) = [Se^{Aa}f](\mathbf{x}, \theta). \quad (5.1.3)$$

Lema 5.2. *Suponhamos que exista uma norma $|||\cdot|||$ (que não vamos definir nestas notas) tal que $|||T||| < 1$. Então*

$$v(\mathbf{x}, \theta) = [Se^{Aa}f](\mathbf{x}, \theta) + [Se^{Aa}Ke^{-Aa}v](I - T)^{-1}Se^{Aa}f(\mathbf{x}, \theta).$$

Demonstração: Sobre a hipótese de que $|||T||| < 1$, temos que $(I - T)$ é inversível. O resultado agora segue de (5.1.2) e de (5.1.3). ■

Sob certas hipóteses, o próximo teorema produz uma fórmula de inversão. Na verdade, mais que isso, garante um método iterativo convergente e estável para obtermos f .

Teorema 5.1. *Definimos $Z = [Se^{Aa}Ke^{-Aa}v](I - T)^{-1}Se^{Aa}$. Sob a hipótese de que $|||Z|||$ é suficientemente pequena, vale a seguinte fórmula de reconstrução*

$$f(\mathbf{x}) = \sum_{j=0}^{\infty} [-\mathcal{N}\mathcal{R}Z]^j \mathcal{N}g(\mathbf{x}). \quad (5.1.4)$$

onde $[\mathcal{N}g](x)$ é o operador definido em (3.2.16).

Demonstração: Escrevemos $\mathbf{x} = s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}$. Fazemos $t \rightarrow \infty$ em (1.5.26) e obtemos que

$$g(\theta, s) = \mathcal{R}_a f(\theta, s) + \mathcal{R}Zf(\theta, s).$$

Aplicamos o operador \mathcal{N} de ambos os lados da igualdade acima e obtemos que

$$f(\mathbf{x}) + \mathcal{N}\mathcal{R}Zf(\mathbf{x}) = \mathcal{N}g(\mathbf{x}).$$

Utilizamos a hipótese que a norma do operador Z é suficientemente pequena, então temos que o operador $I - (-\mathcal{N}RZ)$ é inversível e sua inversa pode ser dada pela série de Newmann (5.1.4). ■

A hipótese feita no Teorema 5.1, para garantir a inversa do operador $(I - Z)$, está relacionada com o fato que o campo de espalhamento seja pequeno. Assim, essa é uma hipótese que não é muito absurda do ponto de vista físico.

A fórmula de inversão obtida no Teorema 5.1 garante o seguinte algoritmo iterativo para a reconstrução de f :

Corolário 5.1. *Sob as hipóteses do Teorema 5.1 existe um algoritmo iterativo para recuperar f no problema inverso relacionado a Tomografia por Fluorescência óptica.*

Demonstração: Definimos $f^0(\mathbf{x}) = \mathcal{N}g$. Para cada $j \geq 0$, defina u_s^j a solução do seguinte problema

$$\boldsymbol{\theta} \cdot \nabla u_s^j(\mathbf{x}, \theta) + au_s^j(\mathbf{x}, \theta) = Ku_s^j(\mathbf{x}, \theta)_f^j(\mathbf{x}). \quad (5.1.5)$$

Atualizamos f por

$$f^{j+1}(\mathbf{x}) = \mathcal{N}g(\mathbf{x}) - \mathcal{N}\mathcal{R}_a e^{Aa} Ku_s^j(\mathbf{x}). \quad (5.1.6)$$

Observamos que, se o espalhamento é pequeno, como assumido no Teorema 5.1, f^j converge (em alguma norma adequada) para f . ■

5.2 Exercícios

Exercício 5.1. *Demonstre o Lema 5.1.*

Exercício 5.2. *Mostre que o operador Z definido no Teorema 5.1 é um operador linear. Dê condições sobre as normas de S, K e T para que Z seja um operador limitado.*

Exercício 5.3. *Dê condições na norma de S, K e T tal que $\|Z\| < 1$.*

Exercício 5.4. *Mostre que, se $\|Z\| < 1$, então $I - Z$ é um operador inversível e a inversa pode ser dada pela série de Newmann.*

Exercício 5.5. *Faça os detalhes da demonstração do Corolário 5.1.*

Capítulo 6

Sistemas de Equações Lineares

Neste capítulo, faremos uso de algumas ferramentas da Álgebra Linear e Teoria Linear de Operadores para introduzir algumas técnicas muito utilizadas em Problemas Inversos.

Os capítulos anteriores nos ensinaram que, dado um problema linear (um sistema de equações lineares, por exemplo), o operador que rege o problema nem sempre é injetivo e, mesmo que seja esse o caso, não é nada recomendável inverter esse operador.

Apresentaremos abaixo dois conceitos relacionados com a solução de problemas inversos. Estes são a *Inversa Generalizada ou Pseudo-Inversa* de operadores lineares e o *Teorema de Decomposição em Valores Singulares (SVD)*¹ ou *Teorema Espectral*. O primeiro destes conceitos nos permitirá definir uma solução, com uma propriedade especial, dentre todas as possíveis soluções (que podem ser infinitas) de um problema inverso. O segundo permite decompor um operador (uma matriz) como a soma de projeções sobre certos subespaços. Além disso, essa teoria nos ajuda a entender a influência dos autovalores de um operador na solução de problemas inversos.

¹O Teorema de Decomposição em Valores Singulares é um dos teoremas mais fortes da matemática. Existem versões deste Teorema para operadores auto-adjuntos não limitados [24].

6.1 Pseudo-Inversa de Operadores Lineares

Na verdade, tudo o que queremos em problemas inversos é: Encontrar uma maneira de aproximar um operador (o operador inverso) por uma família de operadores bem postos. Conseqüentemente, encontrar uma aproximação (a melhor possível) para a solução do problema.

Nesta seção, apresentaremos uma forma de aproximarmos “da melhor maneira” o inverso de um operador linear. Desenvolveremos a teoria para operadores lineares limitados que possuam imagem fechada, que é o caso de operadores em dimensão finita (matrizes) e operadores compactos. Assim, com a teoria linear, cobrimos uma ampla quantidade de casos interessantes. Faz jus mencionarmos que existem versões dos resultados apresentados abaixo para operadores lineares limitados quaisquer. Para uma abordagem completa sobre o assunto, veja [14].

6.1.1 Definições e Propriedades Básicas

Esperamos que o leitor esteja familiarizado com as definições de matrizes **Hermitianas**, **Simétricas**, **Unitárias**, **Normais**, etc. Caso contrário, uma breve leitura em livros de Álgebra Linear é recomendado. Veja, por exemplo, [29, 37]. Nosso objetivo é encurtarmos o caminho. Assim, vamos direto ao Teorema Espectral² (dimensão infinita) para operadores compactos e auto-adjuntos e obteremos, como corolário, o Teorema da SVD (dimensão finita).

Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ um operador linear e limitado, onde \mathcal{H}_1 e \mathcal{H}_2 denotam espaços de Hilbert. Consideraremos o problema fundamental de resolver a equação linear do tipo

$$Ax = y, \tag{6.1.1}$$

onde $y \in \mathcal{H}_2$.

Exemplo 6.1. *Exemplos da equação (6.1.1) são:*

²Teorema Espectral é o nome que se dá ao Teorema de Decomposição em Valores Singulares para operadores em dimensão infinita.

Caso em que $\mathcal{H}_1 = \mathbb{R}^n$, $\mathcal{H}_2 = \mathbb{R}^m$ e $A \in M_{m \times n}(\mathbb{R})$.

Caso em que $\mathcal{H}_1 = \mathcal{H}_2 = L^2[0, 1]$ e A é um operador integral da forma

$$(Ax)(s) = \int_0^1 k(s, t)x(t)dt, \quad s \in [0, 1],$$

e $k(s, t) \in (C[0, 1] \times C[0, 1])$ é o chamado *Kernel*³.

Como já vimos, se o operador A possui uma inversa, então a equação (6.1.1) possui uma única solução $x = A^{-1}y$. Mas, nossa experiência anterior nos ensinou que "nem tudo são rosas", isto é, pode acontecer de que $\mathcal{N}(A) \neq \{0\}$ ou, talvez, $y \notin \text{Im}(A)$.

Um fato confortante é que, mesmo no caso da equação (6.1.1) não possuir uma solução no sentido tradicional, é possível definir uma solução generalizada do problema que é "a melhor" entre as soluções generalizadas de (6.1.1). Para tal, necessitamos de certas hipóteses sobre a imagem do operador A . No caso em que A é um operador compacto, a hipótese que faremos abaixo não é restritiva.

Hipótese 6.1. *Suponha que A é um operador linear limitado e que $\text{Im}(A)$ é fechada em \mathcal{H}_2 .*

Seja $P : \mathcal{H}_2 \rightarrow \text{Im}(A)$ o operador de projeção ortogonal (que está bem definido pelo Exercício 6.6). Assim, $Py \in \text{Im}(A)$ é o vetor mais próximo de y .

Definição 6.1. *Uma solução generalizada de (6.1.1) é qualquer solução $u \in \mathcal{H}_1$ da equação*

$$Ax = Py. \quad (6.1.2)$$

Exemplo 6.2. *Suponha que $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ e $y = (2, 1)^T$.*

Então $\text{Im}(A) = \text{span}\{(1, -1)^T\}$ e $Py = (1/2, -1/2)^T$. Portanto, o conjunto de soluções generalizadas é dada por

$$\left\{ (x_1, x_2)^T : x_2 = \frac{1}{2} + x_1 \right\}.$$

³Não confundir com o núcleo do operador A

Geometricamente, uma solução generalizada, como na Definição 6.1, significa encontrar $u \in \mathcal{H}_1$ solução do problema de minimização

$$u = \arg \min_{x \in \mathcal{H}_1} \frac{1}{2} \|Ax - y\|^2. \quad (6.1.3)$$

Mais geral, denotando por A^* o operador adjunto de A , segue

Teorema 6.1. *Suponha que $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ seja um operador linear limitado, $y \in \mathcal{H}_2$ e a Hipótese 6.1 seja satisfeita. Então, as seguintes condições sobre $u \in \mathcal{H}_1$ são equivalentes:*

$$(i) \quad Au = Py,$$

$$(ii) \quad u = \arg \min_{x \in \mathcal{H}_1} \frac{1}{2} \|Ax - y\|^2,$$

$$(iii) \quad A^*Au = A^*y \quad (\text{conhecidas como Equações Normais}).$$

Demonstração: $(i) \Rightarrow (ii)$: Seja $Au = Py$. Segue do exercício 6.2 e do Teorema de Pitágoras que, dado $x \in \mathcal{H}_1$,

$$\begin{aligned} \|Ax - y\|^2 &= \|Ax - Py\|^2 + \|Py - y\|^2 \\ &= \|Ax - Py\|^2 + \|Au - y\|^2 \geq \|Au - y\|^2. \end{aligned}$$

$(ii) \Rightarrow (iii)$: Por hipótese, existe pelo menos um $x \in \mathcal{H}_1$ solução de (6.1.2). Disto e do Teorema de Pitágoras, temos que

$$\|Au - y\|^2 = \|Au - Py\|^2 + \|Py - y\|^2 \geq \|Au - Py\|^2 + \|Au - y\|^2.$$

Portanto $Au - y = Py - y \in (\text{Im}(A))^\perp = \mathcal{N}(A^*)$. Assim, $A^*(Au - y) = 0$ e (iii) segue.

$(iii) \Rightarrow (i)$: De (iii) satisfeito, obtemos que $Au - y \in (\text{Im}(A))^\perp$ e, assim, $0 = P(Au - y) = Au - Py$. ■

Definição 6.2. *Um vetor $u \in \mathcal{H}_1$ satisfazendo qualquer uma das sentenças do Teorema 6.1 é chamado uma **solução dos mínimos quadrados** da equação $Ax = y$.*

Uma observação importante a ser feita é a de que, sob a Hipótese 6.1, uma solução de mínimos quadrados de (1.2.1) sempre existe $\forall b \in \mathcal{H}_2$ (veja Exercício 6.4). Caso $\mathcal{N}(A) \neq 0$ então, existe uma infinidade de soluções de mínimos quadrados de (1.2.1). De fato, se u é uma solução de mínimos quadrados e $v \in \mathcal{N}(A)$, então $u + v$ também é uma solução de mínimos quadrados.

Estamos buscando um caminho para "inverter" o operador A , associando com cada $b \in \mathcal{H}_2$ uma única solução de mínimos quadrados. Sabemos que, se $\mathcal{N}(A) \neq 0$, não temos tal unicidade. Será que temos alguma alternativa? A resposta é afirmativa: basta para tal, escolhermos entre as (várias) possíveis soluções, uma que tenha uma característica especial. Mas, que característica especial podemos escolher num conjunto que possui, possivelmente, uma infinidade de elementos?

Vamos voltar e analisar o que temos de hipóteses e resultados.

- i) O conjunto de soluções de mínimos quadrados é não-vazio.
- ii) O conjunto de soluções de mínimos quadrados é **convexo e fechado** (ver Exercício 6.3).

Portanto, pelo *Teorema da Projeção* [24], existe uma única solução de mínimos quadrados com norma mínima associada a cada elemento $b \in \mathcal{H}_2$. Logo, temos um caminho para encontrar uma inversa (mesmo que generalizada) para o operador A .

Definição 6.3. *Seja A um operador satisfazendo a Hipótese 6.1. A aplicação*

$$A^\dagger : \mathcal{H}_2 \longrightarrow \mathcal{H}_1, \quad A^\dagger b = u$$

onde u é a **única** solução de mínimos quadrados de norma mínima da equação (6.1.1). Tal solução é chamada de *Inversa Generalizada* de A .

Existem outras definições de inversa generalizada de operadores que são equivalentes a dada acima (veja [14]).

Teorema 6.2. *Se A satisfaz a Hipótese 6.1, então $Im(A^\dagger) = Im(A^*) = Im(A^\dagger A)$.*

Demonstração: Seja $b \in \mathcal{H}_2$. Num primeiro momento, mostraremos que $A^\dagger b \in \mathcal{N}(A)^\perp$ e então usaremos o Exercício 6.8. Suponha que

$$A^\dagger b = u_1 + u_2 \in \mathcal{N}(A)^\perp \oplus \mathcal{N}(A).$$

Então, u_1 é uma solução de mínimos quadrados de $Ax = b$. De fato,

$$Au_1 = A(u_1 + u_2) = AA^\dagger b = Pb.$$

e, portanto, a afirmação está completa usando o Teorema 6.1.

Suponhamos que $u_2 \neq 0$. Então, pelo Teorema de Pitágoras,

$$\|u_1\|^2 < \|u_1 + u_2\|^2 = \|A^\dagger b\|^2$$

contradizendo o fato de $A^\dagger b$ ser uma solução de mínimos quadrados que tem a norma mínima. Logo, $A^\dagger b = u_1 \in \mathcal{N}(A)^\perp$.

Reciprocamente, sejam $u \in \mathcal{N}(A)^\perp$ e $b = Au$. Então,

$$Au = PAu = Pb$$

e, assim, u é uma solução de mínimos quadrados. Se x é qualquer outra solução de mínimos quadrados, então

$$Ax = Pb = Au$$

e, portanto, $x - u \in \mathcal{N}(A)$. Pelo Teorema de Pitágoras,

$$\|x\|^2 = \|u\|^2 + \|x - u\|^2 \geq \|u\|^2.$$

Assim, u é a solução de mínimos quadrados que tem norma mínima, i.e., $u = A^\dagger b$.

Isto prova a primeira das igualdades. Para verificar a segunda, notemos que, para qualquer $b \in \mathcal{H}_2$,

$$A^\dagger b = A^\dagger Pb \in \text{Im}(A^\dagger A). \quad \blacksquare$$

Corolário 6.1. *Se A satisfaz a Hipótese 6.1, então $A^\dagger : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ é linear e limitado.*

Um fato muito importante para o que segue na teoria de regularização para problemas inversos é consequência do seguinte resultado:

Teorema 6.3. *A inversa generalizada A^\dagger possui gráfico $Gr(A^\dagger)$ fechado. Consequentemente, A^\dagger é contínua se, e só se, $Im(A)$ é fechada.*

Demonstração: Uma parte deste Teorema é consequência do Corolário 6.1. A demonstração completa do Teorema foge ao escopo destas notas, pois, usa conceitos fortes de Análise Funcional em particular o Teorema do Gráfico Fechado, [24]. Interessados na demonstração podem consultar [10]. ■

Observação 6.1. *O Teorema 6.3 reforça ainda mais a diferença entre problemas inversos em dimensão finita e infinita. Pois, no caso de dimensão finita, o operador (matriz) A sempre possui a imagem fechada. Assim, temos a garantia de existência e unicidade de uma solução de mínimos quadrados de norma mínima.*

6.2 A Decomposição em Valores Singulares

Um dos principais resultados da Álgebra Linear é o *Teorema de Decomposição em Valores Singulares (SVD)*. Este teorema permite escrever uma matriz qualquer como uma soma de matrizes de projeção de posto 1. Mais geral, o Teorema de SVD vale para operadores lineares em espaços de Hilbert de dimensão infinita que são auto-adjuntos⁴. Neste caso, o Teorema de SVD é conhecido como *Teorema Espectral*, [24]. No caso especial em que A é um operador linear e compacto, o Teorema Espectral se traduz de forma similar ao caso de dimensão finita, [24, 37, 29, 13].

Observação 6.2. *Faremos a hipótese de que o corpo de escalares do espaço vetorial é o corpo dos números complexos. Assim, se temos uma matriz $n \times n$, esta possui n autovalores. Esse fato é importante no que segue.*

⁴Um operador A entre espaços de Hilbert é Auto-Adjunto se $\mathcal{D}(A^*) = \mathcal{D}(A)$ e $\langle x, Ay \rangle = \langle Ax, y \rangle, \forall x, y \in \mathcal{D}(A)$.

Nosso ponto de partida é uma versão simplificada do Teorema SVD, a qual faremos a demonstração. Formulações mais gerais podem ser encontradas em [24, 37, 29, 13].

Teorema 6.4. [*Diagonalização*] *Seja A uma matriz quadrada de ordem $n \times n$ com um conjunto de n autovetores linearmente independentes. Então, A é similar a uma matriz diagonal ou diagonalizável.*

Demonstração: Construa a matriz S tendo como colunas os vetores $\{v_1, \dots, v_n\}$. Assim:

$$\begin{aligned} AS &= A \begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & | & \cdots & | \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \\ | & | & \cdots & | \end{pmatrix} \\ &= S.D(\lambda_1, \lambda_2, \dots, \lambda_n), \end{aligned}$$

onde, $D(\lambda_1, \lambda_2, \dots, \lambda_n)$ representa a matriz diagonal formada por $(\lambda_1, \lambda_2, \dots, \lambda_n)$. Como S é invertível, temos

$$A = SD(\lambda_1, \lambda_2, \dots, \lambda_n)S^{-1}. \quad \blacksquare$$

Corolário 6.2. *Seja A uma matriz com de ordem $n \times n$ que possui n autovalores distintos. Então, A é diagonalizável.*

Pergunta: Será que toda matriz quadrada é diagonalizável? Não, pelo menos, no sentido do Teorema 6.4. Veja a resposta no Exercício 6.18.

O Teorema 6.4 é a versão mais simples do Teorema de Decomposição em Valores Singulares. Passaremos agora para uma versão mais geral.

Teorema 6.5. *Todo operador compacto possui no máximo uma quantidade enumerável de autovalores que formam uma sequência cujos valores absolutos convergem para zero. Os autovalores de um operador auto-adjunto são reais.*

Teorema 6.6. [Teorema Espectral⁵ - *A compacto e auto-adjunto*] Seja $A : \mathcal{H} \rightarrow \mathcal{H}$ um operador compacto e auto-adjunto. Então, existe um sistema ortonormal completo $\{e_j\}$ de \mathcal{H} tal que $Ae_j = \lambda_j e_j$ e $\lambda_j \rightarrow 0$ quando $n \rightarrow \infty$.

Demonstração: Não faremos a demonstração, pois foge das nossas pretensões. Para a demonstração sugerimos que o leitor consulte [24].

■

Teorema 6.7. [Teorema Espectral - *A compacto*] Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ um operador linear compacto. Então, existem conjuntos ortogonais (não necessariamente completos) $\{e_1, \dots, e_m\}$ de \mathcal{H}_1 e $\{f_1, \dots, f_m\}$ de \mathcal{H}_2 e de números $\sigma_1, \dots, \sigma_m$ com $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$, tal que

$$Ax = \sum_{j=1}^m \sigma_j \langle x, e_j \rangle f_j, \quad x \in \mathcal{H}_1. \quad (6.2.4)$$

No caso da imagem do operador A ter dimensão infinita, temos que considerar $m \rightarrow \infty$. Neste caso, $\sigma_m \rightarrow 0$.

Demonstração: Como A é compacto sabemos que A^*A é compacto e auto-adjunto. Pelo Teorema Espectral 6.6, existe um conjunto ortogonal $\{e_1, \dots, e_m\}$ de \mathcal{H}_1 tal que $A^*Ae_j = \lambda_j e_j$, onde $0 \leq \lambda_j \in \mathbb{R}$. Defina $\sigma_j = \sqrt{\lambda_j}$ e $f_j = \frac{1}{\sigma_j} Ae_j$ (para $\sigma_j > 0$). Um cálculo simples mostra que $\{f_1, \dots, f_m\}$ é um conjunto ortonormal e que a equação (6.2.6) é satisfeita. ■

Definição 6.4. Os valores $\sigma_1, \dots, \sigma_m$ são chamados de **valores espectrais** de A . Chamamos de **sistema singular** de A a tripla (σ_j, e_j, f_j) .

Corolário 6.3. [Teorema espectral em dimensão finita - SVD.] Seja $A \in \mathbb{R}^{m,n}$ com $n \leq m$. Então, existem matrizes unitárias $U \in M_{m \times m}$,

⁵O Teorema Espectral como enunciado também é conhecido como Teorema de Hilbert-Schmidt. Existem várias versões deste Teorema,[24].

$V \in M_{n \times n}$ e uma matriz diagonal com entradas não-negativas $\Sigma := D\{\sigma_1, \dots, \sigma_n\}$ tais que

$$A = U\Sigma V^T.$$

Os passos para demonstrar o teorema SVD estão nos Exercícios 6.23 e 6.24.

6.2.1 Funções de Operadores: O Teorema da Aplicação Espectral

Daremos agora uma breve introdução ao *Cálculo Funcional*, como é conhecida a teoria que trata de funções de Operadores Lineares. Essa importantíssima ferramenta matemática nos ensinará a derivar as chamadas de *Funções Filtro* que são a peça chave para o entendimento dos Métodos de Regularização dos capítulos a seguir.

Por simplicidade, daremos somente a ideia intuitiva em dimensão finita. Para interessados em detalhes mais aprofundados, recomendamos [24] e livros de Análise Funcional em geral.

Sejam $t \in [0, T]$ e $g(t) = a_0 + a_1t + \dots + a_nt^n$ um polinômio de ordem n em t .

Definição 6.5. *Seja A um operador linear limitado ($\|A\| \in [0, T]$), definimos um polinômio do operador A por*

$$g(A) = a_0 + a_1A + \dots + a_nA^n$$

onde $g(t)$ é o polinômio acima.

O que acontece com uma função contínua de um operador linear limitado? A resposta é dada pelo Teorema da Aplicação Espectral. Elucidaremos suas consequências através de um exemplo. Para interessados na demonstração consulte [24].

Exemplo 6.3. *Uma função de operadores muito especial é a exponencial de um operador linear limitado $\exp(A)$. Dar sentido a esse tipo de operações tem uma importância enorme na caracterização de soluções para sistemas de EDO's e na Teoria de Semigrupos associados a operadores diferenciais parciais.*

Vamos considerar o caso especial em que A é uma matriz e, mais ainda, esta satisfaz as hipóteses do Teorema 6.4.

Sabemos que a função $\exp(t)$ possui uma expansão em séries de potências dada por

$$\exp(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^k}{k!} + \cdots .$$

que converge uniformemente $\forall t \in \mathbb{R}$. Pelo o Exercício 6.26, temos que

$$\sum_{j=0}^{\infty} \frac{A^j}{j!} = S \cdot D(\exp(\lambda_1), \cdots, \exp(\lambda_n), \cdots) \cdot S^{-1} =: \exp(A).$$

Como o operador A é limitado, a série $\sum_{j=0}^{\infty} \frac{A^j}{j!}$ converge uniformemente na norma dos operadores e, assim, $\exp(A)$ está bem definida.

No caso especial em que A é uma matriz quadrada e injetiva, temos, do Teorema SVD, que

$$A^{-1} = V \Sigma^{-1} U^T, \quad (6.2.5)$$

onde $\Sigma^{-1} = D\{\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_n}\}$.

Observação 6.3. Notemos que a inversão de uma matriz pode ser pensada como a função $g(t) = t^{-1}$ aplicada a matriz.

De fato, podemos provar que este resultado não é mera coincidência. Denotemos por $\Sigma(A)$ o espectro de A .

Teorema 6.8. [Teorema da Aplicação Espectral.] Seja A um operador linear limitado e $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função contínua (à direita ou à esquerda). Então, $\Sigma(g(A)) = g(\Sigma(A))$.

Em particular, se A é compacto, então

$$g(A)x = \sum_{j=0}^m g(\sigma_j) \langle x, e_j \rangle f_j, \quad x \in \mathcal{H}.$$

É extremamente importante entender este resultado, pois servirá no entendimento, para a construção das estratégias de regularização (veja Capítulo 7) e para entender a relação existente entre mal condicionamento de um problema com os respectivos valores espectrais do operador associado.

6.2.2 Relação entre Má-colocação e Valores Espectrais

Segue do Teorema Espectral que, se o operador linear A é compacto e possui inversa, então

$$A^{-1}y = \sum_{j=1}^m \frac{1}{\sigma_j} \langle y, f_j \rangle e_j, \quad y \in \text{Im}(A). \quad (6.2.6)$$

Assim, se a dimensão do problema é grande, ou melhor, se σ_m esta próximo de zero, significa que $\frac{1}{\sigma_m}$ é grande. Portanto, pequenas perturbações na direção de um autovetor associado ao autovalor σ_m implicam numa grande variação na solução.

Explicaremos esse fenômeno de maneira mais clara através da Decomposição em Valores Singulares em dimensão finita.

Consideremos o problema inverso (em dimensão n) de recuperar x na equação matricial

$$Ax = y^\delta,$$

para um dado com ruídos $y^\delta = (y_1, \dots, y_n + \frac{1}{n})$. Assim, o erro nos dados é da ordem de $\frac{1}{n}$.

Suponhamos que a matriz A tenha sua decomposição espectral dada por (6.2.5) onde os valores singulares são $\sigma_j = \mathcal{O}(\frac{1}{j})^6$, $j = 1, 2, \dots, n$. Logo, A é inversível.

Da equação (6.2.5), temos que a solução com ruídos é dada por

$$x^\delta = A^{-1}y^\delta = V\Sigma^{-1}U^T y^\delta.$$

Como as matrizes U e V são unitárias e todas as normas são equivalentes (espaço de dimensão finita), temos a seguinte estimativa para a

⁶ $\mathcal{O}(a)$ significa que $\frac{\mathcal{O}(a)}{a} = \text{constante}$.

solução

$$\begin{aligned} \|x - x^\delta\|^2 &= \|V\Sigma^{-1}U^T y - V\Sigma^{-1}U^T y^\delta\|^2 = \sum_{j=1}^n (\mathcal{O}(j)(y_j - y_j^\delta))^2 \\ &= \left(\frac{\mathcal{O}(n)}{n}\right)^2 = (\mathcal{O}(1))^2. \end{aligned} \quad (6.2.7)$$

Note que o erro na solução é muito grande, se $(\mathcal{O}(1))^2$ for grande.

Vamos agora a um critério de solvabilidade de equações lineares governadas por operadores compactos. Muitos autores referem-se a este resultado como sendo o **Critério de Picard**, [3, 4, 10].

Teorema 6.9. *Seja A um operador compacto e (σ_j, e_j, f_j) um sistema singular de A . Dado $y \in \mathcal{H}_2$, as seguintes condições são equivalentes:*

- a) $y \in \text{Im}(A)$,
- b) $y \in \overline{\text{Im}(A)}$; $\sum_{j=0} \sigma_j^{-2} |\langle y, f_j \rangle|^2 < \infty$.

Demonstração: Daremos uma ideia da prova.

a) \Rightarrow b) De $y \in \text{Im}(A) \subset \overline{\text{Im}(A)}$. Seja $x \in \mathcal{H}_1$ tal que $Ax = y$. Segue do Teorema 6.7 que $A^*f_j = \sigma_j e_j$. Assim, usando a desigualdade de Bessel,

$$\begin{aligned} \sum_{j=0} \sigma_j^{-2} |\langle y, f_j \rangle|^2 &= \sum_{j=0} \sigma_j^{-2} |\langle Ax, f_j \rangle|^2 = \sum_{j=0} \sigma_j^{-2} |\langle x, A^*f_j \rangle|^2 \\ &= \sum_{j=0} |\langle x, e_j \rangle|^2 \leq \sum_{j=0} \|x\|^2 < \infty. \end{aligned}$$

b) \Rightarrow a) Defina $x_n := \sum_{j=0}^n \sigma_j^{-1} \langle y, f_j \rangle e_j$. Portanto, para $m, n \in \mathbb{N}$ temos:

$$\|x_n - x_m\|^2 = \sum_{j=n+1}^m \sigma_j^{-2} |\langle y, f_j \rangle|^2$$

e, portanto, $\{x_n\}$ é uma sequência de Cauchy, cujo limite denotaremos por $x := \lim_{n \rightarrow \infty} x_n$. Pela continuidade de A e definição de x_n , segue que

$$Ax = \sum_{j=0}^{\infty} \langle y, f_j \rangle f_j \quad \text{e} \quad \|Ax\| \leq \|y\|.$$

Definimos $z := y - \sum_{j=0}^{\infty} \langle y, f_j \rangle f_j$. Segue facilmente que

$$\|z\|^2 = \|y\|^2 - \sum_{j=0}^{\infty} |\langle y, f_j \rangle|^2; \quad \langle z, f_j \rangle = 0, \forall j \in \mathbb{N} \quad \text{e} \quad A^*z = 0.$$

Portanto, como $y \in \overline{\text{Im}(A)} = \mathcal{N}(A^*)^\perp$, temos

$$\langle z, y \rangle = \|y\|^2 - \sum_{j=0}^{\infty} |\langle y, f_j \rangle|^2 = \|z\|^2.$$

Logo, $y = \sum_{j=0}^{\infty} \langle y, f_j \rangle f_j = Ax$. ■

Observação 6.4. *Notemos que o Teorema de Picard 6.9 sugere que uma tentativa de solução para a equação $Ax = y$ é*

$$x = \sum_{j=0}^{\infty} \sigma_j^{-1} \langle y, f_j \rangle e_j. \quad (6.2.8)$$

Suponhamos que temos somente o dado perturbado $y^\delta = y + \delta f_k$. Substituindo na equação (6.2.8), obtemos como solução perturbada $x^\delta = x + \delta \sigma_k^{-1} e_k$. Assim,

$$\|x - x^\delta\|^2 = \left\| \frac{\delta}{\sigma_k} e_k \right\|^2 = \frac{\delta^2}{\sigma_k^2}. \quad (6.2.9)$$

Como $\sigma_j \rightarrow 0$, segue que a estimativa (6.2.9) pode ser muito ruim, mostrando o efeito de mal-condicionamento causado pelos valores singulares de um operador compacto.

Continuamos com o filosófico conceito de precisão. Quanto mais **precisos** procuramos ser (aproximando melhor o operador A^{-1}) mais longe ficamos de uma solução **precisa**, uma vez que, os dados não são medidos de maneira **precisa**.

6.3 Exercícios

Exercício 6.1. *Faça um paralelo entre as definições de matrizes Hermitianas, Simétricas, Unitárias, Normais e operadores adjuntos e auto-adjuntos.*

Exercício 6.2. *Suponha que a Hipótese 6.1 seja satisfeita e seja P o operador de projeção ortogonal sobre $\mathcal{R}(A)$. Prove que, dado $y \in \mathcal{H}_2$, $Py - y \in (\text{Im}(A))^\perp$.*

Exercício 6.3. *Mostre que o conjunto de soluções de mínimos quadrados pode ser escrito como*

$$\{u \in \mathcal{H}_1 : A^*Au = A^*b\}. \quad (6.3.10)$$

Também, prove que este conjunto é convexo e fechado.

Exercício 6.4. *Assuma que A satisfaz a Hipótese 6.1. Prove que existe pelo menos uma solução de mínimos quadrados. Dê condições sobre o operador A para que a solução de mínimos quadrados seja única.*

Exercício 6.5. *Prove que ambos os exemplos apresentados acima para a equação (6.1.1) são operadores compactos. Sugestão: Para o caso de A ser o operador integral, comece supondo que $k(s, t) \in (C[0, 1] \times C[0, 1])$ e use o Teorema de Ascoli-Arzelá. Use a densidade de $C[0, 1]$ em $L^2[0, 1]$.*

Exercício 6.6. *Seja \mathcal{H} um espaço de Hilbert e $C \subset \mathcal{H}$ um conjunto convexo e fechado. Prove que para todo $b \in \mathcal{H}$, existe uma única projeção de b sobre C . Prove ainda que a projeção de b sobre C é o vetor de C mais próximo de b .*

Exercício 6.7. *Mostre que, se A possui uma inversa, então $A^\dagger = A^{-1}$.*

Exercício 6.8. *Mostre que, se A satisfaz a Hipótese 6.1, então $\text{Im}(A^*)$ é fechada e $\mathcal{N}(A)^\perp = \text{Im}(A^*)$.*

Exercício 6.9. *Demonstre o Corolário 6.1.*

Exercício 6.10. Prove que se A é um operador linear entre espaços de dimensão finita, então $\text{Im}(A)$ é fechada.

Exercício 6.11. Suponha que A satisfaça a Hipótese 6.1. Prove que $A^\dagger : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ é o único operador linear limitado satisfazendo

$$AA^\dagger = P_{\text{Im}(A)} \text{ e } A^\dagger A = P_{\text{Im}(A^\dagger)}.$$

Esta é a definição de Moore para Inversa Generalizada. (Sugestão: Consulte [14])

Exercício 6.12. Prove que uma matriz quadrada A possui, no máximo n , autovalores. Dê um exemplo de uma matriz que não possui autovalores.

Exercício 6.13. Prove que, se estamos considerando o espaço vetorial sobre o corpo dos números complexos, então uma matriz quadrada $A_{(n \times n)}$ possui n autovalores.

Exercício 6.14. Dê um exemplo, em dimensão infinita, de um operador linear que não possui autovalores.

Exercício 6.15. Mostre que, dado um conjunto linearmente independentes de vetores, sempre existe um conjunto ortogonal. Mostre ainda que o espaço gerado pelos dois conjuntos são iguais. Sugestão: Use o Processo de Gram-Schmidt.

Exercício 6.16. Justifique, de maneira adequada, que a matriz S no Teorema 6.4 é de fato inversível.

Exercício 6.17. Prove o Corolário 6.2.

Exercício 6.18. Mostre que a matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ não é diagonalizável no sentido do Teorema 6.4.

Exercício 6.19. Mostre que, se λ_j satisfaz $A^* A e_j = \lambda_j e_j$, então $\lambda_j \in \mathbb{R}$ e $\lambda_j \geq 0$.

Exercício 6.20. Mostre que o conjunto $\{f_1, \dots, f_m\}$ definido como no Teorema 6.7 é ortonormal.

Exercício 6.21. *Mostre que se A^* é um operador linear compacto, então*

$$A^*y = \sum_{j=1}^m \sigma_j \langle y, f_j \rangle e_j, \quad x \in \mathcal{H}_1. \quad (6.3.11)$$

Exercício 6.22. *Mostre que se A^* é um operador linear compacto, então*

$$A^\dagger y = \sum_{j=1}^m \sigma_j^{-1} \langle y, f_j \rangle f_j, \quad y \in \mathcal{D}(A^\dagger). \quad (6.3.12)$$

Exercício 6.23. *Mostre que todo operador linear cuja imagem possui dimensão finita é compacto. Consequentemente, toda matriz é um operador linear compacto.*

Exercício 6.24. *Mostre que se A é uma matriz, então AA^* e A^*A são operadores compactos e auto-adjuntos.*

Exercício 6.25. *Demonstre o Corolário 6.3.*

Exercício 6.26. *Seja A uma matriz como no Teorema 6.4. Mostre que*

$$A^2 = SD(\lambda_1^2, \dots, \lambda_n^2)S^{-1}.$$

Use indução para mostrar que $A^p = SD(\lambda_1^p, \dots, \lambda_n^p)S^{-1}$, para qualquer $p \in \mathbb{N}$.

Capítulo 7

Regularização para Problemas Inversos

Passaremos a desenvolver as chamadas *Estratégias ou Métodos de Regularização* para Problemas Inversos. Sempre que procuramos solucionar um Problema Inverso, temos que contornar empecilhos como instabilidade e mal-condicionamento (má-colocação). Exemplos desta situação foram apresetados na Seção 1.3.

A maneira natural de solucionar um problema do tipo (1.2.1) (para dados exatos ou para dados perturbados por ruídos) é inverter o operador A . Mas, como foi exemplificado e como acontece na prática, nos deparamos muitas vezes (quase sempre) com operadores que não possuem inversa. Ou, se a inversa existe, esta é mal-condicionada (no caso de problemas em dimensão finita) ou ilimitada (no caso de dimensão infinita). Assim, o axioma de dependência contínua dos dados falha, produzindo soluções inadequadas para o problema.

Como vimos no Capítulo 6, mesmo que o operador A não possua inversa, podemos utilizar a solução de mínimos quadrados que possua norma mínima, ou seja, aproximar a solução $x^\dagger = A^\dagger y$ do problema (1.2.1). Relembrando, os dados y para o problema, em geral, não são conhecidos exatamente. Na prática, que temos em mão são dados aproximados y^δ , obtidos por medições, com o nível de ruídos δ ,

satisfazendo

$$\|y - y^\delta\| \leq \delta. \quad (7.0.1)$$

No caso em que A não possui, necessariamente, uma imagem fechada, o Teorema 6.3 nos ensina que $A^\dagger y^\delta$ não é uma boa aproximação para $A^\dagger y$, pois A^\dagger não é limitado. Portanto, temos que elaborar outras estratégias para solucionar estes problemas inversos. É isso que faremos nesse capítulo.

7.1 O Conceito de Regularização

Denominamos por *estratégia de regularização* o artifício matemático¹ de obtermos uma solução aproximada, digamos x_α^δ , de maneira **estável** e que **convirja** (em topologias adequadas), quando o nível de ruídos converge para zero, para a solução x^\dagger do problema inverso considerado. Além disso, o parâmetro α deve ser escolhido de maneira apropriada (seja lá o que isso signifique). Em termos gerais, uma *estratégia ou método de regularização* consiste em aproximar uma solução x^\dagger de um problema mal posto (1.2.1) por uma família (a um parâmetro α) de problemas bem postos. Mais precisamente:

Definição 7.1. *Sejam $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ um operador linear e limitado e $\alpha_0 \in (0, +\infty)$. Para todo $\alpha \in (0, \alpha_0)$, seja*

$$R_\alpha : \mathcal{H}_2 \longrightarrow \mathcal{H}_1$$

*um operador contínuo (não necessariamente linear). A família $\{R_\alpha\}$ é chamada de uma **regularização** ou **uma família de operadores de regularização** (para A^\dagger) se, para todo $y \in \mathcal{D}(A^\dagger)$, existir uma regra para escolha do parâmetro $\alpha := \alpha(\delta, y^\delta)$ tal que*

$$\limsup_{\delta \rightarrow 0} \{\|R_\alpha y^\delta - A^\dagger y\| : y^\delta \in \mathcal{H}_2, \|y - y^\delta\| \leq \delta\} = 0 \quad (7.1.2)$$

é satisfeita para $\alpha := \alpha(\delta, y^\delta) \xrightarrow{\delta \rightarrow 0} 0$.

¹Nós, matemáticos, gostamos de denominar os truques, as estratégias e outros artifícios por **métodos**.

Observação 7.1. *Notemos que não estamos requerendo que a família de operadores de regularização $\{R_\alpha\}$ seja de operadores lineares. No caso em que $\{R_\alpha\}$ é linear, então dizemos que o método de regularização é linear.*

Definição 7.2. *Uma estratégia de regularização (R_α, α) é dita **convergente** se $x_\alpha := R_\alpha y$ converge para x^\dagger .*

A arte de aplicar métodos de regularização está sempre relacionada com o compromisso entre precisão e estabilidade. Ou seja, procuramos por aproximações x_α^δ de x^\dagger , que dependam continuamente dos dados com ruídos y^δ (estabilidade) e que convirjam para x^\dagger , se o nível de ruídos δ convergir para zero. Aliado a isso tudo, o parâmetro de regularização α deve ser escolhido de forma adequada. Existem basicamente duas formas de escolha do parâmetro de regularização. Estas formas de escolha para α ficarão mais claras logo abaixo.

Queremos enfatizar que um método de regularização consiste:

- a) de uma estratégia para aproximar o operador inverso A^{-1} de maneira a evitar o mal condicionamento.
- b) de uma regra para escolha de parâmetros de regularização, no sentido que, se o parâmetro de regularização é escolhido de acordo com essa regra, então a solução regularizada converge (em alguma norma) para a solução do problema, quando o nível de ruído tende para zero.
- c) do conceito de solução que estamos considerando e da topologia em que esse conceito de solução está imerso.

7.2 Resultados de Convergência

Diante ao apresentado até então, surgem as seguintes questões:

- (i) Como construir uma família de operadores de regularização?
- (ii) Como obter uma escolha de parâmetros para que um tal método de regularização convirja?

(iii) É possível obtermos alguma performance “ótima” nesse caminho?

Para responder as questões (i) e (ii) acima, apresentaremos, logo abaixo, alguns métodos de regularização, divididos em duas classes: contínuos (por exemplo o Método de Tikhonov) e iterativos (por exemplo o Método de Landweber). Estes respondem, pelo menos em parte, as primeiras duas questões.

Para responder a (iii), ou seja, para assegurar que um método de regularização aplicado ao problema inverso (1.2.1) converge a uma solução e para expressar essa convergência em termos de taxas, é necessário obtermos algumas informações *a-priori* sobre a solução exata x^\dagger ou sobre y . Essas informações *a-priori* são formuladas em termos das *condições de fonte* (*source conditions*).

Como o próprio nome sugere, uma condição de fonte é algum tipo de informação *a priori* sobre a solução do problema. Em geral, aparece na forma de uma representação da solução x^\dagger em termos da imagem do operador A^* (ou A), ou como potências da imagem do mesmo [3, 4, 15, 10, 23].

No nosso contexto, podemos dizer o seguinte:

Teorema 7.1. *Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ linear e limitado.*

i) Se $x \in \text{Im}(A^)$ e $\|Ax\| \leq \tau$, então*

$$\|x\| \leq C_1 \tau^{\frac{1}{2}}. \quad (7.2.3)$$

*ii) Se $x \in \text{Im}(A^*A)$ e $\|Ax\| \leq \tau$, então*

$$\|x\| \leq C_2 \tau^{\frac{2}{3}}. \quad (7.2.4)$$

Demonstração: *i)* De $x \in \text{Im}(A^*)$, segue que existe $y \in \mathcal{H}_2$, com $x = A^*y$ e $\|y\| \leq C_1^{\frac{1}{2}}$. Assim,

$$\|x\|^2 = \langle x, x \rangle = \langle x, A^*y \rangle = \langle Ax, y \rangle \leq \|Ax\| \|y\| \leq C_1^{\frac{1}{2}} \tau.$$

ii) De $x \in \text{Im}(A^*A)$, segue que existe $z \in \mathcal{H}_1$, com $x = A^*Az$ e $\|z\| \leq C_2^{\frac{1}{2}}$. Portanto,

$$\begin{aligned} \|x\|^2 &= \langle x, A^*Az \rangle = \langle Ax, Az \rangle \leq \tau \|Az\| = \tau (\langle Az, Az \rangle)^{\frac{1}{2}} \\ &= \tau (\langle z, x \rangle)^{\frac{1}{2}} \leq \tau (\|z\| \|x\|)^{\frac{1}{2}}. \quad \blacksquare \end{aligned}$$

Uma interpretação da condição i) no Teorema acima é a de que a inversa de $A|_{K_C} : K_C \rightarrow A(K_C)$ é contínua em y , onde $K_C := \{x : x = A^*y, \|y\| \leq C\}$. Note que a limitação em ii) é melhor pois estamos assumindo mais condições na solução x .

Exemplo 7.1 (Diferenciação nos dados). *Consideremos o operador*

$$\begin{aligned} A : L^2[0, 1] &\rightarrow L^2[0, 1] \\ x &\mapsto (Ax)(t) = \int_0^t x(s) ds = y(t), \quad t \in [0, 1]. \end{aligned} \quad (7.2.5)$$

Suponha que tenhamos uma informação a priori de que $x \in \{v \in AC[0, 1] : v(1) = 0 \text{ e } v' \in L^2[0, 1]\}$ e que a norma da primeira derivada seja estimada por:

$$\|x'\|_{L^2[0,1]}^2 \leq C.$$

Claramente, o operador adjunto é dado por

$$\begin{aligned} A^* : L^2[0, 1] &\rightarrow L^2[0, 1] \\ y &\mapsto (A^*y)(s) = - \int_s^1 y(r) dr. \end{aligned} \quad (7.2.6)$$

Assim, se $\|Ax\| \leq \tau$, então

$$\begin{aligned} \|x\|_{L^2[0,1]}^2 &= - \int_0^1 x'(t) \int_0^t x(s) ds dt = - \int_0^1 x'(t) Ax(t) dt \\ &\leq \|x'\| \|Ax\| \leq C\tau. \end{aligned}$$

Existem basicamente duas formas de escolha do parâmetro de regularização: uma escolha *a-priori*, ($\alpha = \alpha(\delta)$) ou uma escolha *a-posteriori* ($\alpha = \alpha(\delta, y^\delta)$), dependendo do método de regularização utilizado. Passaremos a estudar cada um dos casos com um pouco mais de detalhes.

Uma pergunta que pode ser feita é: Existe uma escolha de parâmetros de regularização α que dependa somente dos dados y^δ e não dependa do nível de ruídos δ ? A resposta é a seguinte:

Proposição 7.1. *Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ um operador linear limitado. Suponha que existe $\{R_\alpha\}$ uma regularização para A^\dagger , com uma escolha de parâmetros α que dependa somente de y^δ mas não dependa de δ , tal que o método de regularização (R_α, α) seja convergente, para todo $y \in \mathcal{D}(A^\dagger)$. Então A^\dagger , é limitado.*

Demonstração: Seja $\alpha = \alpha(\delta)$. Pela definição de método de regularização convergente, temos que

$$\limsup_{\delta \rightarrow 0} \{ \|R_{\alpha(y^\delta)} y^\delta - A^\dagger y\| : y^\delta \in \mathcal{H}_2, \|y - y^\delta\| \leq \delta \} = 0, \quad (7.2.7)$$

e, portanto, $R_{\alpha(y)} y = A^\dagger y$ para todo $y \in \mathcal{D}(A^\dagger)$. Logo, de (7.2.7), segue que para qualquer sequência $\{y_n\} \in \mathcal{D}(A^\dagger)$ que converge para $y \in \mathcal{D}(A^\dagger)$,

$$A^\dagger y_n = R_{\alpha(y_n)} y_n \rightarrow A^\dagger y,$$

e, assim, A^\dagger é limitado no $\mathcal{D}(A^\dagger)$. Pelo Teorema 6.3, temos que $\mathcal{D}(A^\dagger) = \mathcal{H}_2$. ■

A Proposição 7.1 nos ensina que uma escolha de parâmetros para um método de regularização convergente para problemas mal postos deve, obrigatoriamente, levar em conta o nível de ruídos δ .

7.2.1 Escolha *a priori* do Parâmetro de Regularização

Uma escolha do parâmetro de regularização *a-priori* ($\alpha = \alpha(\delta)$) é, teoricamente, feita antes de qualquer cálculo numérico na tentativa de resolver o problema inverso. Desta forma, não depende do cálculo atual, digamos o resíduo $\|Ax_\alpha^\delta - y^\delta\|$.

Por enquanto, para o caso de A ser um operador linear, temos:

Proposição 7.2. *Para todo $\alpha > 0$, suponha que R_α um operador contínuo. Então a família R_α é uma regularização para A^\dagger se*

$$R_\alpha \rightarrow A^\dagger \text{ pontualmente no } \mathcal{D}(A^\dagger) \text{ quando } \alpha \rightarrow 0. \quad (7.2.8)$$

Neste caso, existe, para cada $y \in \mathcal{D}(A^\dagger)$, uma escolha a priori para α tal que $x_\alpha := R_\alpha y$ convirja para uma solução de $Ax = y$.

Demonstração: Fixemos $y \in \mathcal{D}(A^\dagger)$ qualquer. Pela hipótese, existe um função monótona $\gamma : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ com $\lim_{\varepsilon \rightarrow 0} \gamma(\varepsilon) = 0$ tal que, para todo $\varepsilon > 0$,

$$\|R_{\gamma(\varepsilon)}y - A^\dagger y\| \leq \frac{\varepsilon}{2}.$$

Da continuidade de $R_{\gamma(\varepsilon)}$, para cada $\varepsilon > 0$ existe um $\rho(\varepsilon)$ tal que

$$\|z - y\| \leq \rho(\varepsilon) \quad \Rightarrow \quad \|R_{\gamma(\varepsilon)}z - R_{\gamma(\varepsilon)}y\| \leq \frac{\varepsilon}{2}.$$

Sem perda de generalidade, podemos assumir que a função $\rho(\varepsilon)$ é estritamente monótona, contínua e $\lim_{\varepsilon \rightarrow 0} \rho(\varepsilon) = 0$. Portanto, a inversa ρ^{-1} existe na imagem de ρ , é estritamente monótona, contínua e satisfaz $\lim_{\delta \rightarrow 0} \rho^{-1}(\delta) = 0$.

Definimos $\alpha := \gamma(\rho^{-1}(\delta))$. Notemos que α é monótona e satisfaz $\lim_{\delta \rightarrow 0} \alpha(\delta) = 0$. Como $\|y - y^\delta\| \leq \delta$, uma simples desigualdade triangular mostra que

$$\|R_{\alpha(\delta)}y^\delta - A^\dagger y\| \leq \varepsilon.$$

Isso demonstra a nossa afirmação. ■

Quando estamos resolvendo um problema inverso, temos que ter sempre em mente o quanto queremos, de fato, aproximar a solução do problema inverso. Vamos ser mais específicos.

Suponha que $Im(A)$ não é fechada, assim, A^\dagger é não limitada (pelo Teorema 6.3). Seja $\{R_{\alpha(\delta)}\}$ uma estratégia linear de regularização para o problema $Ax = y^\delta$. Seja $y^\delta \in \mathcal{H}_2$ satisfazendo $\|y - y^\delta\| \leq \delta$, então

$$\begin{aligned} \|R_{\alpha(\delta)}y^\delta - A^\dagger y\| &\leq \|x_{\alpha(\delta)} - A^\dagger y\| + \|R_{\alpha(\delta)}y^\delta - x_{\alpha(\delta)}\| \\ &= \|x_{\alpha(\delta)} - A^\dagger y\| + \|R_{\alpha(\delta)}y^\delta - R_{\alpha(\delta)}y\| \quad (7.2.9) \\ &\leq \|x_{\alpha(\delta)} - A^\dagger y\| + \delta \|R_{\alpha(\delta)}\|. \end{aligned}$$

Notamos que, na estimativa (7.2.9) temos dois efeitos competindo. O primeiro termo é o efeito da regularização: quanto menor for $\alpha(\delta)$,

melhor é a solução aproximada $x_{\alpha(\delta)}$ para x^\dagger . O segundo termo é o efeito da má-colocação do problema inverso: quando $\alpha(\delta) \rightarrow 0$, $\|R_{\alpha(\delta)}\| \rightarrow \infty$ (pelo Exercício 7.3). Faça uma comparação entre a estimativa (7.2.9) e a estimativa (1.3.3). Qual o papel de h em (1.3.3)?

O caso em que temos igualdade na equação (7.2.9) é, sem sombra de dúvidas, o **piores caso**. Mas, temos que trabalhar com a hipótese de que o pior caso aconteça. Assim, a importância de escolher α de forma apropriada (e positivo) fica evidente, mesmo que tenhamos que abrir mão de aproximar, o tanto quanto queríamos, o problema original.

Algumas técnicas de escolha *a priori* para o parâmetro α são bem conhecidas e amplamente usadas. Uma delas é a chamada de **curva L**. Não entraremos em detalhes aqui. Para interessados sugerimos [4, 10].

Assim, para obtermos uma estratégia ótima de solução para problemas inversos com uma escolha *a priori* do parâmetro de regularização, temos que ser capazes de escolher valores apropriados para α de forma a balancearmos a equação (7.2.9). Uma maneira de construir uma família de regularização adequadamente será apresentada logo mais na Seção 7.3 em termos dos valores espectrais. Essa técnica está baseada na construção das chamadas *Funções Filtro*, [3, 4, 23], ou nas funções de truncamento. Nos Capítulos 8 e 9 apresentaremos outras formas de escolher o parâmetro de regularização.

Da estimativa (7.2.9), segue que:

Proposição 7.3. *Seja $\{R_\alpha\}$ uma estratégia linear de regularização. Para cada $y \in \mathcal{D}(A^\dagger)$, seja α uma escolha a priori para o parâmetro de regularização. Então, (R_α, α) é uma estratégia de regularização convergente se e só se*

$$\lim_{\alpha \rightarrow 0} \alpha(\delta) = 0 \quad e \quad \lim_{\delta \rightarrow 0} \delta \|R_{\alpha(\delta)}\| = 0.$$

7.2.2 Escolha *a posteriori* do Parâmetro de Regularização

Uma escolha *a posteriori* do parâmetro de regularização é feita via uma comparação entre o resíduo (ou a **discrepância**), i.e.,

$$\|Ax_\alpha^\delta - y^\delta\| \leq \tau\delta \tag{7.2.10}$$

e o nível de ruídos δ . Esta escolha é chamada de *Princípio da Discrepância*.

Observação 7.2. *Uma motivação heurística para tal escolha é a seguinte: Queremos resolver $Ax = y$, mas só conhecemos y^δ com $\|y - y^\delta\| \leq \delta$. Assim, pedir que, para uma solução aproximada x_α^δ , com $\|Ax_\alpha^\delta - y^\delta\| < \delta$, não faz sentido. Ou seja, o melhor que podemos esperar é que tenhamos um resíduo da ordem de δ .*

Voltando a analisar a equação (7.2.9), vemos que quanto menor o parâmetro de regularização, pior é a estabilidade. Assim, devemos escolher α *a posteriori* o **maior** possível tal que a *discrepância* (7.2.10) seja satisfeita. Notamos que, se $\delta = 0$, então o princípio da discrepância nunca é atingido. Neste caso, tomamos $\alpha := \alpha(y, \delta = 0) = +\infty$. Disto segue o Teorema:

Teorema 7.2. *Um método de regularização (R_α, α) , onde $\alpha := (\delta, y^\delta)$ é escolhido de acordo com o princípio da discrepância (7.2.10), é convergente $\forall y \in Im(A)$.*

Demonstração: Veja [10, Teorema 4.17]. ■

7.3 Regularização por Truncamento

Nesta seção construiremos o primeiro **método de regularização** específico destas notas. Nos deteremos aos detalhes do caso particular em que o operador A é linear e compacto. Para o caso em que A é um operador linear limitado qualquer, os resultados são muito bem apresentados em [10].

Vamos direto ao problema a ser considerado.

Seja A um operador linear e compacto com um sistema singular (σ_n, e_n, f_n) . Considere o problema de encontrar x na equação

$$Ax = y^\delta.$$

Como y^δ pode não pertencer a $Im(A)$, temos que nos contentar em encontrar uma melhor aproximação $x^\dagger = A^\dagger y^\delta$ da solução exata x . Ou,

equivalentemente, encontrar entre as soluções das equações normais

$$A^*Ax = A^*y^\delta,$$

a solução x^\dagger que tem a menor norma.

Do Teorema de Picard 6.9, temos que uma possibilidade de solução seria

$$x^\delta = \sum_{j=1}^{\infty} \sigma_j^{-2} \langle A^*y^\delta, e_j \rangle e_j. \quad (7.3.11)$$

Vimos na Observação 6.4 que usar a equação (7.3.11) não é uma boa alternativa para calcular uma aproximação para a solução x^\dagger do problema inverso acima, uma vez que A^*A também é compacto e, assim, $\sigma_j \rightarrow \infty$.

Para $\alpha \in (0, \alpha_0)$ e $\lambda \in [0, \|A\|^2]$, defina a função (contínua a direita)

$$f_\alpha(\lambda) = \begin{cases} \frac{1}{\lambda}, & \text{se } \lambda \geq \alpha \\ 0, & \text{se } \lambda < \alpha. \end{cases}$$

Portanto, pelo Teorema da Aplicação Espectral,

$$\begin{aligned} x_\alpha^\delta &:= f_\alpha(A^*A)A^*y^\delta = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle A^*y^\delta, e_j \rangle e_j = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle y^\delta, Ae_j \rangle e_j \\ &= \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle y^\delta, \sigma_j f_j \rangle e_j = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-1} \langle y^\delta, f_j \rangle e_j. \end{aligned} \quad (7.3.12)$$

A definição de x_α^δ como na equação (7.3.12) pode ser vista como uma versão *truncada* da expansão em valores singulares (6.3.12).

Definição 7.3. *O método dado pela equação (7.3.12) é chamado de expansão truncada em valores singulares.*

Observação 7.3. *Calcular x_α^δ por $x_\alpha^\delta = A_\alpha^\dagger y^\delta$ é um método de projeção sobre os auto-espacos de A^*A . Ainda, o nível de truncamento α , que decide quando os valores singulares são trocados por 0, age como uma escolha a priori do parâmetro de regularização.*

Teorema 7.3. *O método de expansão truncada em valores singulares é um método de regularização.*

Demonstração: Notemos que A_α satisfaz a Proposição 7.2. ■

7.4 Exercícios

Exercício 7.1. *Prove que o adjunto do operador A definido pela equação (7.2.5) é o operador dado pela equação (7.2.6).*

Exercício 7.2. *Suponha A um operador linear limitado com $\text{Im}(A)$ fechada. Construa um operador de regularização para o problema inverso $Ax = y$. Que regra foi usada na construção do parâmetro de regularização? (Sugestão: Usar o Teorema 6.3)*

Observação 7.4. *O exercício acima mostra que, no caso em que o operador A possui imagem fechada, a **pseudo-inversa** é uma possível regularização para o problema inverso.*

Exercício 7.3. *Mostre que se uma família de operadores lineares converge uniformemente para um operador linear, então este limite é contínuo. Use isto para justificar porquê não podemos requerer a convergência uniforme na Proposição 7.2. Ainda, mostre que, no caso em que $\text{Im}(A)$ não é fechada, então $\|R_\alpha\| \xrightarrow{\alpha \rightarrow 0} +\infty$.*

Exercício 7.4. *Mostre que x_α^δ dado em (7.3.12) pode ser calculado como $x_\alpha^\delta = A_\alpha^\dagger y^\delta$, onde A_α é um operador com imagem de dimensão finita definido por*

$$A_\alpha x = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j \langle x, e_j \rangle f_j.$$

Exercício 7.5. *Preencha os detalhes da demonstração do Teorema 7.3.*

Capítulo 8

Regularização de Tikhonov

Com a mesma filosofia da teoria geral de regularização para problemas inversos, a **regularização de Tikhonov** é um compromisso entre precisão e estabilidade. O objetivo deste capítulo é estabelecermos tal compromisso. Nestas notas, trataremos exclusivamente do caso em que o operador A é linear. O estudo do método de regularização de Tikhonov para problemas não lineares pode ser encontrado em [4, 8, 10, 36].

8.1 Convêrgencia

Nesta seção, consideraremos o caso em que o operador A é linear e limitado. Com essa hipótese, estamos interessados em encontrar, de forma estável, uma aproximação para a solução do problema inverso

$$Ax = y^\delta,$$

para medidas conhecidas do erro $\|y - y^\delta\| \leq \delta$.

Assim, como estudamos nos Capítulos 7, uma solução regularizada requer uma estratégia mais adequada que tentar resolver as equações normais

$$A^*Ax = A^*y^\delta. \tag{8.1.1}$$

Ou, de forma equivalente, encontrar um mínimo para o problema variacional de mínimos quadrados

$$J(x) = \frac{1}{2} \|Ax - y^\delta\|^2.$$

Lembramos que tudo o que queremos é "inverter" o operador A de maneira estável. Mais que isso, não queremos errar muito ao fazer essa inversão, isto é, queremos manter o resíduo $\|Ax - y^\delta\|$ controlado. Veja a estimativa (7.2.9). Então, formalmente, gostaríamos de inverter de forma estável o operador A^*A . Pelo Teorema da Aplicação Espectral 6.8, isto equivale a calcular $g(A^*A)$, onde $g(\lambda) = \frac{1}{\lambda}$, $\lambda \neq 0$. Portanto, uma solução aproximada para a equação (8.1.1) é dada por $x^\delta = g(A^*A)A^*y^\delta$.

Como o espectro de A^*A pode estar muito próximo de ZERO e até mesmo conter o 0, uma estratégia de calcular x^δ , como $x^\delta = g(A^*A)A^*y^\delta$ não é possível, ou, se é possível, é muito instável.

Qual é a estratégia para obter uma solução aproximada de (1.2.1) de maneira estável? Afastar o espectro de A^*A de zero.

Seja $0 < \alpha \in [0, \alpha_0]$, defina

$$f_\alpha(\lambda) := g(\lambda + \alpha) = \frac{1}{\lambda^2 + \alpha}. \quad (8.1.2)$$

A função $f_\alpha(\cdot)$ é dita ser a *Função Filtro* para o método de regularização de Tikhonov.

Do Teorema da Aplicação Espectral, temos que

$$f_\alpha(A^*A) = (A^*A + \alpha I)^{-1}. \quad (8.1.3)$$

Segue do Exercício 8.8 que a escolha de x_α^δ , da forma

$$x_\alpha^\delta = (A^*A + \alpha I)^{-1} A^*y^\delta \quad (8.1.4)$$

é uma solução regularizada, definida via a equação linear

$$(A^*A + \alpha I)x_\alpha^\delta = A^*y^\delta. \quad (8.1.5)$$

Esta última equação pode ser pensada como uma regularização para as equações normais (8.1.1). Este método é chamado de **Regularização de Tikhonov**¹.

¹Muitas vezes é chamado de *Regularização de Tikhonov-Phillips*.

Exercício 8.1. *Seja A um operador linear e compacto entre espaços de Hilbert com um sistema singular dado por (σ_j, e_j, f_j) . Mostre que a solução regularizada x_α^δ na equação (8.1.4) tem a forma*

$$x_\alpha^\delta = \sum_{j=1}^{\infty} \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle y^\delta, f_j \rangle e_j. \quad (8.1.6)$$

*Use o mesmo raciocínio para mostrar que $x^\delta = g(A^*A)A^*y^\delta$ satisfaz*

$$x^\delta = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \langle y^\delta, f_j \rangle e_j. \quad (8.1.7)$$

Observação 8.1. *Fazemos uma comparação entre (8.1.7) e (8.1.6) e então vemos claramente o resultado de estabilidade da equação (8.1.6): o erro em $\langle y^\delta, f_j \rangle$ é propagado com um fator de $\frac{\sigma_j}{\sigma_j^2 + \alpha}$, o qual é sempre limitado quando $j \rightarrow \infty$. Em (8.1.7), o fator de propagação é $\frac{1}{\sigma_j}$.*

Como nem sempre estamos trabalhando com operadores compactos e, mesmo se esse for o caso, a determinação de um sistema singular de um operador é uma tarefa muito custosa (tanto do ponto de vista numérico quanto matemático). Seria ideal termos uma outra forma de determinar uma solução pela regularização de Tikhonov. Isto é dado pelo teorema abaixo que trata de uma versão variacional da regularização de Tikhonov:

Teorema 8.1. *Seja x_α^δ como na equação (8.1.4). Então x_α^δ é o único minimizador do **funcional de Tikhonov***

$$J_\alpha(x) := \|Ax - y^\delta\|^2 + \alpha\|x\|^2. \quad (8.1.8)$$

Demonstração: Para $\alpha > 0$, o funcional J_α é estritamente convexo e coercivo. Assim, J_α possui um único minimizador que deve satisfazer a condição necessária (e neste caso também suficiente) de primeira ordem

$$J'_\alpha(x).h = 0 \quad \text{para todo } h \in \mathcal{H}_1. \quad (8.1.9)$$

Disto, segue que

$$\begin{aligned} 0 &= J'_\alpha(x) \cdot h = 2\langle Ax - y^\delta, Ah \rangle + 2\alpha\langle x, h \rangle \\ &= 2\langle A^*Ax - A^*y^\delta + \alpha Ix, h \rangle \end{aligned} \quad (8.1.10)$$

para todo $h \in \mathcal{H}_1$. Portanto, (8.1.10) é equivalente a (8.1.4). ■

Observação 8.2. *Notemos que, qualquer minimizador do funcional de Tikhonov (8.1.8) pertence a $\mathcal{N}(A)^\perp$. Com efeito, pois, caso contrário, poderíamos fazer a segunda parte do funcional crescer, mantendo a primeira parte constante.*

O parâmetro α no funcional (8.1.8) é o parâmetro de regularização. Minimização em (8.1.8) é um compromisso entre minimizar a norma do resíduo $\|Ax - y^\delta\|$ e tomar o tamanho do **termo de penalização** $\|x\|$ pequeno e, assim, forçar a estabilidade. A escolha apropriada do parâmetro α é ainda um problema e deve ser feita *a priori*. Como α é o parâmetro que estabelece este compromisso, muitos autores sugerem sua escolha através da chamada **curva L**. Um boa referência para a construção da curva L é [4].

Observação 8.3. *Notemos que a definição de x_α^δ como em (8.1.4) só tem sentido para operadores lineares. Mas, o problema de minimização (8.1.8) pode ser formulado para operadores não-lineares [4, 8, 10, 36].*

8.2 Semi-Convergência

A definição da solução regularizada, pela minimização do funcional de Tikhonov (8.1.8), nos dá diretamente resultados de convergência e estabilidade, como:

Teorema 8.2. *Seja x_α^δ definida por (8.1.4), $y \in \text{Im}(A)$ com $\|y - y^\delta\| \leq \delta$. Se $\alpha := \alpha(\delta)$ é tal que*

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0 \quad e \quad \lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0, \quad (8.2.11)$$

então

$$\lim_{\delta \rightarrow 0} x_{\alpha(\delta)}^\delta = A^\dagger y. \quad (8.2.12)$$

Demonstração: Seja $\delta_n \rightarrow 0$ qualquer. Definimos $\alpha_n := \alpha(\delta_n)$ e $x_n := x_{\alpha_n}^{\delta_n}$. Seja J_n o funcional de Tikhonov (8.1.8) com $\alpha = \alpha_n$ e x_n seu correspondente minimizante (que existe e é único, pelo Teorema 8.1). Seja $x^\dagger := A^\dagger y$. Então, por definição de x_n , temos

$$\begin{aligned} \alpha_n \|x_n\|^2 \leq J_n(x_n) &\leq J_n(x^\dagger) = \|Ax^\dagger - y^{\delta_n}\|^2 + \alpha_n \|x^\dagger\|^2 \\ &\leq \delta_n^2 + \alpha_n \|x^\dagger\|^2, \end{aligned}$$

e, assim,

$$\|x_n\|^2 \leq \frac{\delta_n^2}{\alpha_n} + \|x^\dagger\|^2. \quad (8.2.13)$$

Portanto, por hipótese, $\{x_n\}$ é uniformemente limitada. Pelo Teorema de Banach-Alaoglu [23], $\{x_n\}$ possui uma subsequência que converge fraco² para $z \in \mathcal{H}$. Como A é linear e limitado,

$$Ax_{n_k} \rightharpoonup Az. \quad (8.2.14)$$

Novamente, a definição de x_{n_k} implica que

$$\|Ax_{n_k} - y^{\delta_{n_k}}\|^2 \leq J_{n_k}(x_{n_k}) \leq \delta_{n_k}^2 + \alpha_{n_k} \|x^\dagger\|^2 \xrightarrow{k \rightarrow \infty} 0.$$

Segue de (8.2.14) que

$$Az = y. \quad (8.2.15)$$

Da Observação 8.2, temos que $x_n \in \mathcal{N}(A)^\perp$ e, assim, $z \in \mathcal{N}(A)^\perp$ (Prove). Pelo Teorema 6.2 e sua demonstração, obtemos $z = x^\dagger$. Assim, $x_{n_k} \rightharpoonup x^\dagger$. Aplicando o mesmo argumento para todas as subsequências, obtemos que

$$x_n \rightharpoonup x^\dagger. \quad (8.2.16)$$

²Uma sequência em um espaço de Hilbert é **fracamente convergente** para $z \in \mathcal{H}$ se para todo $h \in \mathcal{H}$ temos $\langle x_n, h \rangle \rightarrow \langle z, h \rangle$. Notação: $x_n \rightharpoonup z$.

Afirmação: $x_n \rightarrow x^\dagger$.

Assuma que exista $\varepsilon > 0$ tal que $\|x_{n_k}\| \leq \|x^\dagger\| - \varepsilon$. Então, o mesmo argumento de extração de subsequências acima implica que $\|z\| \leq \|x^\dagger\| - \varepsilon$, contradizendo (8.2.16). Logo,

$$\liminf \|x_n\| \geq \|x^\dagger\|. \quad (8.2.17)$$

De (8.2.13), temos

$$\liminf \|x_n\| \leq \|x^\dagger\|. \quad (8.2.18)$$

Juntando as últimas três estimativas, temos que $x_n \rightarrow x^\dagger$. \blacksquare

8.2.1 Taxas de convergência

Segue da definição de solução regularizada pelo método de Tikhonov que

$$\|x_\alpha^\delta - x^\dagger\| \leq \sup_{\lambda \in \Sigma(A)} |\lambda f_\alpha(\lambda)| \|y - y^\delta\| \leq \frac{\delta}{\sqrt{\alpha}}. \quad (8.2.19)$$

Assim, se $\alpha \sim \delta$, obtemos a seguinte ordem de convergência

$$\|x^\dagger - x_\alpha^\delta\| = \mathcal{O}(\sqrt{\delta}). \quad (8.2.20)$$

8.3 Regularização de Tikhonov e a Inversa Generalizada

De acordo com a definição variacional, $A^\dagger y$ é o vetor $x \in \mathcal{H}_1$ de menor norma que minimiza o funcional $\|Ax - y\|$. Por outro lado, a ideia da regularização de Tikhonov é minimizar ambos: o funcional, $\|Ax - y\|$, e a norma, $\|x\|$, através da minimização do funcional de Tikhonov (8.1.8).

Pelo Teorema 8.1, a única solução x_α satisfaz

$$x_\alpha = (A^*A + \alpha I)^{-1} A^* y.$$

Segue, do Teorema 8.2, que

$$\lim_{\alpha \rightarrow 0} x_\alpha = A^\dagger y.$$

Com esse resultados obtemos:

Corolário 8.1. *Seja A um operador linear e limitado satisfazendo a hipótese 6.1. Então,*

$$A^\dagger = \lim_{\alpha \rightarrow 0} (A^*A + \alpha I)^{-1} A^*,$$

uniformemente no espaço $\mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ dos operadores lineares limitados

Para obtermos um limitante para o erro com este método, considere $\lambda \in \Sigma(A^*A)$. Então temos que

$$|\lambda f_\alpha(\lambda) - 1| = \frac{\alpha}{\lambda + \alpha} \leq \frac{\alpha}{\|A^\dagger\|^{-2} + \alpha}.$$

Assim, segue que

$$\|(A^*A + \alpha I)^{-1} A^* - A^\dagger\| \leq \frac{\alpha \|A^\dagger\|^3}{1 + \alpha \|A^\dagger\|^2}. \quad (8.3.21)$$

Exercício 8.2. *Preencha os detalhes na demonstração da estimativa (8.3.21).*

8.4 Regularização de Tikhonov para Tomografia Computadorizada

Como vimos antes, o método de regularização de Tikhonov produz soluções estáveis e convergentes para a solução do problema $Ax = y^\delta$, como função do nível de ruídos nos dados.

Nesta seção, adaptaremos a teoria desenvolvida anteriormente para o problema da TC. Para tal, assumiremos que os dados p_i^δ são obtidos por medidas, tal que, $p_j = \mathcal{R}_j f$ na norma de $L^2(\Omega)$, para um conjunto

de raios- X com $j = 1, \dots, N$. De acordo com a Subseção 1.5.2, temos que

$$\mathcal{R}_j f = (A(:, j), f) \quad j = 1, \dots, N,$$

onde (\cdot, \cdot) representa o produto interno em $L^2(\Omega)$ e $A(:, j)$ representa a coluna j da matriz A , dada na Subseção 1.5.2.

Lema 8.1. *Definindo, $\mathcal{R} : L^2(\Omega) \rightarrow \mathbb{R}^N$ por*

$$\mathcal{R} = (\mathcal{R}_1, \dots, \mathcal{R}_N)^T \quad (8.4.22)$$

temos que o operador adjunto $\mathcal{R}^ : \mathbb{R}^N \rightarrow L^2(\Omega)$ é dado por*

$$\mathcal{R}^* r = \sum_{j=1}^N A(:, j) r_j. \quad (8.4.23)$$

Demonstração: Deixamos como exercício para o leitor. ■

Como \mathcal{R} é linear, o funcional de Tikhonov (8.1.8) consiste em minimizar

$$J(f) := \|\mathcal{R}f - p^\delta\|^2 + \alpha \|f\|^2.$$

Pelo Teorema 8.1, a única solução regularizada para TC, dada pelo método de regularização de Tikhonov, é

$$f_\alpha = \mathcal{R}^* r \quad (S + \alpha I)r = p^\delta, \quad (8.4.24)$$

onde $S = \mathcal{R}\mathcal{R}^* = (A(:, k), A(:, j))$ para $k, j = 1, \dots, N$.

8.5 Exercícios

Exercício 8.3. *Dê um exemplo de um operador linear que possui espectro não-vazio e que não possui autovalores. Isso pode acontecer no caso em que A é uma matriz, isto é, em dimensão finita?*

Exercício 8.4. *Seja A um operador linear limitado entre espaços de Hilbert. Mostre que o espectro de A^*A é real e positivo.*

Exercício 8.5. Dê um exemplo de um operador linear limitado A tal que 0 seja um elemento do espectro de A . Existe operadores auto-adjuntos tais que 0 seja um autovalor? Se afirmativo, dê exemplos. Sugestão: procure exemplos nos espaços de seqüências.

Exercício 8.6. Mostre que se λ é um autovalor de A , então λ^2 é um autovalor de A^*A . Seja A seja uma matriz inversível. Mostre que $1/\lambda$ é um autovalor de A^{-1} .

Exercício 8.7. Suponha que A^*A seja inversível. Considere $x^\dagger = g(A^*A)A^*y$ e $x^\delta = g(A^*A)A^*y^\delta$. Use o exercício acima para mostrar que

$$\|x^\dagger - x^\delta\|^2 \leq (g(A^*A))^2 \|A^*\|^2 \|y - y^\delta\|^2.$$

O que acontece com $\|x^\dagger - x^\delta\|$ como uma função dos autovalores de A^*A , para dados com ruído, satisfazendo $\|y - y^\delta\| \leq \delta$.

Exercício 8.8. Seja A um operador linear e limitado entre espaços de Hilbert. Mostre que, para todo $0 < \alpha \in \mathbb{R}$, o operador $A^*A + \alpha I$ é linear, limitado, injetivo e sobrejetivo e, assim, $(A^*A + \alpha I)^{-1}$ existe e é limitada. Sugestão: Use o Teorema do Gráfico Fechado, [24], para mostrar a limitação da inversa.

Exercício 8.9. Um funcional $J : \mathcal{H} \rightarrow \mathbb{R}$ é dito ser **coercivo** se $\lim_{\|x\| \rightarrow \infty} J(x) = +\infty$. Mostre que, se J é convexo e coercivo, então, este atinge um mínimo. Ainda, se J é estritamente convexo e coercivo, mostre que o mínimo é único.

Exercício 8.10. Mostre que dado um funcional J como no exercício acima, então a condição de otimalidade de primeira ordem é necessária e também suficiente para x ser um minimizante de J .

Exercício 8.11. Verifique a desigualdade em (8.2.19) no caso em que A é compacto e com um sistema singular dado por (σ_j, e_j, f_j) .

Exercício 8.12. Demonstre o Corolário 8.1. Sugestão: Use o que você sabe da função $f_\alpha(x) = (\alpha + x)^{-1}$, $\alpha > 0$, e o Teorema da Aplicação Espectral.

Exercício 8.13. *Mostre que $S = \mathcal{R}\mathcal{R}^* = (A(:, k), A(:, j))$ para $k, j = 1, \dots, N$.*

Exercício 8.14. *Use o método de regularização de Tikhonov para encontrar uma solução aproximada para o problema do Exercício 1.2. Como deve ser escolhido α nesse caso?*

Capítulo 9

Regularização por Métodos Iterativos

Como vimos no Capítulo 7, solucionar o problema inverso (1.2.1) de forma estável, isto é, de forma que condiz com a realidade dos dados obtidos, implica em encontrar uma forma de aproximar o operador inverso por uma família de operadores contínuos. Ou seja, necessitamos de um método de **regularização**.

Uma alternativa para regularização de operadores são os *métodos iterativos de regularização*. A vantagem de tais métodos é que eles possuem propriedades regularizantes [21, 23, 4, 10].

Gauss demonstrou que a melhor maneira de determinar um parâmetro desconhecido de uma equação do tipo (1.2.1), é minimizando a soma dos quadrados dos resíduos¹, isto é,

$$\min_{x \in \mathcal{H}_1} \frac{1}{2} \|A(x) - y\|^2. \quad (9.0.1)$$

Assumindo algumas propriedades do operador A , podemos provar que o minimizador de (9.0.1), caso exista, deve satisfazer a condição necessária de primeira ordem

$$A'(x)^* A(x) = A'(x)^* y. \quad (9.0.2)$$

¹Tal método é conhecido hoje em dia como **método de Mínimos Quadrados**.

Uma possibilidade para encontrar uma solução de (9.0.2) é interpretá-la como uma iteração de ponto fixo

$$x_{k+1} = \Phi(x_k) \quad (9.0.3)$$

para o operador

$$\Phi(x) = x + A'(x)^*(y - A(x)). \quad (9.0.4)$$

Em muitos exemplos práticos (e até mesmo teóricos) é quase impossível verificar analiticamente quando o operador Φ é contrativo ou não. No contexto de problemas não-lineares, esta dificuldade é ainda maior, [21, 10].

Muitos métodos iterativos para resolver (1.2.1) são baseados na solução da equação normal (9.0.2), via sucessivas iterações, partindo de um chute inicial x_0 . Observamos que tal chute, em particular para problemas não-lineares, costuma conter informações *a priori* sobre a solução do problema.

Nesse capítulo, nos dedicaremos a alguns métodos do tipo gradiente, mais especificamente, ao método de Landweber, ao método de descida máxima e algumas estratégias do tipo Kaczmarz (ART e Landweber Kaczmarz).

Neste manuscrito, nos deteremos ao estudo de métodos iterativos, considerando o caso em que o operador A é linear. O caso em que A é não linear possui tratamento detalhado em [8, 10, 21].

9.1 Método de Landweber

Nessa seção, assumiremos que o operador A na equação (1.2.1) é linear e limitado. Com essa hipótese, uma maneira de resolvermos (9.0.2) é considerarmos a iteração

$$x_{k+1} = x_k + \gamma A^*(y - Ax_k), \quad k = 0, 1, 2, \dots \quad (9.1.5)$$

em que $\|A\|^{-2} \geq \gamma > 0$ é um parâmetro de relaxação, de forma que a iteração tenha a propriedade de descida. Esse foi o ponto de partida de Landweber [25], em 1951, ao estudar equações integrais de primeira

espécie, quando propôs o método que hoje leva o nome de **Método de Regularização de Landweber**.

No caso de dados com ruído y^δ , denotando as iterações por x_k^δ , chegamos a *iteração de Landweber*

$$x_{k+1}^\delta = x_k^\delta + A^*(y^\delta - Ax_k^\delta). \quad (9.1.6)$$

Observação 9.1. *Notemos que a equação (9.1.6) é equivalente a pré-multiplicar a equação $Ax = y^\delta$ por $\gamma^{\frac{1}{2}}$ e, então, iterar como em (9.1.5). Dada a limitação de γ , sem perda de generalidade, podemos supor que $\|A\| \leq 1$ e iterar como em (9.1.6).*

Nesta seção, a menos que façamos menção em contrário, admitiremos que $\|A\| \leq 1$.

Abaixo, apresentaremos resultados de convergência da iteração de Landweber (9.1.6). As técnicas utilizadas na demonstração de convergência são relativamente simples e bem conhecidas de métodos iterativos para problemas diretos. Observamos ainda que obtemos a mesma iteração (9.1.6) partindo de um chute inicial $x_0 = 0$.

9.1.1 Convergência

Nesta subseção, provaremos que a sequência $\{x_k\}$ de iterados pelo método de Landweber converge para uma solução aproximada do problema (1.2.1). Para tal, temos que definir o que entendemos por uma solução aproximada. Denotaremos por $x^\dagger := A^\dagger y$, como a solução de mínimos quadrados com norma mínima para o problema (1.2.1). Aqui A^\dagger denota a pseudo-inversa de A , definida no Capítulo 7.

Começaremos dando condição necessárias e suficientes para a iteração (9.1.5) convergir.

Teorema 9.1. *Se $y \in \mathcal{D}(A^\dagger)$, então a sequência x_k gerada pela iteração de Landweber (9.1.5) converge para $x^\dagger = A^\dagger y$ quando $k \rightarrow \infty$. Se $y \notin \mathcal{D}(A^\dagger)$, então $\|x_k\| \rightarrow \infty$ quando $k \rightarrow \infty$.*

Demonstração: De forma recursiva, podemos escrever x_k em (9.1.5)

como

$$x_k = \sum_{j=0}^{k-1} (I - A^*A)^j A^* y. \quad (9.1.7)$$

Como $y \in \mathcal{D}(A^\dagger)$, então $A^*y = A^*Ax^\dagger$. Assim,

$$x^\dagger - x_k = x^\dagger - A^*A \sum_{j=0}^{k-1} (I - A^*A)^j x^\dagger = (I - A^*A)^k x^\dagger. \quad (9.1.8)$$

Definimos $r_k(\lambda) = (1 - \lambda)^k$. Como, por hipótese $\|A\| \leq 1$, segue que o espectro de A^*A é um subconjunto de $(0, 1]$. Notemos que, para $\lambda \in \Sigma(A^*A) \subset (0, 1]$, $r_k(\lambda)$ converge para zero uniformemente quando $k \rightarrow \infty$. Pelo Teorema da Aplicação Espectral [24, 37, 29],

$$x^\dagger - x_k = r_k(A^*A)x^\dagger. \quad (9.1.9)$$

Logo, $x_k \xrightarrow{k \rightarrow \infty} x^\dagger$. ■

O Teorema 9.1 nos ensina que $\{x_k\}$, gerada pela iteração de Landweber, converge para uma solução de mínimos quadrados da equação (1.2.1) quando $y \in \mathcal{D}(A^\dagger)$. Como, em geral, dados perturbados y^δ são tal que $y^\delta \notin \mathcal{D}(A^\dagger)$, então, ainda do Teorema 9.1, sabemos que a sequência x_k^δ diverge. A pergunta é: Qual é o fator de propagação destes erros?

Lema 9.1. *Sejam y, y^δ com $\|y - y^\delta\| \leq \delta$ e x_k e x_k^δ obtidos pelas respectivas iterações de Landweber (9.1.5) e (9.1.6). Então,*

$$\|x_k - x_k^\delta\| \leq \sqrt{k}\delta, \quad k \geq 0. \quad (9.1.10)$$

Demonstração: Pela linearidade de A , temos que

$$x_k - x_k^\delta = \sum_{j=0}^{k-1} (I - A^*A)^j A^*(y - y^\delta).$$

Como $\|A\| \leq 1$, segue que $(I - A^*A)$ é um operador semi-definido positivo com $\|I - A^*A\| \leq 1$. Assim,

$$\begin{aligned} \left\| \sum_{j=0}^{k-1} (I - A^*A)^j A^* \right\|^2 &= \left\| \sum_{j=0}^{k-1} (I - A^*A)^j (I - (I - A^*A)^k) \right\| \\ &\leq \left\| \sum_{j=0}^{k-1} (I - A^*A)^j \right\| \leq k \end{aligned}$$

e o Lema segue. ■

Notemos que, na presença de erro nos dados temos

$$\|A^\dagger y - x_k^\delta\| \leq \|A^\dagger y - x_k\| + \|x_k - x_k^\delta\|. \quad (9.1.11)$$

Esta é a estimativa fundamental para a iteração de Landweber.

A estimativa (9.1.11) nos mostra que o erro total possui duas componentes, um *erro de aproximação* que diminui lentamente e um *erro nos dados* que cresce na ordem de no máximo $\sqrt{k}\delta$. Isso nos leva a seguinte conclusão: Para valores de k pequenos, o erro nos dados é desprezível e a iteração parece convergir para a solução exata $A^\dagger y$. Quando $\sqrt{k}\delta$ atinge a magnitude da ordem do erro de aproximação, o erro propagado nos dados torna-se grande e a aproximação tende a piorar. Veja Figura 1.1. Assim, segue que a propriedade de regularização por métodos iterativos, para problemas mal postos, depende fortemente de um **critério de parada** que detecte a transição entre convergência e divergência do método. O índice da iteração faz o papel do parâmetro de regularização. Já, o critério de parada faz o papel da regra de escolha do parâmetro. Consequentemente, um critério de parada apropriado deve levar em conta a informação adicional do nível de ruídos δ .

Lema 9.2. *A norma do resíduo $y^\delta - Ax_k^\delta$ é sempre monótona não-crescente durante a iteração.*

Demonstração: De fato, pela definição da iteração de Landweber

$$\begin{aligned} y^\delta - Ax_k^\delta &= y^\delta - A(x_{k-1}^\delta + A^*(y^\delta - Ax_{k-1}^\delta)) \\ &= (I - A^*A)(y^\delta - Ax_{k-1}^\delta). \end{aligned}$$

Como $\|I - A^*A\| \leq 1$, o Lema segue. \blacksquare

Por outro lado, se $y^\delta \notin \mathcal{D}(A)$, pelo Teorema 9.1, a iteração x_k^δ diverge para infinito. Portanto, um resíduo pequeno não implica que a aproximação para solução é melhor. Veja a estimativa (9.1.11).

Uma alternativa para a escolha do critério de parada é o **princípio da discrepância**: a iteração é parada no índice $k_* = k(\delta, y^\delta)$ quando, pela primeira vez,

$$\|y^\delta - Ax_{k(\delta, y^\delta)}\| \leq \tau\delta, \quad \tau > 2 \text{ fixo.} \quad (9.1.12)$$

O próximo Teorema garante que, enquanto a discrepância (9.1.12) não é atingida, a aproximação para a solução não piora.

Teorema 9.2. [Monotonia] *Sejam $y \in \mathcal{D}(A)$, x^\dagger a solução de norma mínima de (1.2.1) e y^δ satisfazendo $\|y - y^\delta\| \leq \delta$. Se (9.1.12) é satisfeita, então*

$$\|x_{k+1}^\delta - x^\dagger\| \leq \|x_k^\delta - x^\dagger\|. \quad (9.1.13)$$

Demonstração: Fazemos

$$\begin{aligned} \|x_{k+1}^\delta - x^\dagger\|^2 - \|x_k^\delta - x^\dagger\|^2 &= 2\langle x_k^\delta - x^\dagger, x_{k+1}^\delta - x_k^\delta \rangle + \|x_{k+1}^\delta - x_k^\delta\|^2 \\ &= 2\langle x_k^\delta - x^\dagger, A^*(y^\delta - Ax_k^\delta) \rangle + \langle A^*(y^\delta - Ax_k^\delta), A^*(y^\delta - Ax_k^\delta) \rangle \\ &= 2\langle Ax_k^\delta \pm y^\delta - y, y^\delta - Ax_k^\delta \rangle + \langle y^\delta - Ax_k^\delta, AA^*(y^\delta - Ax_k^\delta) \rangle \\ &= 2\langle y^\delta - y, y^\delta - Ax_k^\delta \rangle - \|Ax_k^\delta - y^\delta\|^2 \\ &\quad - \langle y^\delta - Ax_k^\delta, (I - AA^*)(y^\delta - Ax_k^\delta) \rangle. \end{aligned}$$

Como $I - A^*A$ é semi-definido positivo, segue que

$$\|x_{k+1}^\delta - x^\dagger\|^2 - \|x_k^\delta - x^\dagger\|^2 \leq \|Ax_k^\delta - y^\delta\|(2\|y - y^\delta\| - \|Ax_k^\delta - y^\delta\|).$$

Como y^δ satisfaz (1.2.2) e $k < k_*$, a afirmativa segue. \blacksquare

Como comentado anteriormente, no caso de dados corrompidos por ruídos, a iteração de Landweber deve ser parada após uma quantidade finita de passos. O próximo Teorema mostra que o princípio da discrepância implica nessa importante propriedade para a iteração de Landweber.

Teorema 9.3. *Seja $\tau > 1$ em (9.1.12). Então, o princípio de discrepância determina um índice de parada $k_* = k_*(\delta, y^\delta)$, é finito para a iteração de Landweber, com $k(\delta, y^\delta) = \mathcal{O}(\delta^{-2})$.*

Demonstração: Seja $\{x_k^\delta\}$ dada por (9.1.5). Sabemos que

$$\|x^\dagger - x_k^\delta\|^2 - \|x^\dagger - x_{k+1}^\delta\|^2 \geq \|y - Ax_k^\delta\|^2.$$

Somando sobre $k = 1$ até j e levando em conta a monotonia dos resíduos dado pelo Lema 9.2, temos

$$\|x^\dagger - x_j^\delta\|^2 - \|x^\dagger - x_{j+1}^\delta\|^2 \geq \sum_{k=1}^j \|y - Ax_k^\delta\|^2 \geq k \|y - Ax_j^\delta\|^2.$$

Indutivamente, podemos escrever $y - Ax_j^\delta = (I - AA^*)^j(y - Ax_0)$, segue que

$$\|(I - AA^*)^j(y - Ax_0)\| = \|y - Ax_j^\delta\| \leq k^{-\frac{1}{2}} \|x^\dagger - x_1^\delta\|.$$

Assim,

$$\begin{aligned} \|y^\delta - Ax_j^\delta\| &= \|(I - AA^*)^j(y^\delta - Ax_0)\| \leq \|(I - AA^*)^j(y^\delta - y)\| \\ &\quad + \|(I - AA^*)^j(y - Ax_0)\| \leq \delta k^{-\frac{1}{2}} \|x^\dagger - x_1^\delta\|. \end{aligned}$$

Consequentemente, o lado direito é menor que $\tau\delta$, se $k > \frac{\|x^\dagger - x_1\|^2}{(\tau-1)^2\delta^2}$. Logo, $k(\delta, y^\delta) \leq c\delta^{-2}$, onde c só depende de τ . ■

Para obtermos taxas é preciso fazermos hipóteses sobre a solução x^\dagger . Tais hipóteses são conhecidas como **condições de fonte**. Condições de fonte mais comuns na literatura impõem que x^\dagger pertença a imagem do operador A^* ou pertença a imagem de potências do operador A^*A . Por exemplo, considere a condição de fonte

$$x^\dagger \in \text{Im}((A^*A)^\theta), \quad \theta \in \mathbb{R}_*^+. \quad (9.1.14)$$

Esta condição impõem que a solução x^\dagger pertença a imagem do operador $(A^*A)^\theta$.

Exercício 9.1. *Mostre que o operador $(A^*A)^{\frac{1}{2}}$ está bem definido. O domínio e imagem de $(A^*A)^{\frac{1}{2}}$ são subconjuntos de que espaços vetoriais? Interprete geometricamente a condição de fonte (9.1.14).*

Teorema 9.4. *Suponha a condição de fonte (9.1.14) satisfeita para $\theta = 1/2$. Se $y \in \text{Im}(A)$ e o princípio de discrepância (9.1.12) é válido, então, a iteração de Landweber possui ordem de convergência $k(\delta, y^\delta) = \mathcal{O}(\delta^{-1})$.*

Demonstração: Veja [10, Teorema 6.5]. ■

9.2 Método de Máxima Descida e a Inversa Generalizada

Agora, vamos nos ater no método conhecido como **Máxima Descida ou Steepest Descent** proposto por Cauchy por volta de 1847, para resolver sistemas de equações não-lineares. Com a ideia de representar a solução de um sistema de equações não-lineares por meio do mínimo de um funcional não negativo, Cauchy construiu uma sequência, de forma iterativa, que passa do iterado corrente para o seguinte na direção em que o funcional decresce mais rápido. Por essa propriedade, o método de Cauchy também é conhecido como método de Máxima Descida.

Aqui, apresentaremos alguns dos resultados para aproximarmos $A^\dagger y$, onde A é um operador que satisfaz a Hipótese 6.1, obtidos por Nashed [3]. Nashed aplicou o método de máxima descida juntando as equações lineares como uma única equação de operador linear $Ax = y$. Antes disso, vamos a algumas considerações iniciais.

Seja $J : \mathcal{H} \rightarrow \mathbb{R}$ um funcional não-negativo. Denotemos por $x^* \in \mathcal{H}$ um ponto tal que

$$J(x^*) = \inf\{J(x) : x \in \mathcal{H}\}. \quad (9.2.15)$$

Suponhamos que J seja Fréchet diferenciável em cada ponto de \mathcal{H} .

Dado um ponto x_0 , queremos minimizar J andando na direção em que J decresce de forma mais rápida. Assim, devemos escolher uma

direção $z \in \mathcal{H}$ tal que a derivada direcional

$$DJ(x_0, z) = \langle z, \nabla J(x_0) \rangle,$$

é a menor possível.

Pela desigualdade de Cauchy-Schwarz

$$- \|z\| \|\nabla J(x_0)\| \leq \langle z, \nabla J(x_0) \rangle DJ(x_0, z)$$

e a igualdade só acontece se z é um múltiplo positivo de $-\nabla J(x_0)$. Assim, começando em x_0 , a direção de maior descida de J é $z = -\nabla J(x_0)$. Se $\alpha > 0$ é o tamanho do passo na direção $-\nabla J(x_0)$, obtemos uma nova aproximação para x^* por

$$x_1 = x_0 - \alpha_0 \nabla J(x_0).$$

O parâmetro α_0 é escolhido de forma que x_1 minimize J sobre a reta que passa por x_0 na direção $-\nabla J(x_0)$ e, assim, deve satisfazer

$$\frac{d}{d\alpha} J(x_0 - \alpha \nabla J(x_0))|_{\alpha=\alpha_0} = 0.$$

A sequência iterativa $\{x_k\}$ gerada pelo método de máxima descida é dada por

$$x_{k+1} = x_k - \alpha_k \nabla J(x_k),$$

onde

$$\frac{d}{d\alpha} J(x_k - \alpha \nabla J(x_k))|_{\alpha=\alpha_k} = 0.$$

Suponhamos que o operador A satisfaz a Hipótese 6.1. Pelo Teorema 6.1, a solução de mínimos quadrados de $Ax = y$ é justamente o mínimo do funcional $J(x) = \frac{1}{2} \|Ax - y\|^2$. Desta forma, somos levados a aplicar o método de de descida máxima para o funcional J .

Exercício 9.2. Defina $r := \nabla J(x) = A^*(Ax - y)$ (que está bem definido pelo exercício 9.6). Mostre que o valor ótimo de α é dado por

$$\alpha = \frac{\|r\|^2}{\|Ar\|^2}.$$

Assim, a sequência $\{x_k\} \subset \mathcal{H}_1$ pelo método de descida máxima é dada por

$$\begin{aligned} x_{k+1} &= x_k - \alpha_k r_k \\ r_k &= A^*(Ax_k - y) \\ \alpha_k &= \frac{\|r_k\|^2}{\|Ar_k\|^2}. \end{aligned} \quad (9.2.16)$$

Observação 9.2. *Observamos que se $r_k = 0$, então x_k é uma solução de mínimos quadrados e o método pára em x_k . Se tivermos um critério para garantir que x_k é a solução de menor norma entre todas as soluções de mínimos quadrados, obteremos que $x_k = x^\dagger$.*

Exercício 9.3. *Mostre que se $Ar_k = 0$, então $r_k = 0$. Assim, a sequência gerada pelo algoritmo (9.2.16) está bem definida.*

Seguiremos os passos de Nashed [3], na demonstrações abaixo.

Lema 9.3. *Seja A satisfazendo a Hipótese (6.1). Então,*

$$\lim_{k \rightarrow \infty} r_k = 0.$$

Demonstração: De

$$J(x_{k+1}) = \frac{1}{2} \|Ax_k - \alpha_k Ar_k - y\|^2 = J(x_k) - \frac{1}{2} \frac{\|r_k\|^2}{\|Ar_k\|^2}.$$

Recursivamente, obtemos que

$$J(x_{k+1}) = J(x_0) - \frac{1}{2} \sum_{j=0}^k \frac{\|r_j\|^2}{\|Ar_j\|^2}.$$

Como $J(x) \geq 0$, $\forall x$, temos que

$$\|A\|^{-2} \sum_{j=0}^{\infty} \|r_j\|^2 \leq \sum_{j=0}^{\infty} \frac{\|r_j\|^2}{\|Ar_j\|^2} \leq 2J(x_0).$$

Portanto, $\|r_k\| \rightarrow 0$ quando $k \rightarrow \infty$. ■

Teorema 9.5. *Seja A satisfazendo a Hipótese 6.1. Então, a sequência gerada pelo método de descida máxima converge para uma solução de mínimos quadrados de $Ax = y$, para qualquer $x_0 \in \mathcal{H}_1$. A sequência converge a x^\dagger se e somente se, $x_0 \in \text{Im}(A^*)$.*

Demonstração: Escrevemos a iteração (9.2.16) de forma recursiva, então temos que

$$x_{k+1} = x_0 - \sum_{j=0}^k \alpha_j r_j. \quad (9.2.17)$$

Afirmção $\{x_n\}$ é uma sequência de Cauchy. De fato

$$x_m - x_n = - \sum_{j=n}^m \alpha_j r_j \in \text{Im}(A^*) = \mathcal{N}(A)^\perp. \quad (9.2.18)$$

Portanto, existe $\eta > 0$ tal que

$$\eta^2 \|x_m - x_n\|^2 \leq \langle A^* A(x_m - x_n), x_m - x_n \rangle = \|A(x_m - x_n)\|^2.$$

Assim,

$$\begin{aligned} \langle A^* A(x_m - x_n), x_m - x_n \rangle &\leq \|A^*(Ax_m - Ax_n)\| \|x_m - x_n\| \\ &\leq \eta^{-1} (\|A^*(Ax_m - y)\| + \|A^*(Ax_n - y)\|) \|A(x_m - x_n)\|. \end{aligned}$$

Como $r_j \rightarrow 0$, temos que $r_m - r_n = A^* A(x_m - x_n) \rightarrow 0$ quando $m, n \rightarrow \infty$. Como $\text{Im}(A) = \mathcal{N}(A^*)^\perp$ é fechada, segue do Teorema do Gráfico Fechado [24], que A^* possui inversa limitada quando restrito a $\text{Im}(A) = \mathcal{N}(A^*)^\perp$.

Portanto, $\|A(x_m - x_n)\| \leq M$ e

$$\eta^3 \|x_m - x_n\|^2 \leq M(r_m + r_n) \rightarrow 0 \quad m, n \rightarrow \infty.$$

Logo, $\{x_k\}$ é Cauchy e, assim, converge para algum $u \in \mathcal{H}_1$. Pela continuidade de A , temos

$$\|A^* Au - A^* y\| = \lim_{k \rightarrow \infty} \|A^* Ax_k - A^* y\| = \lim_{k \rightarrow \infty} r_k = 0$$

donde concluimos que u é uma solução de mínimos quadrados.

Para finalizar, lembremos que qualquer solução de mínimos quadrados é da forma $x^\dagger \oplus \mathcal{N}(A)$ (Teorema 6.1). Como $\text{Im}(A^*) = \mathcal{N}(A)^\perp$, segue que x^\dagger é a única solução de mínimos quadrados na $\text{Im}(A^*)$. Se $x_0 \in \text{Im}(A^*)$, então, de (9.2.17), segue que $x_k \in \text{Im}(A^*)$ para todo k . Como $\text{Im}(A^*)$ é fechada, $u = x^\dagger$.

Caso $x_0 \notin \text{Im}(A^*)$, então $x_0 = x'_0 + P_{\mathcal{N}(A)}x_0$, onde $x'_0 \in \mathcal{N}(A)^\perp = \text{Im}(A^*)$ e $P_{\mathcal{N}(A)}x_0 \neq 0$. Como $A^*AP_{\mathcal{N}(A)}x_0 = 0$, de (9.2.17), temos:

$$x_k = x'_0 - \sum_{j=0}^{k-1} \alpha_j r_j + P_{\mathcal{N}(A)}x_0 \longrightarrow x^\dagger + P_{\mathcal{N}(A)}x_0.$$

Isto completa a demonstração. ■

9.3 Métodos tipo Kaczmarz

Dedicaremos esta seção a um método de regularização para resolver problemas inversos que aparecem na forma de sistemas de equações

$$A_i(x) = y_i \quad i \in \{1, \dots, M\}. \quad (9.3.19)$$

Uma maneira de resolver o problema (9.3.19) é considerarmos o operador $\mathbb{A} := (A_1, \dots, A_M)$ e $Y = (y_1, \dots, y_M)$ e resolver a equação

$$\mathbb{A}(x) = Y,$$

usando as técnicas que foram apresentadas acima.

A principal ideia de métodos tipo Kaczmarz é resolver o sistema (9.3.19), de maneira cíclica, onde cada equação do sistema é considerado em separado. Nas próximas subseções apresentaremos o método ART (Algebraic Reconstruction Techniques) e o método de Landweber-Kaczmarz como exemplos de tais estratégias.

9.3.1 Método ART

Uma das técnicas mais usadas em diagnósticos médicos para Tomografia Computadorizada, até pouco tempo, é o método ART. Este

método representa a ideia de métodos do tipo Kaczmarz de uma maneira simples de entender. Faremos a construção do método iterativo ART para o caso em que $A_i = (a_{i1}, \dots, a_{iN})$, onde $i \in \{1, \dots, M\}$, representa uma linha de um sistema linear escrito na forma (9.3.19). Desta maneira, a iteração ART usa conceitos básicos de Álgebra Linear.

Iteração: Dado \vec{x}_0 um ponto inicial, projetamos \vec{x}_0 ortogonalmente sobre o hiperplano determinado pela primeira equação do sistema (9.3.19), isto é, sobre $a_{11}x_1 + \dots + a_{1N}x_N = y_1$. Ao vetor projeção ortogonal, chamamos de primeiro iterado \vec{x}_1 . De posse do vetor \vec{x}_1 projetamos ortogonalmente sobre a segunda linha do sistema (9.3.19) e obtemos \vec{x}_2 . Aplicamos este procedimento, sucessivamente, até M , obtemos \vec{x}_M . De maneira cíclica, projetamos $\vec{x}_{mod(M)}$ sobre a equação $mod(M)$ do sistema (9.3.19). A figura mostra a iteração ART para um sistema quadrado de ordem 2.

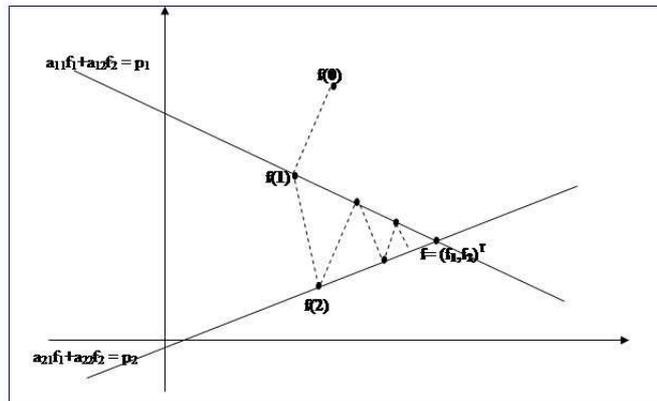


Figura 9.1: Geometria do algoritmo ART.

Assim, a iteração é dada por

$$\vec{x}_n = \vec{x}_{n-1} - \left(\frac{\langle A_i, \vec{x}_{n-1} \rangle - y_i}{\|A_i\|^2} \right) (A_i)^T \quad (9.3.20)$$

Observamos que um dos defeitos do método ART é seu alto custo

computacional. Ainda, a taxa de convergência pode ser muito lenta ou até mesmo divergir. Assim, às vezes é necessária a introdução de parâmetros de correção na equação (9.3.20) de forma a melhorar a performance do método².

9.3.2 Método de Landweber-Kaczmarz

O método de Landweber-Kaczmarz consiste em aplicar o método de Landweber para resolver (9.3.19) de forma cíclica. Isto é, a iteração é definida como

$$x_{k+1}^\delta = x_k^\delta + \omega_k A'_{[k]}(x_k^\delta)(y^\delta - A_{[k]}(x_k^\delta)), \quad (9.3.21)$$

onde $[k] := k \bmod(M) \in \{0, \dots, M-1\}$ e $i = [k] + 1$. O parâmetro ω_k é definido como

$$\omega_k := \begin{cases} 1 & \text{se } \|A_{[k]}(x_k^\delta) - y_{[k]}^\delta\| > \tau\delta, \\ 0 & \text{caso contrário.} \end{cases}$$

O parâmetro ω_k determina o critério de parada para o método de Landweber-Kaczmarz. A iteração é parada no primeiro índice $k_* = k_*(\delta, y_{[k]}^\delta)$ tal que

$$\omega_{k_*+j} = 0 \quad \text{para } j = \{0, \dots, M-1\}. \quad (9.3.22)$$

Notamos que, um tal critério de parada dado pela equação (9.3.22) significa dizer que

$$x_{k_*}^\delta = x_{k_*+1}^\delta = \dots = x_{k_*+M-1}^\delta, \quad (9.3.23)$$

isto é, k_* é escolhido como o índice que faz com que x_k^δ seja igual em um ciclo.

Convergência para a iteração de Landweber-Kaczmarz segue similar ao feito para o método de Landweber. As hipóteses são similares as feitas para o método de Landweber para cada $A_{[k]}$.

²Hoje em dia, novos métodos ganharam terreno, [31, 30]. Atualmente, o método Filtered Backprojection é o mais usado em Tomografia Computadorizada. Este consiste em uma estratégia para resolver a Transformada de Radon Inversa.

9.4 Aplicação: Tomografia Computadorizada

Como aplicação, apresentaremos os resultados obtidos nas Seções 9.2 e 9.3 para o problema de TC com uma quantidade limitada de dados, [30, 31].

Seja $L^2(D)$ o espaço de Hilbert de todas as funções quadrado integráveis no disco unitário $D \subset \mathbb{R}^2$ e $\mathcal{H}_2 := \{y : [0, 2] \rightarrow \mathbb{R} : \|y\|^2 := \int_0^2 y(t)t dt < \infty\}$.

Consideraremos o sistema

$$A_j x = y_j, \quad j = 0, \dots, N-1, \quad (9.4.24)$$

onde cada $A_j := L^2(D) \rightarrow \mathcal{H}_2$, é dada por

$$(A_j x) := \frac{1}{\pi} \int_{S^1} x(\xi_j + t\theta) d\Omega(\theta), \quad t \in [0, 2], \quad (9.4.25)$$

A equação (9.4.25) corresponde a uma versão circular da Transformada de Radon que foi introduzida na Subseção 1.5.1. Resolver (9.4.25) é um passo crucial em TC. Consideraremos o caso em que os dados são obtidos por integração numérica. Ainda, supomos que o centro de integração ξ_j corresponde a posição de cada detector.

Nesta aplicação, consideraremos somente o caso especial em que cada

$$\xi_j = \left(\sin \left(\frac{\pi j}{N-1} \right), \cos \left(\frac{\pi j}{N-1} \right) \right)$$

está uniformemente distribuído no semi-círculo $S_+^1 := \{\xi = (\xi^1, \xi^2) \in \partial D : \xi^1 \geq 0\}$. Assim, os dados são medidos em uma única parte da fronteira de D (veja a figura 9.2). Portanto, temos poucos dados. É provado, por exemplo em [31, 30], que certos detalhes de x fora da região de detecção não podem ser reconstruídos. Tais resultados são conhecidos como invisibilidade.

Temos duas possibilidades. A primeira é considerar

$$\begin{aligned} \mathbb{A} = (A_0, \dots, A_{N-1})^T : L^2(D)^N &\longrightarrow \mathcal{H}_2^N \\ x &\longmapsto \mathbb{A}x = (y_0, \dots, y_{N-1})^T = y. \end{aligned}$$

Com essa configuração podemos aplicar a iteração de Landweber linear (9.1.6), a iteração de Landweber-Kaczmarz ou a iteração de descida máxima (9.2.16).

A segunda possibilidade é aplicar o método de Landweber-Kaczmarz para o sistema de equações lineares (9.4.25). Para nosso teste numérico, optamos pela segunda.

A solução x^\dagger é mostrada no lado esquerdo da figura 9.2. Do lado direito da figura 9.2, vemos a figura que consiste da superposição de funções características e um kernel Gaussiano que representa os dados do problema, com $N = 50$ medidas. Os dados $y_j = A_j x^\dagger$ são calculados por integração numérica de forma a adicionar 4% de ruído.

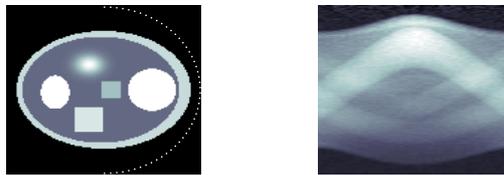


Figura 9.2: Do lado direito a solução e do esquerdo os dados com 4% de ruídos.

A figura 9.3 mostra a solução regularizada $x_{k_*}^\delta$, os métodos Landweber-Kaczmarz e steepest descent com uma estratégia tipo Kaczmarz.

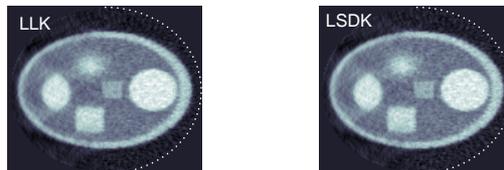


Figura 9.3: Do lado direito a solução aproximada pela regularização de Landweber e do esquerdo por steepest-descent, com estratégias tipo Kaczmarz.

9.5 Exercícios

Exercício 9.4. *Mostre que se $y \in \mathcal{D}(A^\dagger)$, então $A^*y = A^*Ax^\dagger$. Sugestão: use a inversa generalizada. Consulte o Apêndice 6.*

Exercício 9.5. *Faça um gráfico comparando os resultados obtidos nos Lemas 9.1 e 9.2 com a estimativa (9.1.11). Compare com a figura 1.1.*

Exercício 9.6. *Prove que se A é um operador linear e limitado, então $J(x) = \frac{1}{2}\|Ax - y\|^2$ é Fréchet diferenciável e $\nabla J(x) = A^*(Ax - y)$.*

Exercício 9.7. *Mostre o Lema 9.3 sem a hipótese de A possuir imagem fechada.*

Exercício 9.8. *Suponha que existam constantes $C, M, m > 0$ tais que $m\|x\|^2 \leq \langle A^*Ax, x \rangle \leq M\|x\|^2$, $\forall x \in \text{Im}(A^*)$ e que A satisfaça a Hipótese 6.1. Mostre que o método steepest descent possui um limitante de erro dado por*

$$\|x^\dagger + P_{N(A)}x_0 - x_n\| \leq C \left(\frac{M - m}{M + m} \right)^2.$$

Exercício 9.9. *Considere a reta de equação $a_{i1}x_1 + \dots + a_{iN}x_N = y_i$ em \mathbb{R}^N . Mostre que a projeção ortogonal de $v \in \mathbb{R}^N$ sobre a reta acima satisfaz a equação (9.3.20).*

Exercício 9.10. *Mostre que $\langle A_i, \vec{x}_n \rangle = y_i$. Interprete esse resultado geometricamente.*

Exercício 9.11. *Use a iteração (9.3.20) para obter uma solução do problema de tomografia descrita pela figura 1.7.*

Exercício 9.12. *Mostre que o método de Landweber-Kaczmarz (9.3.21), com o critério de parada (9.3.22), é um método de regularização. Sugestão: consulte [21].*

Exercício 9.13. *Use a iteração de Landweber-Kaczmarz (9.3.21) para encontrar uma solução do problema de tomografia proposto na Figura 1.7. Compare com o método ART.*

Exercício 9.14. Prove que cada operador A_j é linear, limitado e satisfaz $\|A_j\| \leq 1$. Mostre ainda que o operador adjunto é dado por $(A_j^*y)(\xi) = y \frac{(|\xi_j - \xi|)}{\sqrt{\pi}}$.

Bibliografia

- [1] M. J. Ablowitz e A. S. Fokas, "Complex Variables", Cambridge University Press, 2000.
- [2] G. Bal, "Introduction to Inverse Problems", Department of Applied Physics and Applied Mathematics, Columbia University, New York, 2004.
- [3] J. Baumeister, "Stable solution of inverse problems", Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1987.
- [4] J. Baumeister e A. Leitão, "Topics in inverse problems", Publicações Matemáticas do IMPA, 25º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 2005.
- [5] M. Bertero e P. Boccacci, "Introduction to inverse problems in imaging", Institute of Physics Publishing, Bristol, 1998.
- [6] T. M. Buzug, "Computer Tomography - From Photom Statistics to Modern Cone-Beam CT", Springer, Berlin Heidelberg, 2008.
- [7] A. De Cezaro e A. Leitão, "Introdução aos Problemas Inversos Lineares", IV Bienal da Sociedade Brasileira de Matemática, Maringá PR, online: www.dma.uem.br/bienalsbm, 2008.
- [8] A. De Cezaro e A. Leitão, "Problemas Inversos: Uma Introdução", I Colóquio de Matemática da Região Sul, Santa Maria, RS, online: w3.ufsm.br/colmatsul/, 2010.

- [9] A. De Cezaro e F. Travessini De Cezaro, "Problemas Inversos e a Matemática da Tomografia Computadorizada", V Bienal da Sociedade Brasileira de Matemática, João Pessoa, PB, online: www.mat.ufpb.br/bienalsbm, 2010.
- [10] H. W. Engl, M. Hanke e A. Neubauer, "Regularization of inverse problems", Mathematics and its Applications, Vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
- [11] C. L. Epstein, "Introduction to the Mathematics of Medical Imaging", Pearson Prentice Hall, Upper Saddle River, NJ, 2003.
- [12] L. C. Evans, "Partial differential equations", Graduate Studies in Mathematics, Vol 19, American Mathematical Society, Providence, 1998.
- [13] G. H. Golub e C. F. Van Loan, "Matrix computations", Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, 1996.
- [14] C. W. Groetsch, "Generalized inverses of linear operators: representation and approximation", Monographs and Textbooks in Pure and Applied Mathematics, No. 37, Marcel Dekker Inc., New York, 1977.
- [15] C. W. Groetsch, "The theory of Tikhonov regularization for Fredholm equations of the first kind", Research Notes in Mathematics, Vol. 105, Pitman (Advanced Publishing Program), Boston, 1984.
- [16] C. W. Groetsch, "Inverse problems in the mathematical sciences", Vieweg Mathematics for Scientists and Engineers, Friedr. Vieweg & Sohn, Braunschweig, 1993.
- [17] F. A. Grünbaum, "Some mathematical problems suggested by limited angle tomography", Inverse problems (New York, 1983), SIAM-AMS Proc., Vol. 14., AMS, Providence, pp . 65 - 77, 1984.
- [18] J. Hadamard, "Lectures on Cauchy's problem in linear partial differential equations", Dover Publications, New York, 1953.

- [19] P. C. Hansen, "Rank-deficient and discrete ill-posed problems", SIAM Monographs on Mathematical Modeling and Computation, SIAM, Philadelphia, 1998.
- [20] V. Isakov, "Inverse problems for partial differential equations", Applied Mathematical Sciences, Vol. 127, Springer, New York, 2006.
- [21] B. Kaltenbacher, A. Neubauer e O. Scherzer, "Iterative regularization methods for nonlinear ill-posed problems", Radon Series on Computational and Applied Mathematics, Vol. 6, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
- [22] L. B. Keller, "Inverse problems", Amer. Math. Monthly, Vol. 83, **2**, pp. 107-118, 1976.
- [23] A. Kirsch, "An introduction to the mathematical theory of inverse problems", Applied Mathematical Sciences, Vol. 120, Springer-Verlag, New York, 1996.
- [24] E. Kreyszig, "Introductory functional analysis with applications", Wiley Classics Library, John Wiley & Sons Inc., New York, 1989.
- [25] L. Landweber, "An iteration formula for Fredholm integral equations of the first kind", Amer. J. Math., Vol. 73, pp. 615 - 624, 1951.
- [26] A. K. Louis, "Inverse problems in medicine", Applications of mathematics in industry and technology (Siena, 1988), Teurner, pp. 227 - 287, 1989.
- [27] A. K. Louis, "Application of the approximate inverse to 3-D X-ray CT and ultrasound tomography", Inverse problems in medical imaging and nondestructive testing (Oberwolfach, 1996), Springer, pp. 120 - 133, 1997.
- [28] A. K. Louis e E. T. Quinto, "Local tomographic methods in sonar", Surveys on solution methods for inverse problems, Springer, pp. 147 - 154, 2000.

- [29] C. Meyer, "Matrix analysis and applied linear algebra", SIAM, Philadelphia, 2000.
- [30] F. Natterer, "The mathematics of computerized tomography", Classics in Applied Mathematics, Vol. 32, SIAM, Philadelphia, 2001.
- [31] F. Natterer e F. Wübbeling, "Mathematical methods in image reconstruction", SIAM Monographs on Mathematical Modeling and Computation, SIAM, Philadelphia, 2001.
- [32] Platão, "A República", Coleção A Obra - Prima de cada Autor, Martin Claret, São Paulo, 2002.
- [33] J. Radon, "On the determination of functions by their integral values along certain manifolds", Conf. Proc. Lecture Notes Math. Phys., IV, 75 years of Radon transform (Vienna, 1992), Int. Press, Cambridge, pp. 324-339, 1994.
- [34] A. G. Ramm e A. I. Katsevich, "The Radon transform and local tomography", CRC Press, Boca Raton, 1996.
- [35] A. Sá Barreto, "Introdução as transformadas de Radon", Lectures Notes: Symposium on Spectral and Scattering Theory, Recife, online: www.math.purdue.edu/~sabarre/papers.html, 2003.
- [36] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier e F. Lenzen, "Variational methods in imaging", Applied Mathematical Sciences, Vol. 167, Springer, New York, 2009.
- [37] G. Strang, "Linear algebra and its applications", Academic Press, New York, 1980.
- [38] C.R. Vogel, "Computational methods for inverse problems", Frontiers in Applied Mathematics, Vol 23, SIAM, Philadelphia, 2002.
- [39] J.P. Zubelli, "An introduction to inverse problems. Examples, methods and questions", 22nd Brazilian Mathematics Colloquium, IMPA, Rio de Janeiro, 1999.

Índice

- Abel
 - equação integral de , 43
- aproximação
 - de Born, 58
- coeficientes
 - de atenuação, 25
- condição
 - de fonte, 115
 - de radiação, 33
- critério de parada, 113, 122
- curva L, 102
- diferenças simétricas, 16
- distribuição
 - de Dirac, 47
- equação
 - de Helmholtz, 32
 - do transporte, 28
 - normal, 99
- Eratóstenes, 11
- erro
 - de aproximação, 113
 - nos dados, 113
- espaço
 - $L^p(\mathbb{R}^n)$, 17
 - de Hilbert, 17
- espectro, 112
- estimativa
 - de erro, 16
 - fundamental, 113
- fórmula
 - de Cauchy, 46
 - de Green, 48
 - de Plemelj, 51
- fator integrante, 29, 35
- função
 - de Hankel, 59
 - filtro, 100
 - radial, 42
- funcional
 - coercivo, 107
 - de Tikhonov, 101
- Gauss, 109
- Hadamard, 14
- identidade
 - de Parseval, 18
- iteração
 - ART, 121
 - de Landweber, 111
- lei de Beer, 21

- método
 - ART, 120
 - das características, 35
 - de Landweber, 111
 - de Landweber-Kaczmarz, 122
 - de mínimos quadrados, 109
 - de separação de variáveis, 21
 - de Tikhonov, 99
 - iterativo de regularização, 109
 - tipo Kaczmarz, 120
- medidas
 - de reflexão, 61
 - de transmissão, 60
- Operador, 14
- parâmetro
 - de regularização, 102
- Platão, 11
- princípio
 - da discrepância, 114
 - de superposição, 58
- problema
 - de espalhamento, 34
 - de espalhamento inversos, 34
 - de Riemann Hilbert, 29, 45
 - direto, 13
 - inverso, 13
 - inverso de identificação, 13
 - inverso de reconstrução, 13
- pseudo-inversa, 111
- Radon, 24
- ruídos, 14
- série de Neumann, 65
- sistemas de equações, 120
- solução
 - de quadrados mínimos, 111
 - fundamental, 47
- teorema
 - da Aplicação Espectral, 100, 112
 - da Projeção, 73
 - das Fatias de Fourier, 38
 - de Fubini, 38
 - do Gráfico Fechado, 107
- termo de penalização, 102
- Tomografia, 19
 - computadorizada, 19
 - PET, 27
 - por difração, 31
 - por emissão, 27
 - por fluorescência óptica, 31
 - SPECT, 27
- Transformada
 - de Fourier, 17
 - de Fourier Inversa, 18
 - de Hilbert, 41
 - de Radon, 24
 - de Radon Atenuada, 29