

Problemas Inversos: Uma Introdução

Adriano De Cezaro¹

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

¹adrianocezaro@furg.br

Conteúdo

Introdução	3
Notação	7
Lista de Figuras	8
1 Introdução por Meio de Exemplos	9
1.1 O Problema Inverso da Diferenciação	14
1.1.1 Reconstrução de uma Força Desconhecida	15
1.1.2 Diferenciação Numérica	17
1.2 A Tomografia por Raio-X: Transformada de Radon	19
1.2.1 Tomografia Computadorizada: Caso Contínuo	20
1.2.2 Tomografia Computadorizada: Caso Discreto	31
1.3 Tomografia Elétrica por Impedância - EIT	34
1.4 Breve Apanhado Histórico do Capítulo	40
2 Métodos de Regularização Iterativos Tipo Gradiente	44
2.1 Método de Landweber Linear	47
2.1.1 Convergência	48
2.2 Método de Landweber Não-Linear	54
2.2.1 Convergência	57

2.3	Método de Landweber Iterativamente Regularizado . . .	62
2.3.1	Taxas de Convergência	67
2.4	Método do Gradiente	68
2.5	Método de Landweber e Estratégias de Kaczmarz . . .	76
2.5.1	Um Pouco Sobre o Método ART	76
2.5.2	Método de Landweber-Kaczmarz	78
2.6	Aplicação: Tomografia Computadorizada	80
2.7	Breve Apanhado Histórico do Capítulo	82
3	Métodos de Regularização Iterativos Tipo Newton	85
3.1	Método de Gauss-Newton	87
3.2	Método de Levenberg-Marquardt	88
3.2.1	Convergência	90
3.3	Método de Gauss-Newton Iterativamente Regularizado	92
3.4	Breve Apanhado Histórico do Capítulo	92
4	Método de Regularização de Tikhonov	95
4.1	Problemas Lineares: Convêrgencia	95
4.2	Problemas Lineares: Semi-Convergência	101
4.2.1	Caso Linear: Taxas de Convergência	103
4.3	Problemas Não-Lineares: Convergência	103
4.4	Problemas Não-Lineares: Semi-Convergência	105
4.4.1	Taxas de Convergência	108
4.5	Breve Apanhado Histórico do Capítulo	108
A	Definições e Resultados	110
A.1	Definições e Resultados Básicos em Espaços Vetoriais.	110
A.1.1	Operadores Lineares	112
A.1.2	Transformada de Fourier	114

B	Sistemas de Equações Lineares	117
B.1	Pseudo - Inversa de Operadores Lineares	118
B.1.1	Definições e Propriedades Básicas	118
B.2	A Decomposição em Valores Singulares	127
B.2.1	Funções de Operadores: Aplicação Espectral	132
B.2.2	Relação entre Má-Colocação e Valores Espectrais	135
C	O Conceito de Regularização	139
C.1	Resultados de Convergência	143
C.2	Escolha do Parâmetro de Regularização	145
C.2.1	Escolha <i>a priori</i>	146
C.2.2	Escolha <i>a posteriori</i>	150
C.3	Regularização por Truncamento dos Valores Singulares	151

Introdução

Nas últimas quatro décadas, uma área da matemática aplicada provocou o interesse de muitos pesquisadores e vem despertando a atenção de setores da indústria pelo assunto estudado e suas aplicações. A essa área, denominamos *Problemas Inversos*.

A bem da verdade, *Problemas Inversos* pertencem a uma área de estudos multi-disciplinar, envolve profissionais de várias áreas, como físicos, químicos, engenheiros, profissionais da saúde e, é claro, **matemáticos**. Dentro da matemática, *Problemas Inversos* envolvem conhecimentos de várias áreas, como cálculo (análise), geometria, equações diferenciais ordinárias e parciais, análise numérica, álgebra linear, etc. Ainda, para fins práticos das aplicações temos que contar com boas implementações computacionais. A interdisciplinariedade do termo *Problemas Inversos* ficará mais clara a medida que o leitor adentrar por essas notas, que estão assim organizadas:

No Capítulo 1, trataremos de algumas aplicações dos *Problemas Inversos*, que tem o intuito de despertar o interesse pelo assunto e servir de motivação para os capítulos posteriores. Acreditamos que essa seja uma maneira eficaz para o posterior desenvolvimento de uma teoria bem fundamentada para a solução dos problemas apresentados, a qual contém Lemas, Proposições e Teoremas. Assim, o Capítulo 1 trata de alguns exemplos dos *Problemas Inversos* e suas

aplicações. Daremos enfoque a alguns problemas que possuam, em primeiro lugar, um caráter didático, e que possam ser seguidos com conhecimentos de cálculo e álgebra linear, mas, por outro lado, que sejam também assunto de interesse atual na pesquisa em *Problemas Inversos*. Nosso objetivo é mostrar que a área de *Problemas Inversos* trata de problemas que tem importância no dia-a-dia das pessoas, que é uma área fértil de pesquisa e que esta possui influências socio/econômicas relevantes. Em particular, nessas notas apresentaremos problemas relacionados à tomografia computadorizada e à tomografia elétrica por impedância (EIT) como exemplos de problemas inversos com aplicações relevantes.

Boa parte dessas notas se dedica ao estudo de métodos de regularização iterativos para problemas inversos. Em especial, o Capítulo 2 contempla métodos iterativos do tipo gradiente, que são bem conhecidos dos cursos de Cálculo (cálculo numérico) e Álgebra Linear. As Seções desse Capítulo se diferenciam por apresentarem estratégias de escolha do passo de descida para métodos do tipo gradiente. Ao final deste, apresentaremos estratégias para a solução de sistemas de equações, os chamados métodos do tipo Kaczmarz. Em especial, motivamos tal estratégia usando o algoritmo conhecido como ART, amplamente utilizado em imagens médicas. O Capítulo 3 é dedicado a métodos iterativos tipo Newton, o qual também conhecemos de cursos elementares. Mais uma vez, as Seções contemplam diferentes estratégias de escolha do passo de decida e uma motivação para tal escolha. Somente no Capítulo 4, nos utilizaremos de técnicas variacionais (minimização de funções) para desenvolver a teoria do métodos de regularização de Tikhonov. Tais estratégias visam imitar o Teorema de Weierstrass, bem conhecido dos cursos de Cálculo. Ainda, a Teoria de regularização de Tikhonov possui uma interpretação geométrica bastante sugestiva, a qual, ficará clara du-

rante as notas. Todos esses capítulos trazem uma análise completa da teoria de regularização para os respectivos métodos ali tratados. Grande parte dos Teoremas, Proposições e Lemas são apresentados com demonstrações completas. Alguns detalhes são deixados ao leitor, como forma de exercícios. Por este manuscrito se tratar de um curso de introdução aos problemas inversos, algumas demonstrações que fogem ao escopo destas notas. Tais demonstrações são deixadas ao leitor como forma de pesquisa. Vários exercícios são apresentados ao longo do texto, como forma de motivação para o leitor. No final de cada capítulo, procuramos ainda deixar um breve apanhado histórico do mesmo. Tal apanhado histórico não é completo, mas não deixa de enriquecer o conhecimento do leitor.

No Apêndice A, apresentamos alguns pré-requisitos básicos de Análise e Álgebra Linear que são utilizados durante essas notas. O Apêndice B traz a teoria de regularização usando o conceito de pseudo-inversa de um operador. Consideramos tal apêndice muito instrutivo, uma vez que trata de conceitos bem conhecidos da Álgebra Linear e Análise. Ainda, é possível tratar tais conteúdos de uma forma mais geométrica, o que facilita a compreensão do assunto.

Tendo feito uma motivação para a teoria por meio de exemplos, passaremos a tratar de como resolver tais problemas e estudar formas de transpor as dificuldades inerentes aos problemas inversos. Para isso, o Apêndice C trata do conceito geral de regularização para problemas inversos, que é a base para o desenvolvimento dos métodos de regularização nos Capítulos 2, 3 e 4.

Propomos ao leitor duas maneiras diferentes de seguir essas notas. Um leitor familiarizado com a noção de regularização para problemas inversos e o aparato matemático que permeia tal teoria pode (e deve) se concentrar nos Capítulos 1, 2, 3 e 4 e nos resultados teóricos e práticos desses capítulos. Como forma de mo-

tivação para o estudo, uma leitura preliminar do Capítulo 1 é recomendada. Para o leitor que não está familiarizado com tais noções sugerimos que comece a leitura destas notas pelo Apendice B e siga ao Apêndice C. Uma vez adquirida uma cultura básica em problemas inversos contidos nesses dois apêndices, sugerimos que passe a entender melhor os exemplos e aplicações contidas no Capítulo 1 e depois, siga ao capítulos seguintes, os quais se propõem a desenvolver a teoria necessária para atacar os problemas abordados no Capítulo 1.

Essas notas possuem uma intersecção não vazia com as notas apresentadas na *IV Bienal da Sociedade Brasileira de Matemática* realizada em Maringá - PR durante o período de 29/09 a 03/10 de 2008 [7]. No entanto, as notas atuais tratam de problemas e de teorias não abordadas anteriormente. Em suma, o tratamento dado aqui é mais geral que o apresentado na IV Bienal.

Ao idealizar estas notas, procuramos nos manter fiéis a teoria de regularização para *Problemas Inversos*, mas, acima de tudo, simplificar ao máximo os conceitos envolvidos, de forma a atender os pré-requisitos de um público heterogêneo, que contemple Alunos de Graduação, Pós-Graduação, bem como, Professores e demais interessados pelo assunto.

Gostaria de deixar registrado meus sinceros agradecimentos ao Prof. Antonio Leitão (UFSC) pelas valiosas sugestões, que enriqueceram o assunto abordado nessas notas.

Notação

A	operador linear
$A(\cdot)$	operador não necessariamente linear
A^*	operador adjunto de A
$A'(\cdot)$	derivada de Fréchet de $A(\cdot)$
A^\dagger	pseudo-inversa de A
δ	nível de ruídos
y^δ	dados y com ruído de nível δ
x^\dagger	solução de norma mínima
x^δ	solução aproximada para dados com ruído y^δ
x_k^δ	k -ésimo iterado
x_α^δ	solução regularizada com parâmetro de regularização α .
$\mathcal{R}f$	Transformada de Radon da função f
$\hat{}$	Transformada de Fourier \mathcal{F}
$\check{}$	Transformada de Fourier inversa \mathcal{F}^{-1}
$\mathcal{H}f$	Transformada de Hilbert da função f
$\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$	espaços de Hilbert
$\langle \cdot, \cdot \rangle$	produto interno
$\ \cdot \ $	norma
$\ \cdot \ _\infty$	norma do supremo
\rightarrow	convergência forte
\rightharpoonup	convergência fraca
$C[a, b]$	espaço das funções contínuas em $[a, b]$
$L^2[a, b]$	espaço das funções quadrado integráveis em $[a, b]$
$L^1[a, b]$	espaço das funções módulo integráveis em $[a, b]$
\mathbb{S}^{n-1}	esfera em \mathbb{R}^n
$H(s, w)$	Hiperplano a distância $ s $ da origem
$\mathcal{S}(\mathbb{R}^n)$	espaço de Schwartz
$\frac{\partial^k}{\partial s^k}$	derivada parcial de ordem k com relação a variável s
$rot \vec{E}$	rotacional do campo vetorial \vec{E}

Lista de Figuras

Figura	Página
Modelo.	11
Estimativa de erro.	19
Raio-X.	21
Aparelho de raio-X.	21
Tomografia discreta.	31
Feixe de raio-X.	31
Eletrodos.	38
Dados obtidos através de EIT.	38
Geometria do algoritmo ART.	78
Imagem de um Phantom.	81
Dados com 4% de ruídos.	81
Solução aproximada por Landweber e steepest-descent.	82
Interpretação geométrica da inversa de uma matriz.	125
Estimativa fundamental.	148

Capítulo 1

Problemas Inversos: Introdução por Meio de Exemplos

Problemas que envolvem a determinação de uma **causa** (desconhecida) a partir de um **efeito** (dado) medido ou observado, possuem uma vasta quantidade de aplicações em várias áreas da ciência. Por exemplo, em imagens médicas podemos citar: tomografias [4, 37, 38], eletrocardiologia e ultrassom [34, 35, 42]; em geofísica e ciências ambientais: explorações sísmicas [27, 42], detecção de depósito de petróleo, sedimentos e outras riquezas, monitoramento de poluentes no subsolo [8, 9, 52], em engenharia: testes não-destrutivos em componentes (semi-condutores em nanotecnologia) [1, 3, 8, 23, 49]. Nessa lista podem ser incorporadas ainda aplicações a física [30, 40], química [8, 15, 49], biologia [24, 52], finanças quantitativas [23], etc. Além da relevância das aplicações, a formulação e solução de tais proble-

mas envolvem o conhecimento de vários campos da matemática, de ciências aplicadas e o envolvimento de profissionais dessas áreas.

A esse novo campo de estudos na área da matemática aplicada denominamos *Problemas Inversos*. Dada a interdisciplinariedade e relevância das aplicações, problemas inversos tem atraído uma quantidade grande de pesquisadores com interesse por tais problemas.

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

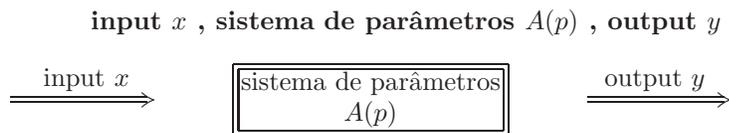
- **fatores sociais:** técnicas de detecção de tumores implicam em prolongar a vida das pessoas.
- **fatores econômicos:** detecção de tumores implica em tratamentos mais eficazes contra o câncer, diminuindo os custos dos mesmos. Ainda, prolonga a vida ativa das pessoas que, consequentemente, geram mais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

Desde o início, o leitor deve estar se perguntando: o que são os tais "**Problemas Inversos**"? Ou ainda, a pergunta pode ser: "**inversos**" do quê? Para J.B. Keller [26], dois problemas são o inverso um do outro, se a formulação de um envolve o conhecimento (mesmo que parcial) do outro, este último, conhecido como o "**Problema Direto**". Assim, a grosso modo, *problemas inversos* estão relacionados com a determinação de **causas**, através da observação (ou medida) de **efeitos**.

Do ponto de vista de 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. Exemplos atuais são: estudos de mudanças climáticas drásticas à milhões de anos, através de medidas observáveis hoje nas camadas glaciais das calotas polares. A segunda é prever os fenômenos futuros, influenciados pelos estados atuais ou por parâmetros de um sistema físico. Ambas motivações são modeladas por equações matemáticas [22, 23, 49].

Sintetizando as ideias acima, podemos assumir que o fenômeno físico, biológico, 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**) corresponde ao output.
2. *O problema de identificação:* Dados input (**causa**) e output (**efeito**), determinar o sistema de parâmetros que relaciona o input ao output.

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

$$A(p)x = y, \tag{1.1}$$

para um dado efeito y ; uma causa a ser determinada x (que pertence a um conjunto de parâmetros admissíveis $p \in \mathcal{U}$) e $A(p)$ representa o modelo que associa a causa ao determinado efeito.

Em termos práticos, os dados y são obtidos por medições e assim, dificilmente, são obtidos de forma exata, dada a natureza da obtenção desses dados (medidas). Portanto, costumamos denotar as medições obtidas por y^δ , das quais assumimos conhecer o nível de **ruídos** δ , satisfazendo

$$\|y - y^\delta\| \leq \delta. \quad (1.2)$$

Numa 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 , com respectivos produtos internos. Veja o Apêndice A para uma definição precisa de espaços de Hilbert. Em uma boa quantidade de aplicações, o operador $A(p)$ é linear e a teoria já é amplamente desenvolvida [2, 8, 14, 27]. Para o caso linear, denotaremos o operador por A . Uma grande quantidade de problemas de interesse envolvem um operador $A(p)$ que é não-linear [3, 8, 17, 25]. No caso do operador $A(p)$ ser não-linear, costumamos denotar $A(p) := A(\cdot)$. Para tais casos, a teoria é bem mais complicada, mas, mesmo assim, nas últimas duas décadas avançou muito. Trataremos da teoria para ambas as classes de problemas inversos nos Capítulos 2, 3 e 4.

Na formulação matemática dos problemas podemos caracterizar:

O problema direto: Dado $x \in \mathcal{H}_1$ e $p \in \mathcal{U}$, encontrar

$$y := A(p)x. \quad (1.3)$$

O problema inverso: Esse aparece, pelo menos de duas for-

mas:

1. *O problema de reconstrução*: Observado $y \in \mathcal{H}_2$ e conhecido o sistema de parâmetros $A(p)$ para $p \in \mathcal{U}$, encontrar

$$x \in \mathcal{H}_1 \quad \text{tal que} \quad A(p)x = y. \quad (1.4)$$

2. *O problema de identificação*: Dados $x \in \mathcal{H}_1$ e $y \in \mathcal{H}_2$, encontrar

$$p \in \mathcal{U} \quad \text{tal que} \quad A(p)x = y. \quad (1.5)$$

Uma característica que diferencia um **problema direto** de um **problema inverso** é que o segundo, em geral, é *mal-posto* no sentido de Hadamard [19]. 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 requisitos acima não seja satisfeito, o problema é dito mal-posto.

Como veremos abaixo, os efeitos da má-colocação ou mal-condicionamento de um problema estão atrelados aos problemas inversos em qualquer dimensão, lineares ou não.

Problemas inversos em dimensão finita aparecem naturalmente na solução de sistemas de equações [20]. Do ponto de vista computacional [20, 51], sempre estamos tratando de problemas inversos em dimensão finita, assim, faz-se jus estudá-los com o devido interesse. Neste caso, os problemas de mal-condicionamento estão intimamente ligados aos autovalores da matriz que representa o sistema [12, 36, 48]. Se os autovalores são muito próximos de zero, ou os dados não pertencem ao espaço solução do problema, podemos enfrentar sérias complicações numéricas [12, 36, 48]. Veja também o Apendice B.

Em termos práticos, esse sempre é o caso, pois, não somos ca-

pazes de representar estruturas não discretas em um programa de simulação computacional, muito embora, a verdadeira solução do problema, em geral, pertença a espaços de dimensão infinita.

Exemplos importantes de problemas inversos lineares são dados por equações integrais de primeira espécie ou por algumas transformadas integrais [4, 14, 20, 37, 42, 47]. Apresentaremos o problema da tomografia por raio-X na Seção 1.2 como um problema inverso formulado por uma equação integral.

Problemas inversos de identificação de parâmetros em equações diferenciais parciais são, em geral, não-lineares, mesmo que o problema direto seja linear [8, 25, 23]. Nestas notas, apresentaremos o problema da *Tomografia Elétrica por Impedância - (EIT)* como um exemplo de problema inverso não-lineares. Veja a Subseção 1.3.

Nas próximas seções, discutiremos alguns exemplos de problemas inversos. Esses exemplos tem o intuito de revelar as dificuldades teóricas e práticas na tentativa de solucionar tais problemas e, assim, motivar o estudo dos métodos de solução regularizadas nos próximos capítulos. Esperamos que o leitor possa acompanhar os raciocínios com conhecimentos básicos de Análise (Espaços Métricos) [31, 32, 33] e Álgebra Linear [12, 36, 48] e generalizações simples desses conceitos para espaços vetoriais de dimensão infinita.

1.1 O Problema Inverso da Diferenciação

Dada a natureza suavizante do operador de integração, costumamos considerar a diferenciação como sendo o problema inverso da integração. Esta afirmativa pode ser comprovada nos exercícios abaixo.

Exercício 1.1. *Seja $f : [a, b] \rightarrow \mathbb{R}$ uma função contínua. Mostre*

que a função

$$g(x) = \int_a^x f(t)dt$$

é uma função diferenciável para todo $x \in]a, b[$.

Exercício 1.2. *Dê um exemplo de uma função $f : [a, b] \rightarrow \mathbb{R}$ contínua que não tenha derivada em algum ponto de $[a, b]$. Construa uma função contínua que não possua derivadas em nenhum ponto de $[a, b]$.*

Nesta seção, apresentaremos dois exemplos que deixam claro a afirmação acima.

1.1.1 Reconstrução de uma Força Desconhecida

Considere um sistema mecânico com uma dinâmica de forças atuantes, cujas medidas não possam ser realizadas diretamente. Assim, tais medidas são dadas pela instabilidade de algum tipo de dinamômetro e portanto, sujeitas a erros.

Um modelo simplificado, onde um único grau de liberdade no sistema mecânico é considerado, pode ser descrito pela equação diferencial ordinária

$$m\ddot{y} + ky = x(t), \quad t > 0, \quad (1.6)$$

onde, m é a massa do sistema, k é a constante de rigidez, \ddot{y} indica segunda derivada, y é a função que descreve o deslocamento e x é a função descrevendo as forças atuando no sistema.

O **problema direto** associado é: dado a dinâmica de forças atuantes $x(t)$, encontrar y solução da EDO (1.6).

Exercício 1.3. *Determine a solução da EDO (1.6), no caso do sistema de forças $x(t) = 0$.*

O **problema inverso** associado é: encontrar a dinâmica de forças x , a partir de medidas das vibrações respostas do sistema, que são obtidas por medições e, portanto, sujeitas a erros.

Caso y seja conhecido exatamente, recuperamos x facilmente, bastando para tal, substituir y na EDO (1.6).

Mas, no caso em que y é contaminado por uma função de ruídos η , isto é, só temos informações sobre $y^\delta = y + \eta$, não podemos substituir y^δ diretamente na EDO (1.6), pois, não sabemos se a função η é duas vezes diferenciável. Ainda que η seja duas vezes diferenciável, $m\ddot{\eta}$ pode ser altamente oscilatória. Esta oscilação leva a soluções muito ruins.

Por exemplo, tome $m = k = 1$ e $y(t) = \exp(-t)$ como sendo a solução da EDO (1.6), (solução para o sistema de forças atuantes $x(t) = 0$). Suponha que conhecemos somente o dado perturbado

$$y^\delta(t) := \exp(-t) + a \sin(\omega t), \quad t > 0.$$

Substituindo y^δ na EDO (1.6), obtemos como resposta

$$x^\delta(t) := a(1 - \omega^2) \sin(\omega t), \quad t > 0.$$

Note que a função $x^\delta(t)$ está muito longe da solução para dados sem ruídos, se ω é muito grande.

Exercício 1.4. *Calcule o erro cometido na solução do problema acima como uma função de ω . Compare com o erro nos dados. O que podemos concluir (em função de w)? Use a norma no espaço das funções contínuas, isto é, a norma do supremo $\|\cdot\|_\infty$.*

Exercício 1.5. *(O conjunto das funções quadrado integráveis em $[a, b]$) Denote por $C[a, b]$ o conjunto das funções contínuas no intervalo $[a, b]$ (e assim, uniformemente contínuas (prove!)). Considere a*

seguinte função:

$$\begin{aligned} \|\cdot\|_2 : C[a, b] \times C[a, b] &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \|x(t) - y(t)\|_2 = \left(\int_a^b (x(t) - y(t))^2 dt \right)^{\frac{1}{2}}. \end{aligned}$$

i) Mostre que $\|\cdot\|_2$ é uma norma em $C[a, b]$.

ii) Mostre que $C[a, b]$ não é completo com a norma $\|\cdot\|_2$.

Definimos por $L^2[a, b]$ o completamento de $C[a, b]$ com relação a norma $\|\cdot\|_2$. Tal completamento sempre existe. Veja [33, 32, 31].

1- Usando os resultados do exercício acima, faça a comparação com o erro nos dados e na solução usando a norma de L^2 .

2- O que podemos concluir (em função de w)?

3- Existe diferença em medir erros com a norma $\|\cdot\|_\infty$ e $\|\cdot\|_2$?

1.1.2 Diferenciação Numérica

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, i.e., para qualquer $\tau \in (0, 1)$ tomar

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

Um simples argumento com a desigualdade triangular fornece

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

Suponha que tenhamos o limitante

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

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

$$\|x^{\delta,h}(\tau) - x(\tau)\|_{\infty} \leq hE + \frac{\delta}{h}. \quad (1.7)$$

A equação (1.7) é típica em problemas inversos e reaparecerá em nossas notas. O que é importante, por agora, é entender o que a equação (1.7) quer nos ensinar. A estimativa (1.7) significa que temos dois termos nessa estimativa de erro: um devido a aproximação da aplicação inversa e o outro devido ao erro de medida. 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, (1.7) 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 (1.7). Ou seja, tomar

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

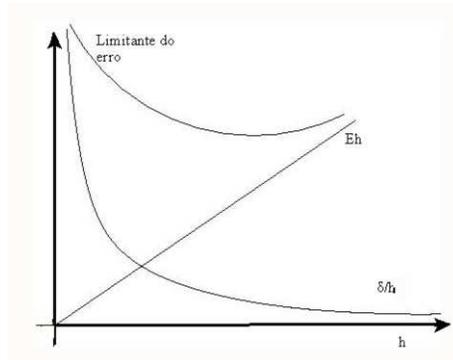


Figura 1.1: Estimativa de erro.

1.2 A Tomografia por Raio-X: Transformada de Radon

Entre os problemas que têm chamado a atenção da comunidade científica estão alguns problemas em Tomografia¹ [37, 38, 42]. As aplicações percorrem os campos de ciências médicas, que têm aplicações em detecção de tumores (tomografias por raio-X, EIT, etc)[10, 35, 34], ciências ambientais com a detecção de depósitos de sedimentos como, por exemplo, prospecção de petróleo (tomografias acústicas) [10, 34, 35, 37, 38, 42]. Estes problemas, em geral, consistem em recuperar a forma e a localização de um objeto imerso (ou de sua densidade) em uma região do plano \mathbb{R}^n a partir de medidas (parciais) sobre a fronteira da região. Essas medidas, na maioria dos casos, são adquiridas por um número reduzido de experimentos [18]. Assim, uma característica comum a esses tipos de problema é a falta de informações nos dados ou então, uma quantidade de ruídos muito

¹Existem várias formas de Tomografias. Nessas notas, daremos ênfase especial a Tomografia por Raio-X e tomografia de impedância elétrica (EIT).

grande nesses dados. Problemas como os citados acima persistem como área de intensa pesquisa em problemas inversos, dada a grande utilidade comercial e de aplicações destas técnicas, bem como, sua complexidade matemática.

A princípio, formularemos o problema para a tomografia por raio-X. Mais adiante, abordaremos o problema da tomografia de impedância elétrica (EIT).

1.2.1 Tomografia Computadorizada: Caso Contínuo

Os exames radiológicos, também conhecidos como tomografias por raio-X, utilizam feixes de raios-X que são disparados por uma máquina e são transmitido através do paciente. Estes feixes impressionam um filme radiológico, o qual, uma vez revelado, proporciona uma imagem que permite distinguir estruturas e tecidos com propriedades de absorção diferenciadas. Durante o exame radiográfico, os raios-X interagem com os tecidos através de efeitos fotoelétrico e radioativos. Dada as propriedades radioativas dos feixes de raio-X, a interação prolongada ou em grande quantidade destes feixes com os tecidos do corpo humano produzem efeitos nocivos a saúde, em particular, em tecidos já debilitados por ações infecciosas (câncer). Por isso, as novas tecnologias em exames radiológicos procuram alternativas mais eficazes e menos prejudiciais ao ser humano, como forma de diagnósticos de tumores. Mas, o raio-X possibilitou o surgimento de exames como a tomografia axial computadorizada que, com ajuda do computador, é capaz de fornecer imagens em vários planos, de forma rápida e razoavelmente precisa, utilizando quantidades mínimas de radiação, a qual é mais eficiente e menos nociva que o raio-X tradicional. Existe uma vasta literatura sobre o assunto e fica impossível citar todas. Para os interessados no assunto, recomendamos uma

pesquisa em [4, 10, 15, 16, 34, 35, 37, 38, 42] e citações.

Descrição do Problema:

Seja f uma função que descreve a densidade de um meio \mathbb{D} em uma região $\Omega \subset \mathbb{R}^2$. Seja L a reta pela qual está direcionado o feixe de ftons de raio-X. Chamaremos de I_0 a intensidade da aplicação (input) do feixe de ftons de raio-X e por I a intensidade após atravessar a região Ω (output). Veja a Figura 1.2.

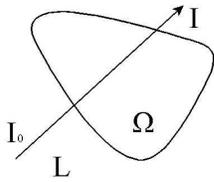


Figura 1.2: Raio-X.

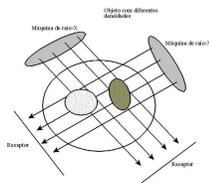


Figura 1.3: Máquina de raioX

Ao percorrer uma distância Δx ao longo de L , a combinação dos efeitos de espalhamento e absorção, implicam em uma atenuação de intensidade ΔI do feixe de photons de raio-X que passa pelo meio.

Esse fenômeno (considerando-se os efeitos de espalhamento nulos) pode ser modelado como

$$\frac{\Delta I}{I} = -f(x)\Delta x, \quad (1.8)$$

onde $f(x)$ é chamado de coeficiente de absorção.

Fazendo $\Delta x \rightarrow 0$ em (1.8), temos que

$$\frac{dI}{I} = -f(x)dx. \quad (1.9)$$

Integrando (1.9) ao longo da reta L obtemos

$$I = I_0 \exp\left(-\int_L f(z)dz\right), \quad (1.10)$$

onde L é o comprimento do feixe.

Observação: A equação (1.10) nos ensina que a intensidade de raio-X é atenuada exponencialmente, ao longo de L .

Considere $I_0 = 1$ (normalizado) na equação (1.10). Para-

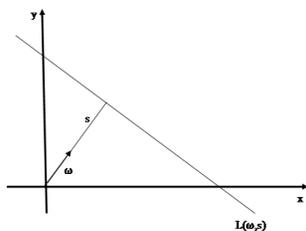


Figura 1.4: Parametrização.

metrizando a reta $L(w, s)$ que é perpendicular a w , e que está a uma distância s da origem, veja Figura 1.4, obtemos

$$\mathcal{R}f(w, s) = \int_{L(w,s)} f dl, \quad s \in \mathbb{R}, \quad \|w\|_2 = 1. \quad (1.11)$$

$\mathcal{R}f$ é chamada de *Transformada de Radon* bi-dimensional de f . Esta é uma aplicação que leva funções de $x \in \mathbb{R}^2$ em funções de $(w, s) \in \mathbb{S}^1 \times \mathbb{R}$, onde \mathbb{S}^{n-1} denota a esfera unitária em \mathbb{R}^n .

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 . Assim, a solução do problema inverso de

reconstruir a densidade f , consiste em inverter o operador \mathcal{R} (em um espaço de funções adequado).

A pergunta é exatamente a mesma feita por *Johann Radon* em seu famoso trabalho de 1917 [41]. É possível determinar o coeficiente de absorção f a partir de sua integral de linha (1.11) ao longo da reta L ?

Observação: O problema inverso associado a tomografia computadorizada é um problema inverso de identificação.

A aplicação desta idéia em detecção de tumores é imediata. Suponha que um médico deseja obter informações sobre a presença de anomalias em diferentes tecidos do corpo humano. A presença de tais anomalias implicam em diferenças nos coeficientes de absorção dos tecidos. Assim, determinar o coeficiente de absorção f , significa determinar a presença de anomalias ou não.

A tentativa de responder a pergunta feita por Radon, 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 $\mathcal{R}f(w, s) = 0$ para $|s| > \rho$ e para todo $w \in \mathbb{S}^1$, é verdade que $f(x, y) = 0$ para todo $x \in \mathbb{R}^2$ com $|x| > \rho$.

Exercício 1.6. *Verifique que a recíproca da questão (4) é verdadeira.*

Do ponto de vista prático, é impossível obter medidas sobre todas as retas que passam por Ω . Deste fato, surge uma outra pergunta, a qual, talvez, seja a mais difícil. **Quantas medidas são necessárias para obtermos uma boa aproximação?** Esse ponto não será abordado nessas notas. Para mais detalhes consulte [18]. Na

próxima seção, apresentaremos uma abordagem discreta da Tomografia Computadorizada onde aparecerá, naturalmente, tal questão.

Agora, abordaremos as questões (1) - (4) do ponto de vista matemático. Os resultados que apresentaremos podem ser encontrados em [44] e referência. Para isso faremos uso da Transformada de Fourier. Para os que não tenham tido contato com essa importante ferramenta matemática, aqui está uma ótima oportunidade para conhecer. Recomendamos uma breve olhada no Apêndice A dessas notas e referências.

O conjunto

$$L(w, s) = \{x \in \mathcal{R}^n : \langle x, w \rangle = s\} \quad (1.12)$$

é um hiperplano a uma distância $|s|$ da origem com vetor normal $w \in \mathbb{S}^{n-1}$. Veja Figura 1.4.

Dada $f \in \mathcal{S}(\mathbb{R}^n)$, considere a integral de superfície da função f ao longo do hiperplano $L(w, s)$, isto é:

$$\mathcal{R}f(w, s) = \int_{L(w, s)} f dS. \quad (1.13)$$

Exercício 1.7. *Mostre que a integral de superfície (1.13) é convergente para $f \in \mathcal{S}(\mathbb{R}^n)$.*

Exercício 1.8. *Mostre que*

$$\mathcal{R}f(w, s) = \mathcal{R}f(-w, -s). \quad (1.14)$$

Note que, dado $k \in \mathbb{N}$

$$\int_{\mathbb{R}} s^k \mathcal{R}f(w, s) ds = \int_{\mathbb{R}} \int_{L(w, s)} (\langle x, w \rangle)^k f(x) dS ds. \quad (1.15)$$

Quando calculamos $(\langle x, w \rangle)^k$, obtemos um polinômio homogêneo de grau k em w , com coeficientes que também são polinômios homogêneos de grau k na variável x . Já, o resultado da integral (1.15) é um polinômio de grau k em w .

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

$$F(w, s) = F(-w, -s) \quad (1.16)$$

e

$$\int_{\mathbb{R}} s^k F(w, s) ds = P_k(w) \quad (1.17)$$

é um polinômio homogêneo de grau k em w .

Exercício 1.9. *Demonstre que*

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

onde $w^\perp = \{\theta \in \mathbb{S}^{n-1} : \langle w, \theta \rangle = 0\}$.

Exercício 1.10. *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 $f(x) = g(\|x\|)$, $\forall x \in \mathbb{R}^n$. Mostre que, se f é radial então*

$$\begin{aligned} \mathcal{R}f(w, s) &= \int_{\mathbb{R}^{n-1}} g(\sqrt{s^2 + y^2}) dy \\ &= 2 \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2)} \int_0^\infty g(\sqrt{s^2 + t^2}) t^{n-2} dt. \end{aligned} \quad (1.18)$$

Lema 1.2.1. *Seja $f \in \mathcal{S}(\mathbb{R}^n)$. Então a transformada de Fourier de*

$\mathcal{R}f(w, s)$ com relação a f satisfaz

$$\widehat{\mathcal{R}f}(w, \lambda) = \int_{\mathbb{R}} \mathcal{R}f(w, s) e^{-i\lambda s} ds = \hat{f}(\lambda w). \quad (1.19)$$

Demonstração : Seja $k \in \mathbb{N}$. Pela desigualdade de Cauchy-Schwarz $|s|^k = |\langle w, s \rangle|^k \leq |x|^k$. Assim,

$$|s|^k |\mathcal{R}f(w, s)| \leq \int_{L(w, s)} |\langle w, s \rangle|^k |f(x)| dS \leq \int_{L(w, s)} |x|^k |f(x)| dS < \infty.$$

Com isso a integral (1.19) converge (justifique). Pelo Teorema de Fubini [33],

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{R}f(w, s) e^{-i\lambda s} ds &= \int_{\mathbb{R}} e^{-i\lambda s} \left(\int_{L(w, s)} f(x) dS \right) ds \\ &= \int_{\mathbb{R}} \left(\int_{L(w, s)} e^{-i\lambda \langle w, x \rangle} f(x) dS \right) ds \\ &= \int_{\mathbb{R}^N} e^{-i\lambda \langle w, x \rangle} f(x) dx = \hat{f}(\lambda w) \end{aligned}$$

■

Exercício 1.11. Mostre que \mathcal{R} é uma aplicação linear.

Lema 1.2.2. Seja $f \in \mathcal{S}(\mathbb{R}^n)$. Então $\mathcal{R}f(w, s) = 0$ se e só se $f(w, s) = 0$.

Demonstração : Como a Transformada de Fourier é um isomorfismo, temos que $\mathcal{R}f(w, s) = 0$ se e só se $\widehat{\mathcal{R}f}(w, s) = 0$. Por (1.19) se e só se $\hat{f}(\lambda w) = 0$. Novamente pelo isomorfismo da Transformada de Fourier, se e só se $f = 0$. ■

Exercício 1.12. *Seja*

$$P : \mathbb{S}^{n-1} \times (0, \infty) \longrightarrow \mathbb{R}^n$$

$$(w, s) \longmapsto sw = x.$$

- i) *Mostre que P é um isomorfismo e $P^{-1}(x) = (x/\|x\|, \|x\|)$ 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 $F \circ P^{-1} = F(x/\|x\|, \|x\|) \in C^\infty(\mathbb{R}^n - \{0\})$.*
- 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:*

$$\frac{\partial^k}{\partial s^k} F(w, s) \Big|_{s=0} = P_k(w),$$

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

Defina o conjunto

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

Temos:

Teorema 1.2.3. *A transformada de Radon é uma aplicação bijetora*

$$\mathcal{R} : \mathcal{S}(\mathbb{R}^n) \longrightarrow M := \left\{ \mathcal{S}(\mathbb{S}^{n-1} \times \mathbb{R}) : F \text{ satisfaz (1.16) e (1.17)} \right\} \quad (1.20)$$

Demonstração : A injetividade foi provada no Lema 1.2.2. Seja

$F(w, s) \in M$. Como

$$\int_{\mathbb{R}} s^k F(w, s) ds = \widehat{s^k F}(w, 0) = i^k \frac{\partial^k}{\partial s^k} \widehat{F}(w, 0)$$

é, por hipótese, um polinômio homogêneo de grau k . Fazendo $\xi = \lambda w$, pelo exercício 1.12, $f(\xi) = \widehat{F}(\xi/\|\xi\|, \|\xi\|) \in C^\infty(\mathbb{R}^n)$. Como $F \in M$, obtemos que $f \in \mathcal{S}(\mathbb{R}^n)$. De (1.19) deduzimos que $F = \mathcal{R}\mathcal{F}^{-1}f$.

■

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 encontrada em (1.19).

Exercício 1.13. *Mostre que:*

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

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 [44]).

Teorema 1.2.4. *Seja $n > 1$ ímpar. Então*

$$f(x) = 2^{-n} (i\pi)^{1-n} \int_{\mathbb{S}^{n-1}} \frac{\partial^{n-1}}{\partial s^{n-1}} \mathcal{R}f(w, \langle x, w \rangle) dS \quad (1.21)$$

Demonstração : Usando a Transformada de Fourier inversa em coordenadas polares e o fato de $n - 1$ ser par temos que

$$\begin{aligned} f(x) &= \frac{1}{(2\pi)^n} \int_{\mathbb{S}^{n-1}} \int_0^\infty e^{i\langle x, \lambda w \rangle} \hat{f}(\lambda w) \lambda^{n-1} d\lambda dS_w \\ &= \frac{1}{2(2\pi)^n} \int_{\mathbb{S}^{n-1}} \int_{-\infty}^\infty e^{i\langle x, \lambda w \rangle} \hat{f}(\lambda w) \lambda^{n-1} d\lambda dS_w. \end{aligned}$$

Usando (1.19) e o exercício 1.13 concluímos a demonstração. ■

A fórmula (1.21) possui propriedades interessantes que devem ser destacadas. Para obtermos $f(x_0)$, é necessário que se conheça os valores de $\mathcal{R}f(w, s)$ para $s = \langle w, x_0 \rangle$. Ou seja, não precisamos conhecer as integrais de f ao longo de todos os planos $L(w, s)$. Basta obtermos informações sobre os que distam $\langle w, x_0 \rangle$ da origem.

Note que a demonstração do Teorema 1.2.4 não é verdadeiro para o caso de n ser par. Neste caso, é preciso introduzir a transformada de Hilbert

Definição 1.2.5 (Transformada de Hilbert). *Seja $f \in \mathcal{S}(\mathbb{R}^n)$. Então*

$$\mathcal{H}f(s) = \mathcal{F}^{-1}(-i \cdot \text{sign}(\lambda) \hat{f}(\lambda)) \quad (1.22)$$

Teorema 1.2.6. *Seja $n > 1$ par. Então*

$$f(x) = 2^{-n} (i\pi)^{1-n} \int_{\mathbb{S}^{n-1}} \left(\frac{\partial^{n-1}}{\partial s^{n-1}} \mathcal{H}\mathcal{R}f \right) (w, \langle x, w \rangle) dS. \quad (1.23)$$

Demonstração : Veja [44]. ■

A idéia do que segue é obter uma fórmula para \mathcal{R}^{-1} para no caso especial $n = 2$. Para tal, segue um resultado importante que relaciona o suporte da Transformada de Radon com o suporte da função f .

Teorema 1.2.7. *Seja $f \in C^\infty(\mathbb{R}^n)$ tal que $f(x) \rightarrow 0$ mais rapidamente que qualquer polinômio. Se $\mathcal{R}f(w, s) = 0$ para $|s| > \rho$ então $f(x) = 0$ para $\|x\| > \rho$.*

Demonstração : Veja [44]. ■

Para nosso entendimento mais profundo, considere Ω como um círculo de raio ρ . Ainda supomos que f é axial-simétrica com respeito a origem. Assim, toda a informação que necessitamos saber está na direção $w_0 = (0, \pm 1)$.

Deste modo, podemos assumir que

$$f(w, s) = f(s), \quad 0 < s \leq \rho \quad \|w\| = 1.$$

Exercício 1.14. *Suponha as hipóteses do Teorema 1.2.7 satisfeitas e $0 < s \leq \rho$. Use (1.18) para mostrar que $\mathcal{R}f(w_0, s)$ satisfaz uma equação integral de Abel de primeira espécie*

$$\mathcal{R}f(w_0, s) = s \int_s^\rho r f(r) / (\sqrt{r^2 - s^2}) dr. \quad (1.24)$$

Teorema 1.2.8. *Suponha que $\mathcal{R}f(w_0, \rho) = 0$. Então*

$$f(s) = -\pi^{-1} \int_r^\rho \frac{d/ds(\mathcal{R}f(w_0, s))}{\sqrt{s^2 - r^2}} ds. \quad (1.25)$$

Demonstração : A demonstração fica como exercício. Qualquer dúvida consulte [44]. ■

O que podemos aprender de (1.21) e de (1.25) é que, ambas as fórmulas envolvem derivadas de $\mathcal{R}f$. Pelo que vimos na Seção 1.1, isto é uma indicação de um problema mal-posto.

Claro que, por exemplo, na equação (1.25), depois de derivarmos, fazemos uma operação suavizante novamente, a saber, integramos. Por outro lado, o kernel² em (1.25) 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. Desenvolveremos alguns desses métodos nos Capítulos 2, 3 e 4. Um tratamento completo pode ser encontrado em [37, 38].

²Aqui, kernel é diferente do conceito de núcleo de uma operação linear.

1.2.2 Tomografia Computadorizada: Caso Discreto

Nesta seção faremos uma breve discussão do aspecto prático da Tomografia Computadorizada e da Transformada de Radon. Como comentado anteriormente, do ponto de vista prático, é impossível obter a integral de linha de f (veja (1.11)) em todas as direções. Na verdade, o conjunto de informações obtidas num processo de Tomografia são valores da Transformada de Radon medidos em M detectores, que denotaremos por $\{p_1, \dots, p_M\}$, distribuídos em torno do objeto de interesse. Veja Figura³ 1.6.

As limitações físicas do processo de medição implicam em uma discretização da imagem tomográfica, isto é, o tamanho e o número N de pixels dentro do campo de visão, que devem ser reconstruídos, consiste de um vetor de variáveis desconhecidas f_j , com $j = \{1, \dots, N\}$. Os f_j são os coeficientes de atenuação. A figura 1.5 mostra, esquematicamente, uma imagem tomográfica a ser reconstituída.

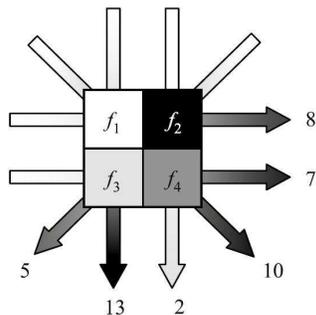


Figura 1.5: Tomografia discreta.

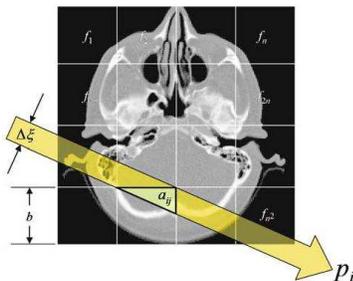


Figura 1.6: Feixe de raio-X.

Fisicamente, cada feixe de raio-X possui uma espessura. Quando o feixe de raio-X passa pela região imageada, temos que levar em

³As Figuras 1.5 e 1.6 foram retiradas de [5].

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 1.6 ilustra a situação.

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

$$a_{ij} = \frac{\text{área iluminada do pixel } j \text{ pelo raio } i}{\text{área total do pixel } j}. \quad (1.26)$$

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

$$\sum_{j=1}^N a_{ij} f_j = p_i \quad i = \{1, \dots, M\}. \quad (1.27)$$

Escrevendo na forma matricial, temos:

$$Af = p, \quad (1.28)$$

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

Fazendo uma comparação direta entre (1.11) e (1.28) obtemos:

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

Algumas das dificuldades de reconstrução no modelo discreto são:

- O sistema (1.28) possui uma solução exata, somente sobre con-

dições ideais. Para dados reais, a presença de ruídos nas medições das intensidades p_i implica em obtermos apenas soluções aproximadas do sistema (1.28), mesmo no caso em que $M = N$. No caso em que $M > N$, isto é, que temos mais informação (medidas p_i) que o número de densidades a serem determinadas f_j (sistemas sobredeterminados), possivelmente, obtemos reconstruções melhores para a densidade de cada pixel da imagem.

- Tipicamente, a matriz A é singular, em geral, é não-quadrada. Isto indica que o problema é mal-posto.
- A matriz A não possui uma estrutura simples. Assim, mesmo que A seja não-singular, é difícil determinar uma maneira de resolver o sistema (1.28) 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.

Exercício 1.15. *Com base na Figura 1.5 determine a matriz A associada. Para o vetor de intensidades p , determine uma solução f para o sistema (1.28).*

Exercício 1.16. *Com base no exercício acima, compare a solução $f = (f_1, f_2, f_3, f_4)$ e a Figura 1.5. O que você tem a dizer?*

Daremos mais detalhes de solução para problemas relacionados com tomografia discreta nos Capítulos 2, 3 e 4.

1.3 Tomografia Elétrica por Impedância - EIT

Por fim, introduzimos o problema da *Tomografia Elétrica de Impedância - Electrical Impedance Tomography - EIT*. Esse é um problema específico da classe de problemas chamados de *problemas de condutividade inversa* [1, 23]. Tais problemas possuem inúmeras aplicações em áreas das mais diversas. Em ciências médicas, com aplicações em detecção de tumores [4, 46], monitoramento de apnéias. Em geofísica e ciências ambientais, com a localização de depósitos de minerais, monitoramento de campos de fluídos. Na engenharia, com a detecção de corrosões em estruturas, perfis de dopagem em nano-condutores, etc. As vantagens sobre outros métodos, como tomografia por raio-X, por exemplo, é que a aplicação de correntes elétricas não causam efeitos colaterais, além de diminuir os custos de testes destrutivos [4, 23, 46].

A *tomografia elétrica por impedância - EIT* é um método de reconstruir propriedades da condutividade elétrica de algum objeto condutor, sabendo-se o valor de algumas medidas elétricas tomadas sobre a fronteira desse objeto. Do ponto de vista de aplicações, objetos diferentes imersos em uma região $\Omega \subset \mathbb{R}^N$, possuem condutividades elétricas diferentes (tecidos afetados por tumores são, em geral, mais densos que tecidos normais do corpo humano), e assim, absorvem mais corrente que outras partes.

Fisicamente, este modelo é descrito pelo sistema de equações de Maxwell para ondas eletromagnéticas de frequência ω , dadas por

$$\begin{aligned} \operatorname{rot} \vec{E} &= -i\omega\mu\vec{H} \\ \operatorname{rot} \vec{H} &= (\sigma + i\omega\varepsilon)\vec{E}, \end{aligned} \quad (1.29)$$

onde \vec{E}, \vec{H} são campos de vetores elétrico e magnético, σ, ε e μ são a condutividade, a permissibilidade elétrica e a permissibilidade magnética do meio, respectivamente.

No corpo humano, a permissibilidade magnética μ é muito pequena e pode ser considerada nula. Assim, obtemos de (1.29) que

$$\operatorname{rot} \vec{E} = 0 \quad \text{em } \Omega. \quad (1.30)$$

Exercício 1.17. *Assuma que Ω é uma região simplesmente conexa. Prove que, se um campo de vetores \vec{E} satisfaz (1.30), então existe uma função u tal que $\nabla u = \vec{E}$.*

A função u como no exercício é chamada de função potencial.

Exercício 1.18. *Prove que $\nabla \cdot (\operatorname{rot} \vec{H}) = 0$, para qualquer que seja o campo vetorial \vec{H} suficientemente diferenciável.*

Usando os exercícios 1.17 e 1.18 substituídos na equação (1.29), temos que a condutividade satisfaz o problema de Dirichlet

$$\begin{aligned} \nabla \cdot ((\sigma + i\omega\varepsilon)\nabla u) &= 0, & \text{em } \Omega \\ u &= f, & \text{na } \partial\Omega, \end{aligned} \quad (1.31)$$

onde a condição de bordo $u = f$ na $\partial\Omega$ significa que o potencial u é descrito pela função f (que é conhecida) no bordo de Ω .

O **problema direto** consiste em: dado Ω um domínio limitado em \mathbb{R}^N , que possui materiais com condutividade elétrica $a(x) = (\sigma(x) + i\omega\varepsilon(x))$ tal que $a(x) \geq \omega_0 > 0$, com fronteira $\partial\Omega \in C^1$ por partes, solucionar o *problema de Dirichlet*, i.e., encontrar a solução u do problema (1.31).

Para o problema direto temos o seguinte resultado:

Teorema 1.3.1. *Existe uma única função $u \in H^1(\Omega)$ solução do problema (1.31), para $a(x) > c > 0$ suficientemente suave.*

Demonstração : A demonstração usa técnicas de equações diferenciais parciais. Em especial de um teorema conhecido como Teorema de Lax-Milgram. Para detalhes nesta área veja [11]. ■

Assumindo que $a = \sigma + i\omega\varepsilon$ é constante nas proximidades da $\partial\Omega$ e que f é suave (e assim, u solução de (1.31) também é suave [23]), então os dados de Neumann

$$a \frac{\partial u}{\partial \eta} = h, \quad \text{na } \partial\Omega, \quad (1.32)$$

estão bem definidos.

O **problema inverso** é: reconstituir a condutividade elétrica $a(x)$, usando o conjunto de valores da corrente elétrica h (dados de Neumann) aplicada à fronteira de Ω e os correspondentes valores do potencial elétrico $u(x)$, na $\partial\Omega$ (dados de Dirichlet $u|_{\partial\Omega} = f$ medidos).

Em muitos problemas importantes ε pode ser considerado nulo e assim, o problema se reduz a determinar o coeficiente real $a = \sigma$. Por isso, tal problema recebe o nome de problema da condutividade inversa.

Uma observação importante, do ponto de vista prático e também matemático da aplicabilidade do processo de EIT é que, tecidos diferentes do corpo humano possuem condutividade diferente e, em geral, constantes. Por exemplo, a condutividade de músculos, pulmões, ossos e sangue são respectivamente, 8.0, 1.0, 0.06 e 6.7. Assim, o processo de determinar a condutividade de uma região é equivalente a determinar a forma desta região. Matematicamente, isto pode ser descrito como, determinar a função característica do conjunto que representa esta região.

O operador que aplica dados de Dirichlet e retorna dados de

Neumann é dito de operador Dirichlet-to-Neumann (DpN) e será denotado por Λ_a . Isto é:

$$\begin{aligned} \Lambda_a : H^{\frac{1}{2}}(\partial\Omega) &\longrightarrow H^{-\frac{1}{2}}(\partial\Omega) \\ f &\longmapsto \Lambda_a(f) := \left(a \frac{\partial u}{\partial \eta}\right) = h. \end{aligned} \quad (1.33)$$

O operador que aplica dados de Neumann e retorna dados de Dirichlet é dito Neumann-to-Dirichlet (NpD).

Exercício 1.19. Use o Teorema 1.3.1 para mostrar que o operador (DpN) como em (1.33) está bem definido, é linear e limitado.

Observação: Este exercício exige o conhecimento do Teorema do Traço. Veja [11].

Na prática, é impossível medir os dados sobre toda a fronteira de Ω . Assim, o problema de reconstrução deve ser encarado com o conhecimento de uma quantidade limitada de dados, obtida por medidas sob a fronteira de Ω . A Figura⁴ 1.7 mostra um exemplo real de tomografia por EIT.

Antes de seguirmos, vamos responder a duas questões:

1. Quantos dados necessitamos conhecer, i.e., se o parâmetro a em (1.31) é unicamente determinado pelo conhecimento de alguns pares de dados $(f, \Lambda_a(f) = h)$ sobre a fronteira ou pelo operador (NpD)?
2. O problema inverso é mal posto no sentido de Hadamard?

As respostas desta perguntas possuem resultados interessantes tanto do ponto de vista matemático, como do ponto de vista de aplicações. O primeiro trabalho nessa área deve-se a Alberto Calderón publicado em 1980 no jornal Sociedade Brasileira de Matemática sob o título *On an inverse boundary value problem* [6]. Este trabalho motivou o desenvolvimento desta área específica dos problemas inversos.

⁴Figuras retiradas de Wikipedia

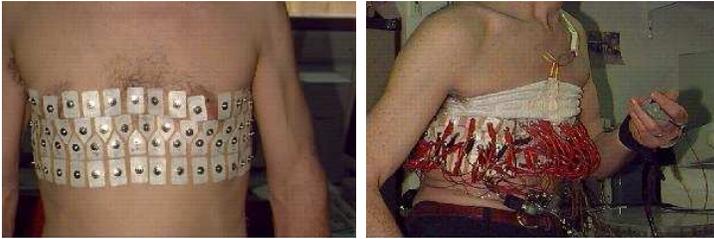


Figura 1.7: Eletrodos.

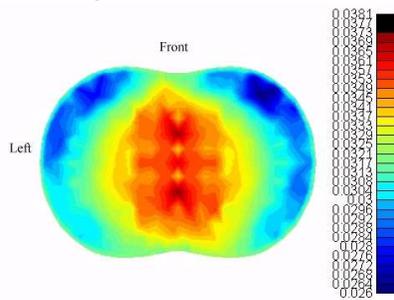


Figura 1.8: Dados obtidos pela EIT

Os resultados de identificabilidade e estabilidade que apresentaremos podem ser encontrados em [23, 45, 50]. Existem resultados mais modernos e que generalizam os que apresentaremos. Veja referências em [45, 50].

Apresentaremos agora alguns resultados sobre unicidade e estabilidade para o problema inverso. Não faremos demonstração, uma vez que as técnicas usadas fogem ao escopo destas notas. Interessados podem procurar os resultados na bibliografia.

As perguntas que temos que responder são:

1. **Unicidade:** Se $\Lambda_{a_1} = \Lambda_{a_2}$, então $a_1 = a_2$?
2. **Estabilidade:** Se Λ_{a_1} está próximo de Λ_{a_2} , então a_1 e a_2 estão

próximos?

3. **Reconstrutibilidade:** Dado as medidas na fronteira Λ_a , então existe uma maneira de recuperar a condutividade a ?
4. **Dados parciais:** Se Γ é uma parte da fronteira de Ω e se $\Lambda_{a_1}(f) = \Lambda_{a_2}(f)$ restrito a Γ , então $a_1 = a_2$?

Observação 1.3.2. *O problema inverso de reconstruir a condutividade elétrica é altamente não linear e mal posto, i.e., o operador não depende continuamente dos dados (veja [1, 3, 4, 23]).*

Partindo do estudo feito por Calderón em 1980, intensa pesquisa nessa linha foi estabelecida. Já existe uma quantidade significativa de resultados e ainda muitos problemas em aberto. Para o caso de $\Omega \subset \mathcal{R}^3$ e condutividades regulares, isto é, pertencentes a $C^2(\overline{\Omega})$, os seguintes resultados são conhecidos. As demonstrações dos resultados abaixo podem ser encontradas em [45] e referências.

Para **unicidade:**

Teorema 1.3.3 (Sylvester-Ulmann 1987). *Se $\Lambda_{a_1}(f) = \Lambda_{a_2}(f)$, então $a_1 = a_2$ em Ω .*

No contexto de **estabilidade:**

Teorema 1.3.4 (Alessandrini 1988). *Seja $a_j \in H^s(\Omega)$ para $s > n/2 + 2$ e assumamos que $\|a_j\|_{H^s(\Omega)} \leq M$ e $1/M \leq a_j \leq M$ para $j = 1, 2$. Então*

$$\|a_1 - a_2\|_{L^\infty(\Omega)} \leq c \|\Lambda_{a_1} - \Lambda_{a_2}\|_{\mathcal{L}(H^{1/2}(\Omega), H^{-1/2}(\Omega))}.$$

Resultados de **reconstrução** são devidos a Nachman 1988:

Teorema 1.3.5. *Existe um algoritmo convergente para reconstruir a a partir de Λ_a .*

No caso de unicidade com **dados parciais** os resultados são bem mais recentes e devem-se a Kenig-Sjöstrand-Uhlmann 2007:

Teorema 1.3.6. *Assuma que Ω é convexo e Γ é qualquer subconjunto aberto de $\partial\Omega$. Se $\Lambda_{a_1}(f) = \Lambda_{a_2}(f)$ em Γ para toda $f \in H^{1/2}(\Omega)$, e se $a_1|_{\partial\Omega} = a_2|_{\partial\Omega}$, então $a_1 = a_2$ em Ω .*

1.4 Breve Apanhado Histórico do Capítulo

O estudo de problemas inversos é muito novo e também muito antigo. Há cerca de dois milênios atrás, no livro VII do diálogo “*República*”, Platão propôs o filosófico problema de reconstruir a “realidade” através da observação da imagem de objetos ou pessoas, cujas sombras eram projetadas na penumbra de uma caverna. Com a idéia de discutir aspectos filosóficos das fontes de conhecimento humano, Platão, também acabou introduzindo o primeiro exemplo de problemas inversos que se tem relato [39].

Em ciências aplicadas, possivelmente, um dos primeiros problemas inversos de que se tem notícia data de 200 anos antes de Cristo. Eratostenes propôs o problema de determinar o diâmetro da terra a partir de medições feitas em duas cidades do tamanho da sombra de um bastão (ao meio dia do dia mais longo do ano), e do conhecimento da distância entre essas cidades [15].

Em 1800, Gauss fez uso do método de quadrados mínimos para reconstruir a órbita de um cometa a partir de dados de órbitas anteriores. O método de quadrados mínimos será melhor estudado mais diante no texto.

A transformada, que hoje em dia chamamos de *Transformada de Radon*, é uma das precursoras do estudo de tomografia por raio-X [38, 37]. Sua formulação matemática foi publicada por Radon em 1917 no famoso trabalho [41]. Em 1961, a solução do problema foi

aplicado pela primeira vez a uma sequência de raios-X onde as intensidades aplicadas a um objeto foram medidas sobre diferentes direções. Allen MacLeod Cormack (1924 - 1998) e Sir Godfrey Hounsfield (1919 - 2004) foram os pioneiros em aplicar essas técnicas à imagens médicas e em 1979 receberam o *Prêmio Nobel de Fisiologia/Medicina* pelas suas contribuições e estudos entre 1960 e 1970.

O problema tratado por Calderón foi o de determinar a condutividade elétrica de um meio fazendo-se medidas de voltagem e corrente na fronteira deste meio. Este problema inverso é conhecido hoje em dia como Electrical Impedance Tomography (EIT). Calderón foi motivado por problemas de prospecção de petróleo, quando trabalhava como engenheiro para a empresa Yacimientos Petrolíferos Fiscales (YPF) do governo Argentino nos anos 40. Seus resultados desta época ficaram sem ser publicados por muito tempo. Em 1980, A. P. Calderón publicou um pequeno paper intitulado "On an inverse boundary value problem" [6] publicado em Brazilian Mathematical Society (SBM) in ATAS of SBM (Rio de Janeiro), pp. 65-73, 1980, onde discutiu resultados de identificabilidade para o problema de EIT. Hoje em dia alguns autores chamam o problema de condutividade inversa como problema de Calderón.

Nas últimas quatro décadas, um novo campo de estudos na área de matemática aplicada tem conquistado um grande número de pesquisadores adeptos. Este campo trata de problemas como os formulados por Platão, Eratóstenes, entre outros, cuja abordagem exige o desenvolvimento de métodos matemáticos como os apresentados por Gauss, Radon e Calderón. A essa área de estudos denominamos *Problemas Inversos*. O súbito crescimento deve-se, certamente, ao grande número de aplicações em outras ciências, como, por exemplo, em geofísica e ciências ambientais (como detecção de depósito de petróleo e outras riquezas ou poluentes no subsolo), ciências médicas

(com ênfase na reconstrução de imagens, tomografia), entre outras, bem como do aparato de novas técnicas e teorias matemáticas envolvidas no ataque a tais problemas.

Citar todas as referências sobre o assunto fica impossível, dada a vastidão das mesmas. No entanto, na Bibliografia destas notas segue uma lista considerável de referências sobre o assunto, onde os interessados em tais problemas podem basear a sua pesquisa na área.

- Platão (428 - 427 a.C.) - filósofo e matemático Grego, autor de diversos diálogos filosóficos e fundador da Academia em Atenas. Tal academia foi a primeira instituição de educação superior do mundo ocidental.
- Eratóstenes (285 - 194 a.C.) - matemático, bibliotecário e astrônomo grego.
- Johann Carl Friedrich Gauss (1777 - 1855) - matemático, astrônomo e físico alemão. Muitos consideram Gauss o maior gênio da história da matemática. Conhecido como o príncipe dos matemáticos.
- David Hilbert (1862 - 1943) - matemático Alemão, reconhecido como um dos mais influentes matemáticos dos séculos XIX e XX.
- Jaques Hadamard (1865 - 1963) - matemático Francês ficou mais conhecido pela prova do Teorema sobre números primos em 1896. Também, introduziu o conceito de problemas bem postos relacionados a teoria de Equações Diferenciais Parciais.
- Johann Peter Gustav Lejeune Dirichlet (1805 - 1859) - matemático alemão a quem se atribui a moderna definição formal de função.

- Johann Radon (1887 - 1956) - matemático austríaco. Foi aluno de David Hilbert.
- John von Neumann (1903 - 1957) - matemático húngaro de origem judaica, naturalizado norte-americano nos anos 30. Desenvolveu importantes contribuições em Teoria dos Conjuntos, Análise Funcional, Teoria Ergódica, Mecânica Quântica, Ciência da Computação, Economia, Teoria dos Jogos, Análise Numérica, Hidrodinâmica das explosões, Estatística e muitas outras áreas da Matemática. De fato é considerado um dos mais importantes matemáticos do século XX.
- Alberto Calderón (1920 - 1998) - matemático argentino, conhecido por seus trabalhos com a teoria das equações diferenciais ordinárias e parciais. Muitos autores chamam o problema de condutividade inversa de problema de Calderón.
- Allen MacLeod Cormack (1924 - 1998) - biólogo sulafricano. Recebeu o Nobel de Fisiologia/Medicina de 1979 pela participação na obtenção de diagnósticos de tumores usando tomografia axial computadorizada.
- Sir Godfrey Newbold Hounsfield (1919 - 2004) - engenheiro britânico. Recebeu o Nobel de Fisiologia/Medicina de 1979 pela participação na obtenção de diagnósticos de tumores usando tomografia axial computadorizada.

Capítulo 2

Métodos de Regularização Iterativos Tipo Gradiente

Como vimos no Capítulo 1, obter uma solução aproximada para um problema inverso (veja a equação (1.1)) que seja estável, isto é, que dependa continuamente dos dados do problema, implica em buscar uma maneira de aproximar o operador inverso por uma família de operadores contínuos (**regularização**). Veja o Apêndice C.

Uma alternativa para regularização de problemas inversos são os *métodos iterativos de regularização*. Tais métodos possuem a vantagem de apresentarem propriedades auto-regularizantes [3, 8, 25, 27, 46, 52]. Quando mencionamos métodos iterativos, nos vem em mente métodos conhecidos e bem desenvolvidos, os quais são amplamente utilizados para resolver problemas bem postos. Dentre estes,

podemos citar *métodos tipo gradiente*, *métodos tipo Newton*¹ e suas variantes. Neste capítulo, nos dedicaremos a alguns métodos do tipo gradiente, mais especificamente, ao método clássico de Landweber e algumas de suas modificações, ao método de steepest descent e a estratégias tipo Kaczmarz, estas últimas associadas a resolução de sistemas de equações mal-postas. No Capítulo 3 retornaremos a métodos tipo Newton.

Deve-se a Gauss a demonstração de que, a melhor maneira de se determinar um parâmetro desconhecido em uma equação do tipo (1.1), é minimizando a soma dos quadrados dos resíduos, isto é,

$$\frac{1}{2} \|A(x) - y\|^2 \rightarrow \min, \quad x \in \mathcal{H}_1. \quad (2.1)$$

Tal método é conhecido hoje em dia como **método de Mínimos Quadrados**.

Assumindo algumas propriedades do operador A , prova-se que o minimizador (caso exista) de (2.1) deve satisfazer a condição necessária de primeira ordem

$$A'(x)^* A(x) = A'(x)^* y. \quad (2.2)$$

A equação (2.2) é chamada de **Equação Normal**.

Exercício 2.1. *Seja A uma transformação linear de \mathbb{R}^n para \mathbb{R}^m . Mostre que a equação (2.2) é a condição de otimalidade (de primeira ordem) para o problema de mínimos quadrados (2.1).*

Generalize para o caso em que A é um operador linear entre espaços de Hilbert.

Observação: Para o caso de A ser um operador linear, as hipóteses sobre o operador A tal que (2.2) seja satisfeita estão

¹Um pouco mais sobre Newton, Gauss e Cauchy, veja Seção 2.7

esboçadas no Exercício 2.1. Para o caso em que $A(\cdot)$ é não linear, temos que exigir que $A(\cdot)$ possua uma derivada de Fréchet.

Uma possibilidade para encontrar uma solução de (2.2) é utilizar uma iteração de ponto fixo

$$x_{k+1} = \Phi(x_k) \quad (2.3)$$

para o operador

$$\Phi(x) = x + A'(x)^*(y - A(x)). \quad (2.4)$$

No entanto, Φ é, em geral, não contrativo.

Métodos iterativos para aproximação de pontos fixos de operadores não expansivos Φ , isto é,

$$\|\Phi(x) - \Phi(\tilde{x})\| \leq \|x - \tilde{x}\|, \quad x, \tilde{x} \in \mathcal{D}(\Phi)$$

têm sido considerados atualmente. Veja referências em [25]. Neste caso, maior ênfase é dada numa prova construtiva de pontos fixos para Φ .

Em muitos exemplos práticos (e até mesmo teóricos), é quase impossível verificar analiticamente quando o operador Φ é contrativo. No contexto de problemas não-lineares, esta dificuldade é ainda maior [8, 25].

Muitos métodos iterativos para resolver (1.1) são baseados na solução da equação normal (2.2), via sucessivas iterações, partindo de um chute inicial x_0 . Tal chute, em particular para problemas não-lineares, costuma conter informações *a priori* sobre a solução do problema.

Iniciaremos o estudo de métodos iterativos, considerando o caso em que o operador A é linear na Seção 2.1. O caso em que

A é não linear será abordado na Seção 2.2.

2.1 Método de Landweber Linear

Nessa seção, assumiremos que o operador A na equação (1.1) é linear e limitado. Com essa hipótese, uma maneira de resolvermos (2.2) é considerarmos a iteração

$$x_{k+1} = x_k + \gamma A^*(y - Ax_k), \quad k = 0, 1, 2, \dots \quad (2.5)$$

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 [29] 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 , denotaremos os iterados por x_k^δ , e a *iteração de Landweber* fica definida por

$$x_{k+1}^\delta = x_k^\delta + A^*(y^\delta - Ax_k^\delta). \quad (2.6)$$

Observação: Note que a equação (2.6) é equivalente a pré-multiplicar a equação $Ax = y^\delta$ por $\gamma^{\frac{1}{2}}$ e, então, iterar como em (2.5). Dada a limitação de γ , sem perda de generalidade, podemos supor que $\|A\| \leq 1$ e iterar como em (2.6).

Neste capítulo, a menos que se faça menção em contrário, admitiremos que $\|A\| \leq 1$.

Abaixo, apresentaremos resultados de convergência da iteração de Landweber (2.6), para o caso de A ser linear e limitado. As técnicas utilizadas na demonstração são relativamente simples e bem conhecidas de métodos iterativos para problemas diretos. Observamos ainda que, para o caso linear, obtemos a mesma iteração (2.6) partindo de

um chute inicial $x_0 = 0$.

2.1.1 Convergência

Nesta subseção provaremos que a sequência $\{x_k\}$ de iterados pelo método de Landeweber converge para uma solução aproximada do problema (1.1). 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 problema (1.1). Aqui A^\dagger denota a pseudo-inversa de A . Veja o Apêndice B para maiores detalhes e para uma interpretação geométrica da solução x^\dagger .

Começaremos dando condição necessárias e suficientes para a iteração (2.5) convergir.

Teorema 2.1.1. *Se $y \in \mathcal{D}(A^\dagger)$, então a sequência x_k gerada pela iteração de Landeweber (2.5) 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 o iterado x_k em (2.5) como

$$x_k = \sum_{j=0}^{k-1} (I - A^* A)^j A^* y. \quad (2.7)$$

Como $y \in \mathcal{D}(A^\dagger)$, então $A^* y = A^* A x^\dagger$. Assim,

$$x^\dagger - x_k = x^\dagger - A^* A \sum_{j=0}^{k-1} (I - A^* A)^j x^\dagger = (I - A^* A)^k x^\dagger. \quad (2.8)$$

Defina $r_k(\lambda) = (1 - \lambda)^k$. Como, por hipótese $\|A\| \leq 1$, segue que o espectro de $A^* A$ denotado por $\Sigma(A^* A)$ é um subconjunto de $(0, 1]$. Note 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

[28, 48, 36],

$$x^\dagger - x_k = r_k(A^*A)x^\dagger. \quad (2.9)$$

Logo, $x_k \xrightarrow{k \rightarrow \infty} x^\dagger$. ■

Exercício 2.2. *Complete os detalhes da demonstração do Teorema 2.1.1.*

Exercício 2.3. *Mostre que se $y \in \mathcal{D}(A^\dagger)$, então $A^*y = A^*Ax^\dagger$. Sugestão: use o que você aprendeu sobre a inversa generalizada. Consulte o Apêndice B.*

O Teorema 2.1.1 nos ensina que, x_k gerada pela iteração de Landweber converge para uma solução de mínimos quadrados da equação (1.1) 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 2.1.1 sabemos que a sequência x_k^δ diverge. A pergunta é: Qual é o fator de propagação destes erros?

Lema 2.1.2. *Sejam y, y^δ com $\|y - y^\delta\| \leq \delta$ e x_k e x_k^δ obtidos pelas respectivas iterações de Landweber (2.5) e (2.6). Então,*

$$\|x_k - x_k^\delta\| \leq \sqrt{k}\delta, \quad k \geq 0. \quad (2.10)$$

Demonstração : Pela linearidade de A temos

$$x_k - x_k^\delta = \sum_{j=0}^{k-1} (I - A^*A)^j A^*(y - y^\delta).$$

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. \blacksquare

Exercício 2.4. *Complete os detalhes da demonstração do Lema 2.1.2.*

Note que, na presença de erro nos dados temos

$$\|x^\dagger - x_k^\delta\| = \|A^\dagger y - x_k^\delta\| \leq \|A^\dagger y - x_k\| + \|x_k - x_k^\delta\|. \quad (2.11)$$

Esta é a estimativa fundamental para a iteração de Landweber.

A estimativa (2.11) nos ensina que o erro total possui duas componentes, um *erro de aproximação* que diminui lentamente e um *erro nos dados* que cresce com uma 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. Compare com a Figura 1.1.

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. Conseqüentemente, um critério de parada apropriado deve levar em conta a informação adicional do nível de ruídos δ .

Lema 2.1.3. *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

$$y^\delta - Ax_k^\delta = y^\delta - A(x_{k-1}^\delta + A^*(y^\delta - Ax_{k-1}^\delta)) = (I - A^*A)(y^\delta - Ax_{k-1}^\delta).$$

Como $\|I - A^*A\| \leq 1$, o lema segue. \blacksquare

Por outro lado, se $y^\delta \notin \mathcal{D}(A)$, pelo Teorema 2.1.1, a iteração x_k^δ diverge para infinito. Portanto, um resíduo pequeno não implica que a aproximação para solução é melhor. Veja a estimativa (2.11) e a Figura 1.1.

Exercício 2.5. *Faça um gráfico comparando os resultados obtidos nos Lemas 2.1.2 e 2.1.3 com a estimativa (2.11). Compare com a figura 1.1.*

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,

$$\|y^\delta - Ax_{k(\delta, y^\delta)}^\delta\| \leq \tau\delta, \quad \tau > 2 \text{ fixo.} \quad (2.12)$$

O próximo Teorema garante que, enquanto a discrepância (2.12) não é atingida, a aproximação para a solução não piora.

Teorema 2.1.4. *[Monotonia] Seja $y \in \mathcal{D}(A)$, x^\dagger a solução de norma mínima de (1.1) e y^δ satisfazendo $\|y - y^\delta\| \leq \delta$. Se (2.12) é satisfeita, então*

$$\|x_{k+1}^\delta - x^\dagger\| \leq \|x_k^\delta - x^\dagger\|. \quad (2.13)$$

Demonstração : Fazendo

$$\begin{aligned}
& \|x_{k+1}^\delta - x^\dagger\|^2 - \|x_k^\delta - x^\dagger\|^2 = 2\langle x_k^\delta - x^\dagger, x_{k+1}^\delta - x_k^\delta \rangle + \|x_{k+1}^\delta - x_k^\delta\|^2 \\
& = 2\langle x_k^\delta - x^\dagger, A^*(y^\delta - Ax_k^\delta) \rangle + \langle A^*(y^\delta - Ax_k^\delta), A^*(y^\delta - Ax_k^\delta) \rangle \\
& = 2\langle Ax_k^\delta \pm y^\delta - y, y^\delta - Ax_k^\delta \rangle + \langle y^\delta - Ax_k^\delta, AA^*(y^\delta - Ax_k^\delta) \rangle \\
& = 2\langle y^\delta - y, y^\delta - Ax_k^\delta \rangle - \|Ax_k^\delta - y^\delta\|^2 \\
& \quad - \langle y^\delta - Ax_k^\delta, (I - AA^*)(y^\delta - Ax_k^\delta) \rangle.
\end{aligned}$$

Como $I - A^*A$ é semi-definido positivo, segue que

$$\|x_{k+1}^\delta - x^\dagger\|^2 - \|x_k^\delta - x^\dagger\|^2 \leq \|Ax_k^\delta - y^\delta\| (2\|y - y^\delta\| - \|Ax_k^\delta - y^\delta\|).$$

Uma vez que y^δ satisfaz (1.2) e que $k < k_*$, a afirmativa segue.

■

Exercício 2.6. *Complete os detalhes na demonstração do Teorema 2.1.4.*

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 2.1.5. *Seja $\tau > 1$ em (2.12). Então, o princípio de discrepância determina um índice de parada $k_* = k_*(\delta, y^\delta)$, que é finito para a iteração de Landweber, com $k(\delta, y^\delta) = \mathcal{O}(\delta^{-2})$.*

Demonstração : Seja $\{x_k^\delta\}$ como em (2.5). Como no Teorema 2.1.4,

$$\begin{aligned}
& \|x^\dagger - x_k^\delta\|^2 - \|x^\dagger - x_{k+1}^\delta\|^2 = \|y - Ax_k^\delta\|^2 \\
& \quad - \langle y - Ax_k^\delta, (I - AA^*)(y - Ax_k^\delta) \rangle \geq \|y - Ax_k^\delta\|^2.
\end{aligned}$$

Somando, sobre $k = 1$ até j , e levando em conta a monotonia dos resíduos dado pelo Lema 2.1.3, 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 $y - Ax_j^\delta = (I - AA^*)^j(y - Ax_0)$, donde segue que

$$\|(I - AA^*)^j(y - Ax_0)\| = \|y - Ax_j^\delta\| \leq k^{-\frac{1}{2}}\|x^\dagger - x_1^\delta\|.$$

Assim,

$$\begin{aligned} \|y^\delta - Ax_j^\delta\| &= \|(I - AA^*)^j(y^\delta - Ax_0)\| \\ &\leq \|(I - AA^*)^j(y^\delta - y)\| + \|(I - AA^*)^j(y - Ax_0)\| \\ &\leq \delta k^{-\frac{1}{2}}\|x^\dagger - x_1^\delta\|. \end{aligned}$$

Conseqüentemente, 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 precisamos fazer hipóteses sobre a solução x^\dagger . Tais hipóteses são conhecidas como **condições de fonte**. Condições de fonte comuns na literatura impõem que, x^\dagger pertença a imagem do operador A^* , ou a imagem de potências do operador A^*A . Por exemplo, cosidere a condição de fonte

$$x^\dagger \in \mathcal{R}((A^*A)^\theta), \quad \theta \in \mathbb{R}_*^+. \quad (2.14)$$

Tal condição impõem que a solução x^\dagger pertença a imagem do operador $(A^*A)^\theta$.

Exercício 2.7. *Mostre que o operador $(A^*A)^{\frac{1}{2}}$ está bem definido. Os conjuntos domínio e imagem de $(A^*A)^{\frac{1}{2}}$ são subconjuntos de que*

espaço vetorial? Interprete geometricamente a condição de fonte (2.14).

Teorema 2.1.6. *Suponha que a condição de fonte (2.14) seja satisfeita para $\theta = 1/2$. Se $y \in \mathcal{R}(A)$ e o princípio de discrepância (2.12) é válido, então, a iteração de Landweber possui ordem de convergência $k(\delta, y^\delta) = \mathcal{O}(\delta^{-1})$.*

Demonstração : Veja [8, Teorema 6.5]. ■

2.2 Método de Landweber Não-Linear

Na seção anterior, fizemos uma análise completa de convergência para a iteração de Landweber no caso linear. Nesta seção, passaremos a natural generalização dos resultados obtidos na seção anterior, para o caso em que o operador A seja não-linear, o qual denotaremos por $A(x)$.

A ideia é aplicar o método de aproximações sucessivas para a equação de ponto fixo (2.4) e obter uma extensão natural para a iteração (2.6), no caso de problemas não lineares.

Assumiremos, durante esta seção, que o operador $A(x)$ possui uma derivada de Fréchet contínua (veja Apêndice A para definição), a qual denotaremos por $A'(\cdot)$.

A **iteração de Landweber não-linear** é definida por

$$x_{k+1}^\delta = x_k^\delta + A'(x_k^\delta)^*(y^\delta - A(x_k^\delta)), \quad (2.15)$$

com y^δ representando os dados com ruídos e satisfazendo (1.2).

Como é bem conhecido, métodos iterativos para problemas não lineares, em geral, não possuem a propriedade de convergirem globalmente. Assim, temos que dar certas condições, que são mais restritivas que as que foram usadas na Seção 2.1. Em especial, o

ponto inicial $x_0^\delta = x_0$ deve incorporar algum conhecimento **a priori** da solução exata do problema. Portanto, supomos que exista uma bola fechada $B_{2\rho}(x_0) \subset \mathcal{D}(A)$ de raio $2\rho > 0$ e de centro x_0 de forma que as hipóteses abaixo sejam satisfeitas.

Hipótese 2.1. *Assumiremos que:*

1. O operador $A(\cdot)$ é contínuo e o problema (1.1) possui uma solução x_* que pertence a $B_{2\rho}(x_0)$.
2. A derivada de Fréchet de $A(\cdot)$ satisfaz

$$\|A'(x)\| \leq 1 \quad x \in B_{2\rho}(x_0). \quad (2.16)$$

3. Ainda, necessitamos assumir que valha a condição local de não-linearidade chamada de **condição do cone tangente**

$$\begin{aligned} \|A(x) - A(\tilde{x}) - A'(x)(x - \tilde{x})\| &\leq \eta \|A(x) - A(\tilde{x})\|, \quad (2.17) \\ \eta &< \frac{1}{2}, \quad x, \tilde{x} \in B_{2\rho}(x_0). \end{aligned}$$

Como veremos abaixo, tais hipóteses são suficientes para provar que a iteração de Landweber (2.15) é bem definida e que converge a uma solução de (1.1).

Exercício 2.8. *Mostre que, se $A(x)$ satisfaz (2.17) então também satisfaz a estimativa*

$$\frac{1}{1+\eta} \|A'(x)(\tilde{x} - x)\| \leq \|A(\tilde{x}) - A(x)\| \leq \frac{1}{1-\eta} \|A'(x)(\tilde{x} - x)\|. \quad (2.18)$$

Exercício 2.9. *Mostre que, se $A'(x)$ é Lipschitz contínuo, então a*

estimativa de erros

$$\|A(x) - A(\tilde{x}) - A'(x)(x - \tilde{x})\| \leq c\|x - \tilde{x}\|^2 \quad (2.19)$$

é satisfeita. Sugestão: Use o Teorema de Taylor.

Observamos que, a estimativa (2.19) carrega muito pouca informação para provarmos convergência da iteração de Landweber (2.15). De fato, o lado esquerdo da estimativa (2.19) pode ser muito pequeno se comparado com o seu lado direito. Esse assunto será tratado nos próximos exercícios.

Exercício 2.10. *Suponha que $A(\cdot)$ é fracamente fechado, compacto e Fréchet diferenciável. Mostre que $A'(\cdot)$ é compacto. Sugestão: Mostre o exercício para o caso em que A é linear. Generalize para o caso não linear. Para ajuda, consulte [28].*

Exercício 2.11. *Mostre que, se $\{x_n\}$ é uma sequência limitada em um espaço de Hilbert, então existe uma subsequência $\{x_{n_k}\}$ que converge fraco para algum x . Sugestão: Esse resultado é conhecido como Teorema de Banach-Alaoglu, veja [28] ou o Apêndice A.*

Exercício 2.12. *Mostre que, se $A(\cdot)$ é um operador compacto e $\{x_n\}$ é uma sequência que converge fraco para x , então existe uma subsequência $\{x_{n_k}\}$ tal que $\{A(x_{n_k})\}$ converge para $A(x)$. Sugestão: Essa é uma das propriedades mais importantes dos operadores compactos. Veja [28] ou o Apêndice A.*

Exercício 2.13. *Seja A como nas hipóteses dos exercícios acima e $\{x_n\}$ uma sequência tal que $\|x_n - x\| \leq \varepsilon$ para todo $n \in \mathbb{N}$. Use os exercícios acima para concluir que o lado esquerdo da equação (2.19) converge para zero enquanto o lado direito permanece constante e igual a $c\varepsilon^2$.*

2.2.1 Convergência

Da mesma forma que no caso linear, quando os dados estão corrompidos por ruídos, devemos atentar para uma regra de parada para a iteração de Landweber (2.15). Essa é dada pelo princípio da discrepância (2.12), para uma escolha apropriada de τ . Com isso, podemos provar que a iteração possui propriedades de monotonia. Isto é:

Proposição 2.2.1 (Monotonia). *Suponha que a Hipótese 2.1 seja satisfeita e que $x_k^\delta \in B_\rho(x_*)$. Enquanto*

$$\|y^\delta - A(x_k^\delta)\| > 2 \frac{1 + \eta}{1 - 2\eta} \delta, \quad (2.20)$$

a iteração x_{k+1}^δ é uma aproximação melhor de x_ que x_k^δ . Ainda, enquanto (2.20) for satisfeita a iteração (2.15) está bem definida.*

Demonstração : Seja $x_k^\delta \in B_\rho(x_*)$. Usando a desigualdade triangular obtemos que $x_*, x_k^\delta \in B_{2\rho}(x_0)$. Portanto, x_* e x_k^δ satisfazem a Hipótese 2.1. Segue da definição da iteração (2.15) e de (1.2) que

$$\begin{aligned} \|x_{k+1}^\delta - x_*\|^2 - \|x_k^\delta - x_*\|^2 &= 2\langle x_{k+1}^\delta - x_k^\delta, x_k^\delta - x_* \rangle + \|x_{k+1}^\delta - x_k^\delta\|^2 \\ &= 2\langle y^\delta - A(x_k^\delta), A'(x_k^\delta)(x_k^\delta - x_*) \rangle + \|A'(x_k^\delta)^*(y^\delta - A(x_k^\delta))\|^2 \\ &\leq 2\langle y^\delta - A(x_k^\delta), y^\delta - A(x_k^\delta) + A'(x_k^\delta)(x_k^\delta - x_*) \rangle - \|y^\delta - A(x_k^\delta)\|^2 \\ &\leq \|y^\delta - A(x_k^\delta)\| \left[2\delta + 2\eta \|y^\delta - A(x_k^\delta)\| - \|y^\delta - A(x_k^\delta)\| \right] \\ &\leq \|y^\delta - A(x_k^\delta)\| \left[2\delta + 2\eta\delta - (1 - 2\eta) \|y^\delta - A(x_k^\delta)\| \right]. \end{aligned}$$

De (2.20), obtemos que o lado direito da estimativa acima é negativo. Logo,

$$\|x_{k+1}^\delta - x_*\| \leq \|x_k^\delta - x_*\|. \quad (2.21)$$

Para mostrar que iteração (2.15) está bem definida usaremos um argumento indutivo. Por hipótese, $x_0 \in B_\rho(x_*)$. Por (2.21), $x_1^\delta \in B_\rho(x_*)$. Suponha que $x_k^\delta \in B_\rho(x_*)$, satisfazendo (2.20). Por (2.21) $x_{k+1}^\delta \in B_\rho(x_*) \subset B_{2\rho}(x_0)$. ■

Segue de (2.20) que τ em (2.12) deve ser dado em função de η e que satisfaça

$$\tau > 2 \frac{1 + \eta}{1 - 2\eta} > 2. \quad (2.22)$$

O próximo passo é mostrar que o índice de parada dado pelo princípio de discrepância (2.12) com τ dado por (2.22) é finito.

Corolário 2.2.2. *Suponha que a Hipótese 2.1 seja satisfeita. Seja k_* dado pelo princípio da discrepância (2.12) com τ como em (2.22). Então,*

$$k_*(\tau\delta)^2 < \sum_{k=0}^{k_*-1} \|y^\delta - A(x_k^\delta)\|^2 \leq \frac{\tau}{(1-2\eta)\tau - 2(1+\eta)} \|x_0 - x_*\|^2. \quad (2.23)$$

Em particular, se $\delta = 0$, então

$$\sum_{k=0}^{\infty} \|y - A(x_k)\|^2 < \infty. \quad (2.24)$$

Demonstração : Segue da Proposição 2.2.1 que

$$\|x_{k+1}^\delta - x_*\|^2 - \|x_k^\delta - x_*\|^2 \leq \|y^\delta - A(x_k^\delta)\|^2 \left[\frac{2(1+\eta)}{\tau} - (1-2\eta) \right].$$

Somando esta desigualdade em k de 0 até $k_* - 1$, e impondo (2.12) temos (2.23). Obviamente que, se $\delta = 0$, então (2.23) pode ser aplicado a todo $k \in \mathbb{N}$. De (2.23) segue (2.24). ■

Exercício 2.14. *Complete os detalhes da demonstração do Corolário 2.2.2.*

Note que, no caso em que $y = y^\delta$ (dados sem ruídos) a equação (2.24) implica que, se a iteração de Landweber (2.15) converge, então o limite é uma solução de (1.1).

Teorema 2.2.3 (Convergência para dados exatos). *Suponha a Hipótese 2.1 satisfeita. Então a iteração de Landweber (2.15) aplicada a dados sem ruídos y , converge para uma solução de $A(x) = y$.*

Ideia da Demonstração : Defina $e_k := x_k - x^\dagger$. Ideia é mostrar que $\{e_k\}$ é uma sequência de Cauchy. Para isso, dado $j \geq k$, escolha l entre k e j tal que

$$\|y - A(x_l)\| \leq \|y - A(x_i)\|, \quad \forall \quad k \leq i \leq j.$$

Da desigualdade triangular temos que

$$\|e_k - e_j\| \leq \|e_k - e_l\| + \|e_l - e_j\|.$$

Ainda,

$$\begin{aligned} \|e_j - e_l\|^2 &= 2\langle e_l - e_j, e_l \rangle + \|e_j\|^2 - \|e_l\|^2 & (2.25) \\ \|e_j - e_k\|^2 &= 2\langle e_l - e_k, e_l \rangle + \|e_k\|^2 - \|e_l\|^2. \end{aligned}$$

Pela Proposição 2.2.1, $\{\|e_k\|\}$ é uma sequência monótona decrescente que converge a algum valor $\varepsilon \geq 0$. Portanto, o lado direito de ambas as igualdades em (2.25) convergem para $\varepsilon^2 - \varepsilon^2 = 0$ quando $k \rightarrow \infty$.

Usando (2.17) e a definição da iteração de Landweber (2.15)

mostra-se que

$$\begin{aligned} \langle e_l - e_j, e_l \rangle &\leq (1 + 3\eta) \sum_{i=l}^{j-1} \|y - A(x_i)\|^2 \\ \langle e_l - e_k, e_l \rangle &\leq (1 + 3\eta) \sum_{i=k}^{l-1} \|y - A(x_i)\|^2. \end{aligned}$$

De (2.24) segue que $\langle e_l - e_j, e_l \rangle$ e $\langle e_l - e_k, e_l \rangle$ convergem para 0 quando $k \rightarrow \infty$. Logo $\{e_k\}$, e por conseguinte $\{x_k\}$, são sequências de Cauchy. Pela observação feita logo após ao Corolário 2.2.2, o limite de x_k é uma solução de $A(x) = y$. ■

Exercício 2.15. *Mostre que, toda sequência monótona e limitada de números reais é convergente. Use isso para deduzir que $\{\|e_k\|\}$ na demonstração acima é convergente.*

Exercício 2.16. *Mostre que $\langle e_l - e_k, e_l \rangle \leq (1 + 3\eta) \sum_{i=k}^{l-1} \|y - A(x_i)\|^2$ para e_l e e_k definidas no Teorema 2.2.3. Sugestão: veja argumentação em [25, Teorema 2.4].*

Exercício 2.17. *Mostre que, se $\mathcal{N}(A'(x^\dagger)) \subset \mathcal{N}(A'(x))$ para todo $x \in B_\rho(x^\dagger)$, então $\{x_k\}$ dado pela iteração de Landweber (2.15) com dados sem ruídos converge para x^\dagger . Sugestão: veja [25, Teorema 2.4].*

Antes de seguirmos, gostaríamos de ressaltar que, para k fixo o iterado x_k^δ na iteração de Landweber (2.15) depende continuamente dos dados y^δ , uma vez que, x_k^δ é o resultado de uma combinação de operações contínuas.

Como no caso linear, se y^δ não pertence a imagem de $A(\cdot)$, então a iteração x_k^δ em (2.15) não converge. Por outro lado, se a iteração é parada de acordo com o princípio da discrepância, então

temos uma maneira estável de obter uma aproximação para a solução de $A(x) = y$. O próximo resultado implica que a iteração de Landweber (2.15) é um método de regularização, se escolhermos o critério de parada de acordo com o princípio da discrepância (2.12), com τ dado por (2.22).

Teorema 2.2.4. *Sejam as Hipóteses do Teorema 2.2.3 satisfeitas e $k_* = k_*(y^\delta, \delta)$ escolhidos de acordo com o princípio da discrepância (2.12), com τ dado por (2.22). Então a iteração de Landweber $x_{k_*}^\delta$ converge a uma solução de $A(x) = y$, quando $\delta \rightarrow 0$. Se $\mathcal{N}(A'(x^\dagger)) \subset \mathcal{N}(A'(x))$ para todo $x \in B_\rho(x^\dagger)$, então $x_{k_*}^\delta$ converge para x^\dagger quando $\delta \rightarrow 0$.*

Demonstração : Denote por x_* o limite da iteração de Landweber para dados sem ruídos (que existe pelo Teorema 2.2.3). Seja $\delta_n \rightarrow 0$ quando $n \rightarrow \infty$. Denote por $k_n = k_*(\delta_n, y^{\delta_n})$ o índice de parada determinado pelo princípio da discrepância aplicado ao par (δ_n, y^{δ_n}) .

Assuma que k é um ponto de acumulação de $\{k_n\}$. Sem perda de generalidade, suponha que $k_n = k$, para todo $n \in \mathbb{N}$. Portanto, da definição de k_n obtemos

$$\|y^{\delta_n} - A(x_k^{\delta_n})\| \leq \tau \delta_n. \quad (2.26)$$

Como k é fixo, x_k^δ depende continuamente de y^δ . Assim, de (2.26) obtemos

$$x_k^{\delta_n} \rightarrow x_k, \quad A(x_k^{\delta_n}) \rightarrow A(x_k) = y \quad \text{quando } n \rightarrow \infty.$$

Ou seja, x_k é uma solução de $A(x) = y$.

Falta analisar o caso em que $k_n \rightarrow \infty$ quando $n \rightarrow \infty$. Seja

$k_n > k$. Pela Proposição 2.2.1 temos

$$\|x_{k_n}^{\delta_n} - x_*\| \leq \|x_k^{\delta_n} - x_*\| \leq \|x_k^{\delta_n} - x_k\| + \|x_k - x_*\|. \quad (2.27)$$

Pelo Teorema 2.2.3, o último termo do lado direito da estimativa acima é menor que qualquer $\varepsilon > 0$ dado, para k suficientemente grande. Para um tal k , o termo $\|x_k^{\delta_n} - x_k\|$ também se torna pequeno quando $n \rightarrow \infty$, uma vez que, a iteração é estável para k fixo. Logo, $x_{k_n}^{\delta_n} \rightarrow x_*$ quando $n \rightarrow \infty$. ■

Obter taxas de convergência para a iteração de Landweber não linear requer muitas hipóteses. Assim, convidamos o leitor interessado a consultar [8, 25].

2.3 Método de Landweber Iterativamente Regularizado

Uma das características do método de Landweber apresentado nas seções anteriores é que esse converge muito devagar [8]. Nesta seção apresentaremos uma reformulação do método clássico de Landweber, que dentre outros atributos, tende a melhorar as taxas de convergência [25] e diminuir a complexidade das hipóteses.

As motivações de introdução do método que faremos aqui, diferem um pouco da feita em [25]. Começaremos relabrando alguns resultados interessantes do Cálculo.

Exercício 2.18 (Teorema do ponto fixo de Brauer). *Seja $f : [a, b] \rightarrow [a, b]$ uma função. Um **ponto fixo** para f é um ponto $c \in [a, b]$ tal que $f(c) = c$.*

Mostre que, se f for contínua, então existe um ponto fixo para f .

O nosso próximo passo é uma versão um pouco diferente do

Ponto fixo de Brauer.

Exercício 2.19. *Seja $f : [a, b] \rightarrow [a, b]$ uma função contínua e $x_0 \in [a, b]$ qualquer, mas fixo. Mostre que existe $d \in [a, b]$ tal que $f(\beta x_0 + (1 - \beta)d) = \beta x_0 + (1 - \beta)d$, para algum $\beta \in [0, 1]$.*

Verifique que $d = (1 - \beta)^{-1}c - \beta x_0$ onde c é o ponto fixo de f determinado no exercício acima.

Com base nos exercícios acima, vamos a construção do método de Landweber iterativamente regularizado. Começaremos supondo que A é um operador linear limitado.

A motivação para o método de Landweber clássico é, obter uma iteração de ponto fixo para $\Phi(x)$, dada por (2.4). Pelo Exercício 2.19, dado x_0

$$\Phi(\beta x_0 + (1 - \beta)x) = \beta x_0 + (1 - \beta)x + A^*(y - A(\beta x_0 + (1 - \beta)x)) \quad (2.28)$$

também possui ponto fixo. Uma iteração de ponto fixo para (2.28) é

$$x_{k+1} = x_k + (1 - \beta_k)A^*(y - Ax_k) + \beta_k(x_0 - x_k) + \beta_k A^*(y - Ax_0) \quad (2.29)$$

Como $A^*(y - Ax_0)$ independe de x_k , escolhendo $\beta_k \in [0, 1]$ de forma que $(1 - \beta_k)\|A\| < 1$, podemos considerar a iteração de Landweber iterativamente regularizada como

$$x_{k+1}^\delta = x_k^\delta + A^*(y^\delta - Ax_k^\delta) + \beta_k(x_0 - x_k^\delta). \quad (2.30)$$

A generalização para o caso não linear fica

$$x_{k+1}^\delta = x_k^\delta + A'(x_k^\delta)^*(y^\delta - Ax_k^\delta) + \beta_k(x_0 - x_k^\delta). \quad (2.31)$$

Uma dedução diferente para a iteração (2.31) será dada no Capítulo 3. Veja também [25].

Dar as condições necessárias para a escolha de β_k , de forma que a iteração (2.31) seja bem definida, é o próximo passo. As demonstrações são um tanto quanto complicadas nas escolhas dos parâmetros.

Proposição 2.3.1. *Seja $x_* \in B_{\bar{\rho}}(x_0)$ uma solução de $A(x) = y$. Sejam ainda, $\kappa \in (0, 1)$ e $0 < \beta_k \leq \beta_{max} < 1/2$. Defina*

$$c(\rho) := \rho \frac{1 - \beta_{max} + \sqrt{1 + \beta_{max}(2 - \beta_{max})\kappa^{-2}}}{2 - \beta_{max}}.$$

Assuma ainda que a Hipótese 2.1 é satisfeita na bola $B_{\rho+c(\rho)}(x_0)$. Se $x_k^\delta \in B_{c(\rho)}(x_)$ então, uma condição suficiente para $x_{k+1}^\delta \in B_{c(\rho)}(x_*)$ é que*

$$\|y^\delta - A(x_k^\delta)\| > 2(1 - \beta_{max}) \frac{1 + \eta}{E} \delta \quad (2.32)$$

onde $E := 1 - 2\eta - 2\beta_{max}(1 - \eta) - \kappa^2 > 0$.

Demonstração : Veja [25, Proposição 3.11]. ■

Note que, a proposição acima não garante monotonia da iteração (2.31), como foi mostrado para a iteração de Landweber. Veja Seções 2.1 e 2.2.

Por outro lado, a Proposição 2.3.1 implica que um critério de parada para a iteração (2.31) deve ser o princípio da discrepância (2.12) com $\tau > 2(1 - \beta_{max}) \frac{1 + \eta}{E}$. Nessas condições temos que:

Corolário 2.3.2. *Suponha que as hipóteses da Proposição 2.3.1 são satisfeitas e que a iteração (2.31) seja terminada de acordo com o*

princípio da discrepância (2.12) com $\tau > 2(1 - \beta_{max})\frac{1+\eta}{E}$. Então:

$$\begin{aligned} k_*(\tau\delta)^2 &\leq \sum_{k=0}^{k_*-1} \|y^\delta - A(x_k^\delta)\|^2 \\ &\leq \frac{\rho^2}{E - 2\tau^{-1}(1 - \beta_{max})(1 + \eta)} \left(1 + 2(1 + \kappa^{-2}) \sum_{k=0}^{k_*-1} \beta_k \right). \end{aligned} \quad (2.33)$$

Em particular, se $\delta = 0$ (dados livres de ruídos $y = y^\delta$), então

$$\sum_{k=0}^{\infty} \|y - A(x_k)\|^2 \leq E^{-1}\rho^2 \left(1 + 2(1 + \kappa^{-2}) \sum_{k=0}^{\infty} \beta_k \right). \quad (2.34)$$

Demonstração : Veja [25, Corolário 3.12]. ■

Exercício 2.20. Mostre que $E - 2\tau^{-1}(1 - \beta_{max})(1 + \eta) > 0$ e assim, a equação (2.33) está bem definida.

Exercício 2.21. Mostre que, se a série

$$\sum_{k=0}^{\infty} \beta_k \quad (2.35)$$

é convergente, então o critério de parada k_* para a iteração (2.31) é finito. Mostre ainda que, se a série (2.35) converge e a sequência de iterados $\{x_k\}$, dado por (2.31), com dados livres de ruídos, também converge, então o limite é uma solução do problema $A(x) = y$.

Exercício 2.22. Defina $e_k := x_k - x_*$ onde x_k é dado por (2.31). Mostre que

$$\|e_{k+1}\| \leq \|e_k\| + \beta_k \rho \sqrt{1 + \kappa^{-2}}.$$

Use indução para mostrar que

$$\|e_k\| \leq \|e_m\| \prod_{l=m}^{k-1} (1 - \beta_l) + \rho \sqrt{1 + \kappa^{-2}} \left(1 - \prod_{l=m}^{k-1} (1 - \beta_l) \right).$$

Sugestão: Use que $(1 - \prod_{l=m}^n (1 - \beta_l)) = \sum_{j=m}^n \beta_j \prod_{l=j+1}^n (1 - \beta_l)$.

Com a convergência da série (2.35), podemos provar que a sequência de iterados dada por (2.31) com dados sem ruídos converge.

Teorema 2.3.3 (Convergência para dados sem ruídos). *Suponha que as hipóteses da Proposição 2.3.1 são satisfeitas e seja $\{\beta_k\}$ satisfazendo (2.35). Então, $\{x_k\}$ dada pela iteração (2.31) converge. Pelo exercício acima, o limite é uma solução de $A(x) = y$.*

Demonstração : A demonstração difere da apresentada para o método de Landweber uma vez que não podemos usar que a sequência $\|e_k\| := \|x_k - x_*\|$ é monótona. Por outro lado, da Proposição 2.3.1 a sequência $\|e_k\|$ é limitada, portanto, possui uma subsequência $\|e_{k_j}\|$ convergente (prove!), digamos para algum $\varepsilon > 0$. Vamos mostrar que toda a sequência $\{\|e_k\|\}$ converge. E assim converge para $\varepsilon > 0$.

Seja $k_{j-1} < k < k_j$. Defina $\gamma_k := (1 - \prod_{l=k}^{k_j-1} (1 - \beta_l))$. Do exercício 2.22 temos que

$$\begin{aligned} -\gamma_k(\rho\sqrt{1 - \kappa^{-2}} - \|e_{k_j}\|) + \|e_{k_j}\| &\leq \|e_k\| \\ &\leq \|e_{k_{j-1}}\| + (\rho\sqrt{1 - \kappa^{-2}} - \|e_{k_{j-1}}\|)\gamma_k. \end{aligned}$$

Como $\{\|e_{k_j}\|\}$ converge, segue da desigualdade acima que $\{\|e_k\|\}$ converge para algum $\varepsilon > 0$. Agora, a demonstração segue os passos do Teorema 2.2.3. ■

O próximo Teorema garante que a iteração (2.31) é um método de regularização.

Teorema 2.3.4. *Suponha que as hipóteses do Teorema anterior são satisfeitas e $k_* = k_*(\delta, y^\delta)$ dado como no Corolário 2.3.2. Então $x_{k_*}^\delta$ dado pela iteração (2.31) do método de Landweber iterativamente regularizado converge para uma solução do problema $A(x) = y$ quando $\delta \rightarrow 0$.*

Demonstração : Similar a apresentada para o método de Landweber. ■

2.3.1 Taxas de Convergência

Como comentado anteriormente, taxas de convergência para o método de Landweber exigem hipóteses fortes [8, 25]. Por outro lado, para o método de Landweber iterativamente regularizado, tais hipóteses são mais naturais e implicam em taxas de convergência melhores que o método de Landweber clássico.

Hipótese 2.2. *Suponha que os itens (1) e (2) da hipótese (2.1) são satisfeitos. Suponha ainda que x^\dagger satisfaz a seguinte condição de fonte*

$$x^\dagger - x_0 \in \mathcal{R}(A'(x^\dagger))$$

e que $A'(x)$ é localmente Lipschitz, isto é,

$$\|A'(x) - A'(\tilde{x})\| \leq L\|x - \tilde{x}\| \quad x, \tilde{x} \in B_{2\rho}(x_0).$$

Com a hipótese acima temos:

Teorema 2.3.5. *Suponha que a Hipótese 2.2 é satisfeita. Sejam*

a, b, c, d parâmetros satisfazendo

$$\begin{aligned} (1+a)w^2 + b + c + d + 2\beta_0 - 2 &\leq 0, \\ \|x_0 - x^\dagger\|^2 &\leq 2\beta_0\sqrt{qdL^{-1}} \leq \rho^2, \\ \sqrt{qdL^{-1}} &< p \leq \beta_0^{-1} \\ \beta_k(1 - \beta_k(p - \sqrt{dd^{-1}L})) &\leq \beta_{k-1} \end{aligned}$$

com $\{\beta_k\}$ uma sequência monótona decrescente. Então, se a iteração é parada de acordo com o princípio da discrepância (2.12) temos que

$$\|x_{k_*}^\delta - x^\dagger\| = \mathcal{O}(\sqrt{\delta}). \quad (2.36)$$

No caso de dados livres de ruídos ($\delta = 0$), temos que

$$\|x_k - x^\dagger\| = \mathcal{O}(\sqrt{\beta_k}). \quad (2.37)$$

Demonstração : Veja [25, Teorema 3.17]. ■

2.4 Método do Gradiente

Nesta seção, vamos nos ater ao método conhecido como **método do gradiente** ou ainda como steepest descent (passo de máxima descida), 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 de um 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 ou steepest descent também é conhecido como método de máxima descida.

Por questões didáticas, trabalharemos com a seguinte hipótese:

Hipótese 2.3. *O operador A é linear, limitado e possui a imagem fechada.*

Tal hipótese também é utilizada no Apêndice B para construir o conceito de inversa generalizada de forma mais simples. Queremos deixar claro que tais hipóteses podem ser suprimidas. No entanto, outras considerações sobre o operador A devem ser feitas. Veja [13].

Exercício 2.23. *Dê um exemplo de um operador linear limitado cuja imagem não é fechada.*

Vamos a alguns resultados preliminares.

Seja $J : \mathcal{H} \rightarrow \mathbb{R}$ um funcional não-negativo. Denotemos por $x^* \in \mathcal{H}$ um ponto tal que

$$J(x^*) = \inf\{J(x) : x \in \mathcal{H}\}. \quad (2.38)$$

Ainda, supomos que J é Fréchet diferenciável em cada ponto de \mathcal{H} .

Exercício 2.24. *Prove que, se A é um operador linear e limitado, então $J(x) = \frac{1}{2}\|Ax - y\|^2$ é Fréchet diferenciável e*

$$\nabla J(x) = A^*(Ax - y).$$

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

$$DJ(x_0, z) = \langle z, \nabla J(x_0) \rangle,$$

aponte na direção de maior descida.

Pela desigualdade de Cauchy-Schwarz,

$$-\|z\|\|\nabla J(x_0)\| \leq \langle z, -\nabla J(x_0) \rangle$$

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 $z = -\nabla J(x_0)$. Tomando um tamanho de passo $\alpha_0 > 0$ na direção $-\nabla J(x_0)$, obtemos a 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 $-\nabla J(x_0)$. 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 do gradiente é dada por

$$x_{k+1} = x_k - \alpha_k \nabla J(x_k),$$

onde

$$\frac{d}{d\alpha} J(x_k - \alpha \nabla J(x_k))|_{\alpha=\alpha_k} = 0.$$

Suponha que o operador A satisfaz a Hipótese 2.3. Pelo Teorema B.1.4, a solução de quadrados mínimos de $Ax = y$ é justamente o mínimo do funcional $J(x) = \frac{1}{2}\|Ax - y\|^2$. Assim, somos levados a aplicar o método do gradiente para o funcional J .

Exercício 2.25. Defina $r := \nabla J(x) = A^*(Ax - y)$ (o qual está bem definido pelo exercício 2.24). Mostre que o valor ótimo de α é dada

por

$$\alpha = \frac{\|r\|^2}{\|Ar\|^2}.$$

Assim, a sequência $\{x_k\} \subset \mathcal{H}_1$ pelo método do gradiente é dada por

$$\begin{aligned} x_{k+1} &= x_k - \alpha_k r_k \\ r_k &= A^*(Ax_k - y) \\ \alpha_k &= \frac{\|r_k\|^2}{\|Ar_k\|^2}. \end{aligned} \tag{2.39}$$

Observação: Observe que, se $r_k = 0$, então x_k é uma solução de quadrados mínimos e o método para em x_k . Se tivermos um critério para garantir que x_k é a solução de menor norma, entre todas as soluções de quadrados mínimos, obteremos que $x_k = x^\dagger$.

Exercício 2.26. *Mostre que se $Ar_k = 0$, então $r_k = 0$. Assim, a sequência gerada pelo algoritmo (2.39) está bem definida.*

Lema 2.4.1. *Seja A satisfazendo a Hipótese (2.3). Então,*

$$\lim_{k \rightarrow \infty} r_k = 0.$$

Demonstração : De

$$J(x_{k+1}) = \frac{1}{2} \|Ax_k - \alpha_k Ar_k - y\|^2 = J(x_k) - \frac{1}{2} \frac{\|r_k\|^2}{\|Ar_k\|^2}.$$

Repetindo a igualdade acima recursivamente, obtemos

$$J(x_{k+1}) = J(x_0) - \frac{1}{2} \sum_{j=0}^k \frac{\|r_j\|^2}{\|Ar_j\|^2}.$$

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 $k \rightarrow \infty$. ■

Exercício 2.27. *Mostre que o Lema 2.4.1 segue sem a hipótese de A possuir imagem fechada.*

Mostraremos que a iteração (2.39), para dados sem ruídos, converge para uma solução de norma mínima do problema $Ax = y$.

Teorema 2.4.2. *Seja A satisfazendo a Hipótese 2.3. Então, a sequência gerada pelo método do gradiente converge para uma solução de quadrados mínimos de $Ax = y$ para qualquer $x_0 \in \mathcal{H}_1$. A sequência converge a x^\dagger se, e somente se, $x_0 \in \mathcal{R}(A^*)$.*

Demonstração : Escrevendo a iteração (2.39) de forma recursiva, temos

$$x_{k+1} = x_0 - \sum_{j=0}^k \alpha_j r_j. \quad (2.40)$$

Mostraremos que $\{x_n\}$ é uma sequência de Cauchy. Seja $m > n$, então

$$x_m - x_n = - \sum_{j=n}^k \alpha_j r_j \in \mathcal{R}(A^*) = \mathcal{N}(A)^\perp. \quad (2.41)$$

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,

$$\begin{aligned} & \langle A^*A(x_m - x_n), x_m - x_n \rangle \\ & \leq (\|A^*Ax_m - A^*y\| + \|A^*Ax_n - A^*y\|) \|x_m - x_n\| \\ & \leq \eta^{-1} (\|A^*Ax_m - A^*y\| + \|A^*Ax_n - A^*y\|) \|A(x_m - x_n)\|. \end{aligned}$$

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 $\mathcal{R}(A) = \mathcal{N}(A^*)^\perp$ é fechada, segue do Teorema do Gráfico Fechado [28], que A^* possui inversa limitada quando restrito a $\mathcal{R}(A) = \mathcal{N}(A^*)^\perp$.

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 concluímos que u é uma solução de quadrados mínimos.

Para finalizar, lembremos que qualquer solução de quadrados mínimos é da forma $x^\dagger \oplus \mathcal{N}(A)$ (Teorema B.1.4). Como $\mathcal{R}(A^*) = \mathcal{N}(A)^\perp$, segue que x^\dagger é a única solução de quadrados mínimos na $\mathcal{R}(A^*)$. Se $x_0 \in \mathcal{R}(A^*)$, então, de (2.40), segue que $x_k \in \mathcal{R}(A^*)$ para todo k . Como $\mathcal{R}(A^*)$ é fechada, $u = x^\dagger$.

Caso $x_0 \notin \mathcal{R}(A^*)$, então $x_0 = x'_0 + P_{\mathcal{N}(A)}x_0$, onde $x'_0 \in \mathcal{N}(A)^\perp = \mathcal{R}(A^*)$ e $P_{\mathcal{N}(A)}x_0 \neq 0$. Como $A^*AP_{\mathcal{N}(A)}x_0 = 0$, de (2.40), 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. ■

Exercício 2.28. *Complete os detalhes da demonstração do Teorema 2.4.2.*

Exercício 2.29. *Suponha que existem constantes $C, M, m > 0$ tais que $m\|x\|^2 \leq \langle A^*Ax, x \rangle \leq M\|x\|^2$, $\forall x \in \mathcal{R}(A^*)$ e que A satisfaz a Hipótese 2.3. Mostre que o método do gradiente possui um limitante de erro dado por*

$$\|x^\dagger + P_{\mathcal{N}(A)}x_0 - x_n\| \leq C \left(\frac{M-m}{M+m} \right)^2.$$

No caso de dados com ruídos, consideraremos a iteração (2.39) com x_k^δ no lugar de x_k .

Mostraremos que a iteração (2.39) é monótona, enquanto o princípio da discrepância (2.12) não é atingido.

Proposição 2.4.3 (Monotonia). *Seja x^\dagger uma solução de norma mínima do problema $Ax = y$, com y^δ satisfazendo (1.2). Enquanto (2.20) for satisfeito, temos que*

$$\|x_{k+1}^\delta - x^\dagger\| \leq \|x_k^\delta - x^\dagger\|.$$

Ainda, se $k_ = k_*(\delta, y^\delta)$ for escolhido de acordo com o princípio da discrepância (2.12), então a iteração é terminada depois de uma quantidade finita de passos.*

Demonstração : Da definição da iteração (2.39) temos que

$$\begin{aligned}
\|x_{k+1}^\delta - x^\dagger\|^2 - \|x_k^\delta - x^\dagger\|^2 &= 2\langle x_{k+1}^\delta - x_k^\delta, x_k^\delta - x^\dagger \rangle + \|x_{k+1}^\delta - x_k^\delta\|^2 \\
&= 2\alpha_k \langle A^*(y^\delta - Ax_k^\delta), x_k^\delta - x^\dagger \rangle + \alpha_k^2 \|r_k\|^2 \\
&= 2\alpha_k \langle y^\delta - Ax_k^\delta, Ax_k^\delta - y^\delta \rangle + 2\alpha_k \langle y^\delta - Ax_k^\delta, y^\delta - Ax^\dagger \rangle + \alpha_k^2 \|r_k\|^2 \\
&\leq -2\alpha_k \|y^\delta - Ax_k^\delta\|^2 + 2\delta\alpha_k \|y^\delta - Ax_k^\delta\| + \alpha_k^2 \|r_k\|^2 \\
&= \alpha_k \|y^\delta - Ax_k^\delta\| \left[2\delta - \|y^\delta - Ax_k^\delta\| \right] - \alpha_k (\|y^\delta - Ax_k^\delta\|^2 - \alpha_k \|r_k\|^2).
\end{aligned}$$

Seja $k < k_*$, então pelo princípio da discrepância, com $\tau < 2$, temos que

$$\|x_{k+1}^\delta - x^\dagger\|^2 \leq \|x_k^\delta - x^\dagger\|^2 - \alpha_k (\|y^\delta - Ax_k^\delta\|^2 - \alpha_k \|r_k\|^2).$$

Como

$$\alpha_k \|r_k\|^2 = \frac{(\langle Ar_k, y^\delta - Ax_k^\delta \rangle)^2}{\|Ar_k\|^2} \leq \|y^\delta - Ax_k^\delta\|,$$

o resultado segue. \blacksquare

Exercício 2.30. *Mostre que, k_* definido pelo princípio da discrepância (2.12) é finito para a iteração do gradiente.*

O próximo resultado mostra que a iteração do gradiente é um método de regularização.

Teorema 2.4.4. *Seja $k_* = k(\delta, y^\delta)$ definido pelo princípio da discrepância (2.12). Então, $x_{k_*}^\delta$ converge para uma solução de $Ax = y$, quando $\delta \rightarrow 0$.*

Demonstração : Note que, para $k < k_*$ fixo, o iterado x_k^δ dada pela iteração (2.39) depende continuamente dos dados y^δ , uma vez que, x_k^δ é a composição de operadores contínuos. Os passos da demonstração,

agora, são similares as feitas anteriormente e são deixadas ao leitor.

■

Para o caso em que $A(\cdot)$ é não linear, as demonstrações seguem os mesmos passos acima. Basta utilizar hipóteses similares ao método de Landweber não linear estudado na Seção 2.1. Para detalhes veja [25].

2.5 Método de Landweber e Estratégias de Kaczmarz

Dedicaremos esta seção a um método de regularização para resolver problemas inversos que aparecem na forma de sistemas de equações

$$A_i(x) = y_i \quad i \in \{1, \dots, M\}. \quad (2.42)$$

Uma maneira de resolver o problema (2.42) é considerar 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 anteriormente.

A principal idéia de métodos tipo Kaczmarz é resolver o sistema (2.42), de maneira cíclica. A título de motivação, apresentaremos na próxima subseção o método ART (algebraic reconstruction techniques) como forma de motivar métodos tipo Kaczmarz.

2.5.1 Um Pouco Sobre o Método ART

Uma das técnicas mais usadas para resolver numericamente o problema associado a diagnósticos médicos para tomografia com-

putadorizada, até bem pouco tempo atrás, chama-se método ART (algebraic reconstruction techniques). Hoje em dia, novos métodos ganharam terreno², uma vez que o método ART é mais caro do ponto de vista computacional [38, 37].

Por outro lado, o método ART representa a ideia de métodos tipo Kaczmarz de 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 (2.42). Desta maneira, a iteração ART usa conceitos básicos de Álgebra Linear.

Iteração: Dado \vec{x}_0 um ponto inicial, projeta-se \vec{x}_0 ortogonalmente sobre o hiperplano determinado pela primeira equação do sistema (2.42), 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 , projeta-se ortogonalmente sobre a segunda linha do sistema (2.42), obtendo-se \vec{x}_2 e assim sucessivamente até M , obtendo, \vec{x}_M . De maneira cíclica, projeta-se $\vec{x}_{mod(M)}$ sobre a equação de índice $mod(M)$ do sistema (2.42). A figura mostra a iteração ART para um sistema quadrado de ordem 2.

A iteração é dada por

$$\vec{x}_n = \vec{x}_{n-1} - \left(\frac{\langle A_i, \vec{x}_{n-1} \rangle - y_i}{\|A_i\|^2} \right) (A_i)^T \quad (2.43)$$

Exercício 2.31. 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 (2.43).

Exercício 2.32. Mostre que $\langle A_i, \vec{x}_n \rangle = y_i$. Interprete esse resultado geometricamente.

²Hoje em dia, métodos iterativos estatísticos são muito usados para controlar o problema causado por ruídos nos dados.

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

$$\omega_{k_*+j} = 0 \quad \text{para } j = \{0, \dots, M-1\}. \quad (2.45)$$

Note que, um tal critério de parada dado pela equação (2.45) significa que

$$x_{k_*}^\delta = x_{k_*+1}^\delta = \dots = x_{k_*+M-1}^\delta, \quad (2.46)$$

isto é, k_* é escolhido como o índice que faz com que x_k^δ permaneça igual em um ciclo.

Convergência para a iteração de Landweber-Kaczmarz segue similarmente ao feito para o método de Landweber. As hipóteses são semelhantes as feitas para o método de Landweber para cada $A_{[k]}$.

Exercício 2.34. *Mostre que o método de Landweber-Kaczmarz (2.44) com o critério de parada (2.45) é um método de regularização. Sugestão: consulte [25].*

Exercício 2.35. *Use a iteração de Landweber-Kaczmarz (2.44) para encontrar uma solução do problema de tomografia proposto na Figura 1.5. Compare com o método ART.*

2.6 Aplicação: Tomografia Computadorizada

Como aplicação, apresentaremos o resultado obtido para o problema de *tomografia computadorizada foto-acústica* com uma quantidade limitada de dados [37, 38].

Seja $L^2(D)$ o espaço de Hilbert de todas as funções quadrado integráveis no disco unitário $D \subset \mathbb{R}^2$ e $\mathcal{H}_2 := \{y : [0, 2] \rightarrow \mathbb{R} : \|y\|^2 := \int_0^2 y(t)t dt < \infty\}$.

Exercício 2.36. *Mostre que \mathcal{H}_2 é um espaço de Hilbert.*

Consideraremos o sistema

$$A_j x = y_j, \quad j = 0, \dots, N-1, \quad (2.47)$$

onde cada $A_j := L^2(D) \rightarrow \mathcal{H}_2$, é dada por

$$(A_j x) := \frac{1}{\pi} \int_{S^1} x(\xi_j + t\theta) d\Omega(\theta), \quad t \in [0, 2], \quad (2.48)$$

A equação (2.48) corresponde a uma versão circular da Transformada de Radon que foi introduzida na Seção 1.2.1. Resolver (2.48) é um passo crucial em tomografia computadorizada foto-acústica. Para esta aplicação em especial, 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. Ainda, nesta aplicação consideraremos somente o caso especial em que cada $\xi_j = (\sin(\frac{\pi j}{N-1}), \cos(\frac{\pi j}{N-1}))$, 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 2.6). Portanto, temos poucos dados. É provado, por exemplo em [38, 37], que certos detalhes de x fora da região de detecção, não podem ser

reconstruídos. Tais resultados são conhecidos como invisibilidade.

Exercício 2.37. Prove que cada operador A_j é linear, limitado e satisfaz $\|A_j\| \leq 1$. Mostre ainda que o operador adjunto é dado por $(A_j^*y)(\xi) = y \frac{(|\xi_j - \xi|)}{\sqrt{\pi}}$.

Temos duas possibilidades. A primeira é considerar o operador

$$\mathbb{A} = (A_0, \dots, A_{N-1}) : L^2(D)^N \longrightarrow \mathcal{H}_2^N,$$

$$\mathbb{A}x := \begin{pmatrix} A_0 \\ \vdots \\ A_{N-1} \end{pmatrix} x = \begin{pmatrix} y_0 \\ \vdots \\ y_{N-1} \end{pmatrix} = y. \quad (2.49)$$

Com essa configuração podemos aplicar a iteração de Landweber linear (2.6), a iteração de Landweber iterativamente regularizado (2.30) ou a iteração do gradiente (2.39).

A segunda possibilidade é aplicar o método de Landweber-Kaczmarz para o sistema de equações lineares (2.48). Para nosso teste numérico, optamos pela segunda.

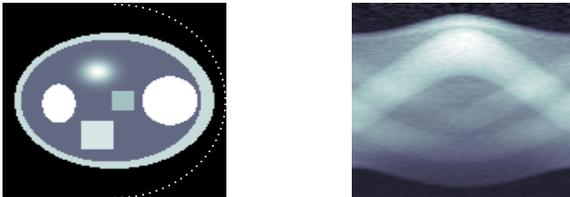


Figura 2.2: Do lado direito a solução e do esquerdo os dados com 4% de ruídos.

A solução x^\dagger é mostrada no lado esquerdo da Figura 2.6. Do lado direito da Figura 2.6, vê-se 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.

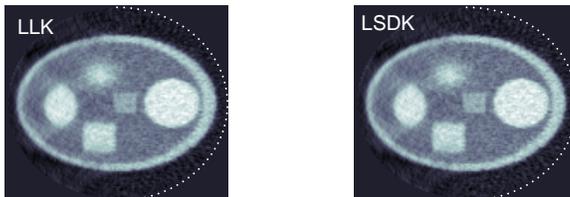


Figura 2.3: Do lado direito a solução aproximada pela regularização de Landweber e do esquerdo por steepest-descent, com estratégias tipo Kaczmarz.

A Figura 2.6 mostra a solução regularizada $x_{k^*}^{\delta}$ para ambos, os métodos Landweber-Kaczmarz e método do gradiente com uma estratégia tipo Kaczmarz.

2.7 Breve Apanhado Histórico do Capítulo

Em 1809, Carl Friedrich Gauss (1777-1855) publicou um artigo no *Werke*, 4, 1-93, demonstrando que a melhor maneira de determinar um parâmetro desconhecido de uma equação é minimizando a soma dos quadrados dos resíduos, mais tarde chamado de Mínimos Quadrados por Adrien-Marie Legendre (1752-1833).

A **desigualdade de Cauchy-Schwarz** para o caso de somas foi publicada por Augustin Cauchy (1821), enquanto a correspondente desigualdade para integrais foi primeiro estabelecida por Viktor Yakovlevich Bunyakovsky (1859) e redescoberta por Hermann Amandus Schwarz (1888).

O método ART foi usado por Hounsfield³ para a reconstrução da primeira imagem por tomografia computadorizada. Na verdade, a iteração ART é uma reinvenção do método de Kaczmarz publicado em 1937.

- Sir Isaac Newton (1643 -1727) - cientista inglês, mais reconhecido como físico e matemático, embora tenha sido também astrônomo, alquimista e filósofo natural. Newton é o autor da obra *Philosophiae Naturalis Principia Mathematica*, publicada em 1687, que descreve a lei da gravitação universal e as Leis de Newton - as três leis dos corpos em movimento que assentaram-se como fundamento da mecânica clássica.
- Augustin Louis Cauchy (1789 - 1857) - matemático Francês. O primeiro avanço na matemática moderna por ele produzido foi a introdução do rigor na análise matemática. Criou a noção moderna de continuidade para as funções de variável real ou complexa. Mostrou a importância da convergência das séries inteiras. Fez definições precisas das noções de limite e integral definida, transformando-as em notável instrumento para o estudo das funções complexas. Sua abordagem da teoria das equações diferenciais foi inteiramente nova, demonstrando a existência e unicidade das soluções, quando definidas as condições de contorno. Foi um dos fundadores da teoria de grupos finitos. Também exerceu grande influência sobre a física.
- Luitzen Egbertus Jan Brouwer (1881 – 1966) - matemático holandês graduado na Universidade de Amsterdam. Trabalhou em topologia, teoria dos conjuntos, medida matemática e análise complexa. O teorema do ponto fixo de Brouwer foi batizado em sua homenagem.

³Relembra a Seção 1.4

- Maurice René Fréchet(1878 – 1973) - matemático francês.
- Stefan Kaczmarz (1895 – 1940) - matemático Polonês. Seu método é a base de muitos métodos usados em tomografia computadorizada.
- Karl Hermann Amandus Schwarz (1843 –1921) - matemático Alemão, conhecido por seu trabalho em análise complexa.

Capítulo 3

Métodos de Regularização Iterativos Tipo Newton

De maneira similar aos resultados conhecidos para resolver problemas diretos (bem postos), métodos do tipo gradiente exigem muitas iterações para convergir. Por outro lado, métodos tipo Newton possuem taxas de convergência melhores. No entanto, a realização numérica de métodos como a iteração de Landweber, é mais fácil que os métodos tipo Newton, em termos de hipóteses.

Neste capítulo faremos a hipótese de que $A(\cdot)$ é Fréchet diferenciável e que a derivada de Fréchet $A'(\cdot)$ é limitada.

Um primeiro candidato a resolver a equação $A(x) = y$ de forma iterativa é o método de Newton

$$x_{k+1}^\delta = x_k^\delta + (A'(x_k^\delta))^{-1}(y^\delta - A(x_k^\delta)), \quad (3.1)$$

partindo de um chute inicial x_0 .

Exercício 3.1. *Suponha que $A(\cdot)$ é Fréchet diferenciável até segunda ordem. Defina $F(x) := A(x) - y$. Mostre que $F(x)$ é diferenciável. Use o Teorema de Taylor para mostrar que, a iteração de Newton (3.1), é dada pela solução da expansão de Taylor de primeira ordem de $F(\cdot)$ em torno de x_k^δ .*

A iteração pelo Método de Newton possui muitas limitações para problemas mal postos. Uma condição necessária para que a iteração (3.1) esteja bem definida, é que $A'(\cdot)$ seja invertível, para todo iterado x_k^δ . Supondo que isso seja verdade, a sua inversa, em geral, é não limitada. Relembre os exercícios da Seção 2.2.

Portanto, para problemas mal postos, a iteração (3.1) é inapropriada, pois, em cada passo estamos resolvendo um problema mal posto.

A ideia principal de métodos tipo Newton é linearizar o problema $A(x) = y$, em torno de uma solução aproximada x_k^δ e então resolver o problema linearizado

$$A'(x_k^\delta)(z - x_k^\delta) = y^\delta - A(x_k^\delta), \quad (3.2)$$

para z .

Note que, se o problema $A(x) = y$ é mal posto, então resolver o problema linearizado (3.2), também é mal posto (Por quê?). Relembre o exercício feito na Seção 2.2. Assim, algum tipo de regularização deve ser aplicada.

3.1 Método de Gauss-Newton

O método de Gauss-Newton pode ser visto como uma modificação do método de Newton. Este consiste em definir

$$r_k^\delta = y^\delta - A(x_k^\delta),$$

e aplicar o método de mínimos quadrados para o problema linearizado

$$A'(x_k^\delta)(z - x_k^\delta) = r_k^\delta.$$

Isto é, minimizar o funcional

$$\|A'(x_k^\delta)(z - x_k^\delta) - r_k^\delta\|^2, \quad (3.3)$$

para z .

Sendo $x_{k+1}^\delta = \tilde{z}$ o mínimo do funcional, obtemos a iteração

$$x_{k+1}^\delta = x_k^\delta + (A'(x_k^\delta)^* A'(x_k^\delta))^{-1} r_k^\delta. \quad (3.4)$$

Exercício 3.2. *Mostre que o funcional 3.3 sempre possui um mínimo.*

Exercício 3.3. *Mostre que a condição de otimalidade para o funcional (3.3) é dada pela equação normal*

$$(A'(x_k^\delta)^* A'(x_k^\delta))(\tilde{z} - x_k^\delta) = r_k^\delta.$$

Exercício 3.4. *Mostre que a condição de otimalidade do exercício anterior é também suficiente.*

Exercício 3.5. *Mostre que, se $A'(x)$ é um operador linear, limitado e compacto, então $A'(x)^* A'(x)$ também é compacto.*

Conforme o exercício acima, a iteração (3.4), mesmo se bem

definida, isto é, mesmo que $A'(x)^*A'(x)$ seja invertível em todos os pontos da iteração, esta é inapropriada para problemas mal postos. Assim, algum tipo de regularização deve ser aplicada.

3.2 Método de Levenberg-Marquardt

A ideia original do *Método de Levenberg-Marquardt* é minimizar o resíduo $\|y^\delta - A(x)\|$, sujeito a **região de confiança** $\|x - x_k^\delta\| \leq h_k$. Pelo método de multiplicadores de Lagrange, isto é equivalente a minimizar o funcional

$$\|y^\delta - A(x_k^\delta) - A'(x_k^\delta)z\|^2 + \alpha_k \|z\|^2. \quad (3.5)$$

para $z = x - x_k^\delta$, onde α_k é o correspondente multiplicador de Lagrange. Então, o processo é iterado com

$$x_{k+1}^\delta = x_k^\delta + z_k^\delta$$

e, possivelmente, um ajuste do raio da região de confiança h_k .

Exercício 3.6. *Mostre que, z_k^δ minimizante de (3.5) está bem definido, isto é, sempre existe.*

Uma das maneiras de atualizar o raio da região de confiança, é fazer uma comparação de concordância entre o resíduo atual e o resíduo previsto, isto é,

$$\|y^\delta - A(x_{k+1}^\delta)\| \quad \text{versus} \quad \|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(z - x_k^\delta)\|.$$

O parâmetro h_k é aumentado (boa concordância) ou diminuído (se a concordância for ruim), por algum fator constante, por exemplo, 2.

Exercício 3.7. *Mostre que, a condição de otimalidade de primeira*

ordem do funcional (3.5) é dada por

$$x_{k+1}^\delta = x_k^\delta + (A'(x_k^\delta)^* A'(x_k^\delta) + \alpha_k I)^{-1} A'(x_k^\delta)^* (y^\delta - A(x_k^\delta)).$$

Uma maneira de atualizar a região de confiança que é bem conhecida em otimização é denominada de *dog-leg*.

Outra justificativa para o método de Levenberg-Marquardt é que estamos aplicando regularização de Tikhonov para o problema linear (3.2). Veja o Capítulo 4.

Exercício 3.8. Seja $x_{k+1}^\delta(\alpha)$ o minimizante de (3.5) para $\alpha_k := \alpha$. Mostre que

$$\begin{aligned} & \|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x_{k+1}^\delta(\alpha) - x_k^\delta)\| \\ &= \alpha \|(A'(x_k^\delta)^* A'(x_k^\delta) + \alpha I)^{-1} (y^\delta - A(x_k^\delta))\|. \end{aligned}$$

Note que, a direção de atualização z_k é a direção de descida do resíduo $\|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(z - x_k^\delta)\|$. Assim, escolhas de α_k de forma que o método convirja devem ser feitas.

Seguiremos os passos de [25] e referências. A proposta é: determinar α_k de acordo com o princípio da discrepância *a posteriori*, da seguinte forma: seja $\alpha := \alpha_k$, tal que

$$p_k^\delta(\alpha) := \|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x_{k+1}^\delta(\alpha) - x_k^\delta)\| = q \|y^\delta - A(x_k^\delta)\|, \quad (3.6)$$

para algum $q \in (0, 1)$.

Note que, enquanto $y^\delta - A(x_k^\delta) \neq 0$, segue de (3.6) e do Exer-

cício 3.8 que $p_k^\delta(\alpha)$ é contínuo e decrescente. Ainda,

$$\begin{aligned}\lim_{\alpha \rightarrow \infty} p_k^\delta(\alpha) &= \|y^\delta - A(x_k^\delta)\| \\ \lim_{\alpha \rightarrow 0} p_k^\delta(\alpha) &= \|(A'(x_k^\delta))^\dagger(y^\delta - A(x_k^\delta))\| \leq \|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x - x_k^\delta)\|,\end{aligned}$$

para qualquer $x \in \mathcal{D}(A)$.

Seja $x^\dagger \in B_\rho(x_0)$ denotando a solução de x_0 -norma mínima e $\gamma > 1$. Se

$$\|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x^\dagger - x_k^\delta)\| \leq \frac{q}{\gamma} \|y^\delta - A(x_k^\delta)\|, \quad (3.7)$$

então (3.6) possui uma única solução α_k (prove! Sugestão: Use o Teorema do Valor Intermediário [33].).

3.2.1 Convergência

De forma similar ao feito para o método de Landweber não-linear na Seção 2.2, nessa seção, vamos provar que o método de Levenberg-Marquardt é um método de regularização.

Primeiro, estabeleceremos monotonia da iteração.

Proposição 3.2.1. *Seja $0 < q < 1 < \gamma$. Suponha que (1.1) possua uma solução e que (3.7) seja satisfeita, de forma que, exista α_k determinado por (3.6). Então:*

$$\begin{aligned}\|x_k^\delta - x^\dagger\|^2 - \|x_{k+1}^\delta - x^\dagger\|^2 &\geq \|x_{k+1}^\delta - x_k^\delta\|^2, \\ \|x_k^\delta - x^\dagger\|^2 - \|x_{k+1}^\delta - x^\dagger\|^2 &\geq 2 \frac{\gamma - 1}{\gamma \alpha_k} \|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta)\|^2 \\ &\geq 2 \frac{(\gamma - 1)(1 - q)q}{\gamma \|A'(x_k^\delta)\|^2} \|y^\delta - A(x_k^\delta)\|^2\end{aligned}$$

Demonstração : Veja [25, Proposição 4.1]. ■

Para mostrar convergência da iteração para dados sem ruídos, temos que supor que vale a seguinte **condição local de cone tangente**

$$\|A(x) - A(\tilde{x}) - A'(x)(x - \tilde{x})\| \leq c \|x - \tilde{x}\| \|A(x) - A(\tilde{x})\|, \quad (3.8)$$

$$x, \tilde{x} \in B_{2\rho}(x_0).$$

Teorema 3.2.2 (Convergência para dados exatos). *Suponha que $A(x) = y$ possui uma solução, que $A'(\cdot)$ é uniformemente limitada na bola $B_\rho(x^\dagger)$ e que $\|x_0 - x^\dagger\| \leq q/c$. Suponha, ainda, α_k determinado por (3.6) e que (3.8) seja satisfeita. Então, o método de Levenberg-Marquardt converge para uma solução de $A(x) = y$, quando $k \rightarrow \infty$.*

Demonstração : Os passos da demonstração são análogos aos feitos para o método de Landweber não-linear. Para detalhes, veja [25, Teorema 4.2]. ■

Teorema 3.2.3. *Suponha que as hipóteses do Teorema 3.2.2 são satisfeitas. Seja $k_* = k_*(\delta, y^\delta)$ dado pelo princípio da discrepância (2.12), com $\tau > 1/q$. Então, a iteração do método de Levenberg-Marquardt, para dados com ruídos, é terminada em um número finito de passos k_* e*

$$k_*(\delta, y^\delta) = \mathcal{O}(1 + |\ln \delta|). \quad (3.9)$$

Ainda, o iterado $x_{k_}^\delta$, dado pelo método de Levenberg-Marquardt, converge para uma solução de $A(x) = y^\delta$, quando $k \rightarrow \infty$.*

Demonstração : Veja [25, Theo 4.3]. ■

Taxas de convergência para o método de Levenberg-Marquardt são complicadas em termos de hipóteses. Assim, deixamos ao leitor interessado consultar [25].

3.3 Método de Gauss-Newton Iterativamente Regularizado

O método de Gauss-Newton iterativamente regularizado consiste em iterar

$$x_{k+1}^\delta = x_k^\delta + (A'(x_k^\delta)^* A'(x_k^\delta) + \alpha_k I)^{-1} (y^\delta - A(x_k^\delta)) + \alpha_k (x_0 - x_k^\delta), \quad (3.10)$$

com x_0 , uma aproximação *a priori* para a solução do problema, α_k uma sequência tendendo a zero e y^δ são dados corrompidos por ruídos satisfazendo (1.2).

O método de Gauss-Newton iterativamente regularizado é similar ao método de Levenberg-Marquardt. De fato, note que x_{k+1}^δ é um minimizante do funcional

$$\|y^\delta - A(x_k^\delta) - A'(x_k^\delta)(x - x_k^\delta)\|^2 + \alpha_k \|x - x_0\|^2. \quad (3.11)$$

Análise de convergência, para o método de Gauss-Newton iterativamente regularizado, requer hipóteses similares às feitas para o método de Landweber não-linear na Seção 2.2. Devido a similaridade dos argumentos desenvolvidos anteriormente, deixamos a análise de convergência para o método de Gauss-Newton iterativamente regularizado como exercício. Para detalhes, consulte [25].

3.4 Breve Apanhado Histórico do Capítulo

Em análise numérica, o método de Newton (também conhecido como método de Newton-Raphson), é um dos métodos mais conhecidos para encontrar zeros (ou raízes) de funções a variáveis reais, de maneira aproximada. Sob hipóteses adequadas na função, o método

de Newton tem convergência local quadrática, o que não pode ser atingido por métodos do tipo gradiente. Por outro lado, se a iteração é iniciada muito longe da raiz exata, a iteração diverge facilmente.

O método de Newton foi proposto por Isaac Newton em *De analysi per aequationes numero terminorum infinitas* de 1669 e publicada posteriormente por William Jones em 1711 em *De methodis fluxionum et serierum infinitarum* em 1671, traduzido e publicado como *Method of Fluxions* por John Colson em 1736. A abordagem inicial de Newton difere muito da atual. Newton aplicou seu método inicialmente só para encontrar raízes de polinômios. De fato, ele não calculava uma aproximação sucessiva x_k , mas calculava uma sequência de polinômios e só no final, ele chegava a uma aproximação para a raiz.

O algoritmo obtido pelo método de Gauss-Newton é um método para se obter soluções de problemas de mínimos quadrados não-lineares, o qual pode ser visto como uma modificação do método clássico de Newton para minimizar (ou maximizar) funções. Diferentemente do método de Newton, o método de Gauss-Newton só pode ser aplicado para minimizar problemas de mínimos quadrados.

O algoritmo de Levenberg-Marquardt produz soluções para o problema de minimizar funções não-lineares sobre o espaço dos parâmetros. Este método pode ser considerado como um método que vive entre o método de Gauss-Newton e o método de gradiente. O algoritmo foi publicado por Kenneth Levenberg, enquanto trabalhava no Frankford Army Arsenal, um dos destacamentos militares Norte-Americano. Este foi redescoberto por Donald Marquardt quando trabalhava como estatístico no DuPont.

- Sir Isaac Newton - veja Seção 2.7.
- Joseph Raphson- matemático Inglês. Ficou conhecido pelo mé-

todo de Newton-Raphson. Pouco se sabe de sua vida. Não se tem informações precisas sobre datas de nascimento e morte.

- Gauss - veja Seção 1.4.
- Kenneth Levenberg - sem informações históricas.
- Donald W. Marquardt (1929 - 1997) - estatístico Americano. Redescobriu o algoritmo que hoje leva o nome de Levenberg-Marquardt.

Capítulo 4

Método de 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 estabelecer tal compromisso, para o método de regularização de Tikhonov. Num primeiro momento, trataremos o caso em que o operador A é linear. Trataremos da teoria para o caso em que $A(\cdot)$ é não linear, logo em seguida.

4.1 Problemas Lineares: Convêrgencia

Nesta seção, consideraremos o caso em que o operador A é linear e limitado. Com essa hipótese, estamos interessados em encon-

trar, 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 nos métodos iterativos estudados nos capítulos anteriores, uma solução regularizada requer uma estratégia mais adequada que tentar resolver as equações normais

$$A^*Ax = A^*y^\delta. \quad (4.1)$$

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

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

Lembrando, tudo o que queremos é "inverter" o operador A de maneira estável. Mais que isso, não queremos errar muito ao fazer essa inversão, i.e., queremos manter o resíduo $\|Ax - y^\delta\|$ controlado. Relembre a estimativa (2.11).

Assim, formalmente, gostaríamos de inverter de forma estável o operador A^*A . Pelo Teorema da Aplicação Espectral B.2.10, isso equivale a calcular $g(A^*A)$, onde $g(\lambda) = \frac{1}{\lambda}$, $\lambda \neq 0$. Portanto, uma solução aproximada para a equação (4.1), é dada por

$$x^\delta = g(A^*A)A^*y^\delta. \quad (4.2)$$

Antes de continuarmos, vamos a algumas propriedades importantes, que enunciaremos em forma de exercício.

Exercício 4.1. *Dê um exemplo de um operador linear, que possui*

espectro não-vazio e que não possui autovalores. Pergunta: isso pode acontecer no caso em que A é uma matriz, isto é, em dimensão finita?

Exercício 4.2. *Seja A um operador linear limitado entre espaços de Hilbert. Mostre que o espectro de A^*A é real e positivo.*

Exercício 4.3. *Dê um exemplo de um operador linear limitado A , tal que, 0 seja um elemento do espectro de A . Existem operadores auto-adjuntos, tais que, 0 seja um autovalor? Se afirmativo, dê exemplos. Sugestão: procure exemplos nos espaços de seqüências.*

Aprendemos dos exercícios acima que o espectro de A^*A pode conter elementos próximos de ZERO e até mesmo o conter. Portanto, a estratégia de calcular x^δ , como $x^\delta = g(A^*A)A^*y^\delta$ não é possível, ou, se é possível, é muito instável.

Exercício 4.4. *Mostre que, se λ é um autovalor de A então λ^2 é um autovalor de A^*A . Suponha que A é uma matriz invertível. Mostre que $1/\lambda$ é um autovalor de A^{-1} .*

Exercício 4.5. *Suponha que A^*A é invertível. 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 $\|y - y^\delta\| \leq \delta$?*

Qual é a estratégia? Afastar os elementos espectro de A^*A de zero, ou seja, para $0 < \alpha \in [0, \alpha_0]$, definimos

$$f_\alpha(\lambda) := g(\lambda + \alpha) = \frac{1}{\lambda^2 + \alpha}. \quad (4.3)$$

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

Exercício 4.6. *Mostre que $f_\alpha(\cdot)$ é contínua pela direita.*

Segue do Teorema da Aplicação Espectral que

$$f_\alpha(A^*A) = (A^*A + \alpha I)^{-1}. \quad (4.4)$$

Exercício 4.7. *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 limitado. Sugestão: Use o Teorema do Gráfico Fechado [28] para mostrar a limitação da inversa.*

Segue do exercício acima que a escolha de x_α^δ da forma

$$x_\alpha^\delta = (A^*A + \alpha I)^{-1} A^* y^\delta, \quad (4.5)$$

é uma solução regularizada, definida via a equação linear

$$(A^*A + \alpha I)x_\alpha^\delta = A^* y^\delta. \quad (4.6)$$

Esta pode ser pensada como uma regularização para as equações normais (4.1).

Este método é chamado de **Regularização de Tikhonov**¹.

Exercício 4.8. *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 (4.5), tem a forma*

$$x_\alpha^\delta = \sum_{j=1}^{\infty} \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle y^\delta, f_j \rangle e_j. \quad (4.7)$$

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

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

$$x^\delta = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \langle y^\delta, f_j \rangle e_j. \quad (4.8)$$

Observação: Fazendo uma comparação entre (4.8) e (4.7) vemos, claramente, o resultado de estabilidade da equação (4.7): o erro em $\langle y^\delta, f_j \rangle$ é propagado com um fator de $\frac{\sigma_j}{\sigma_j + \alpha}$, o qual é sempre limitado quando $j \rightarrow \infty$ e não como um fator de $\frac{1}{\sigma_j}$ como em (4.8).

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, mesmo computacionalmente. Seria ideal termos uma outra forma de determinar uma solução pela regularização de Tikhonov. Temos a seguinte versão variacional da regularização de Tikhonov:

Teorema 4.1.1. *Seja x_α^δ como na equação (4.5). Então x_α^δ é o único minimizador do funcional de Tikhonov*

$$J_\alpha(x) := \|Ax - y^\delta\|^2 + \alpha\|x\|^2. \quad (4.9)$$

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

$$J'_\alpha(x).h = 0 \quad \text{para todo } h \in \mathcal{H}_1. \quad (4.10)$$

Disto, segue que

$$\begin{aligned} 0 = J'_\alpha(x).h &= 2\langle Ax - y^\delta, Ah \rangle + 2\alpha\langle x, h \rangle \\ &= 2\langle A^*Ax - A^*y^\delta + \alpha Ix, h \rangle \end{aligned} \quad (4.11)$$

para todo $h \in \mathcal{H}_1$. Portanto, (4.11) é equivalente a (4.5). ■

Observação: Note que, qualquer minimizador do funcional de Tikhonov (4.9) pertence ao $\mathcal{N}(A)^\perp$. Com efeito, pois, caso contrário, poderíamos fazer a segunda parte do funcional crescer, mantendo a primeira parte constante.

Exercício 4.9. *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. Mostre que se J é estritamente convexo e coercivo, então o mínimo é único.*

Exercício 4.10. *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 .*

O parâmetro α no funcional (4.9) é o parâmetro de regularização. Minimização em (4.9) é um compromisso entre minimizar a norma do resíduo $\|Ax - y^\delta\|$ e manter o tamanho do **termo de penalização** $\|x\|$ pequeno e a ssim, 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 a escolha de α através da chamada **curva L**. Um boa referência para a construção da curva L é [3]. Veja Figura 1.1 para a interpretação geométrica.

Observação: Note que a definição de x_α^δ como em (4.5) só tem sentido para operadores lineares. Por outro lado, o problema de minimização (4.9) pode ser formulado para operadores não-lineares. Nas Seções 4.3 e 4.4 desenvolveremos a teoria de regularização de Tikhonov para problemas não lineares.

4.2 Problemas Lineares: Semi-Convergência

A definição da solução regularizada pela minimização do funcional de Tikhonov (4.9) nos fornece, diretamente, resultados de convergência e estabilidade, como:

Teorema 4.2.1. *Seja x_α^δ definida por (4.5), $y \in \mathcal{R}(A)$ com $\|y - y^\delta\| \leq \delta$. Se $\alpha := \alpha(\delta)$ é tal que*

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0 \quad e \quad \lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0, \quad (4.12)$$

então

$$\lim_{\delta \rightarrow 0} x_{\alpha(\delta)}^\delta = A^\dagger y. \quad (4.13)$$

Demonstração : Seja $\delta_n \rightarrow 0$ qualquer. Defina $\alpha_n := \alpha(\delta_n)$ e $x_n := x_{\alpha_n}^{\delta_n}$. Seja J_n o funcional de Tikhonov (4.9) com $\alpha = \alpha_n$ e x_n seu correspondente minimizante (que existe e é único pelo Teorema 4.1.1). Seja $x^\dagger := A^\dagger y$. Então, pela 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,

$$\|x_n\|^2 \leq \frac{\delta_n^2}{\alpha_n} + \|x^\dagger\|^2. \quad (4.14)$$

Portanto, por hipótese $\{x_n\}$ é uniformemente limitada. Pelo Teorema de Banach-Alaoglu (veja Teorema A.1.10), $\{x_n\}$ possui uma subsequência que converge fraco² para $z \in \mathcal{H}$. Como A é linear e

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: $x_n \rightharpoonup z$.

limitado,

$$Ax_{n_k} \rightharpoonup Az. \quad (4.15)$$

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 (4.12) e (4.15) que

$$Az = y. \quad (4.16)$$

Da Observação 4.1 temos que $x_n \in \mathcal{N}(A)^\perp$ e, assim, $z \in \mathcal{N}(A)^\perp$ (Prove). Pelo Teorema B.1.7 e sua demonstração, obtemos $z = x^\dagger$. Assim, $x_{n_k} \rightharpoonup x^\dagger$. Aplicando o mesmo argumento para todas as subsequências obtemos que

$$x_n \rightharpoonup x^\dagger. \quad (4.17)$$

Afirmção: $x_n \rightarrow x^\dagger$.

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 (4.17). Logo,

$$\liminf \|x_n\| \geq \|x^\dagger\|. \quad (4.18)$$

De (4.14), temos

$$\liminf \|x_n\| \leq \|x^\dagger\|. \quad (4.19)$$

Juntando as últimas três estimativas, temos que $x_n \rightarrow x^\dagger$. ■

Exercício 4.11. Nas hipóteses do Teorema acima, mostre que (4.17)

- (4.19) implica que $x_n \rightarrow x^\dagger$.

4.2.1 Caso Linear: Taxas de Convergência

Segue da definição de solução regularizada pelo método de Tikhonov que

$$\|x_\alpha^\delta - x^\dagger\| \leq \sup_{\lambda \in \Sigma(A)} |\lambda f_\alpha(\lambda)| \|y - y^\delta\| \leq \frac{\delta}{\sqrt{\alpha}}. \quad (4.20)$$

Assim, se $\alpha \sim \delta$, obtemos a seguinte ordem de convergência

$$\|x^\dagger - x_\alpha^\delta\| = \mathcal{O}(\sqrt{\delta}). \quad (4.21)$$

Exercício 4.12. Verifique a desigualdade na equação (4.20), no caso em que A é compacto e com um sistema singular dado por (σ_j, e_j, f_j) .

4.3 Problemas Não-Lineares: Convergência

Nesta seção, estudaremos a teoria de regularização de Tikhonov para o caso em que o operador $A(\cdot)$ é não-linear. Assim, como para os métodos iterativos, as hipóteses para o caso não linear são mais restritivas. Para o caso da regularização de Tikhonov assumiremos que:

Hipótese 4.1. O problema $A(x) = y$ possui uma solução $x_* \in B_\rho(x_0)$, com $x_0 \in \mathcal{D}(A)$. E ainda,

1. $A(\cdot)$ é contínuo em seu domínio $\mathcal{D}(A)$ e
2. $A(\cdot)$ é fracamente (sequencialmente) fechado, isto é, para qualquer sequência $\{x_n\} \subset \mathcal{D}(A)$, com x_n convergindo fraco para x

e $A(x_n)$ convergindo fraco para y , então $x \in \mathcal{D}(A)$ e $A(x) = y$.

Como observado anteriormente, no caso de operadores não lineares, a construção da solução regularizada, pela função filtro f , não se aplica. Por outro lado, podemos utilizar de técnicas variacionais e minimizar um funcional como no caso linear. Como o intuito é obter uma solução regularizada para o problema inverso, isto é, uma solução que dependa continuamente do nível de ruído δ , nos concentraremos na minimização do funcional

$$\mathcal{F}_\alpha(x) = \|A(x) - y^\delta\|^2 + \alpha\|x - x_0\|^2, \quad (4.22)$$

onde $\alpha > 0$ é o parâmetro de regularização e $x_0 \in \mathcal{D}(A)$ é um *a priori* para a solução exata do problema.

O primeiro passo é mostrar que o funcional (4.22) está bem definido, isto é, que possui um minimizante no $\mathcal{D}(A)$.

Lema 4.3.1. *Dado $y \in \mathcal{H}_2$. Se a Hipótese 4.1 é satisfeita, então o funcional (4.22) possui um mínimo.*

Demonstração : Como $x_0 \in \mathcal{D}(A)$ temos que $\mathcal{F}_\alpha(x_0) < \infty$. Ainda, $\mathcal{F}_\alpha(x) \geq 0$ para todo $x \in \mathcal{D}(A)$. Assim, o conjunto imagem do funcional \mathcal{F}_α é não vazio e limitado inferiormente em \mathbb{R} . Portanto, possui um ínfimo. Pela definição de ínfimo, existe uma sequência $x_n \in \mathcal{D}(A)$ com $\mathcal{F}_\alpha(x_n) \rightarrow \inf \mathcal{F}_\alpha$. Da definição de $\mathcal{F}_\alpha(\cdot)$,

$$\|A(x_n) - y\|^2 + \alpha\|x_n - x_0\|^2 < \infty.$$

Portanto, as sequências $\{\|A(x_n) - y\|\}$ e $\{\|x_n - x_0\|\}$ são limitadas. Assim, existem subsequências $\{x_{n_k}\}$ e $A(x_{n_k})$ convergindo fraco para x e z , respectivamente. Pela Hipótese 4.1 $x \in \mathcal{D}(A)$ e $A(x) = z$. Pela

continuidade fraca inferior da norma em espaços de Hilbert

$$\mathcal{F}_\alpha(x) = \liminf \mathcal{F}_\alpha(x_n) = \inf \mathcal{F}_\alpha. \quad \blacksquare$$

4.4 Problemas Não-Lineares: Semi-Convergência

Agora, estamos em condições de mostrar que os minimizantes do funcional (4.22) dependem continuamente dos dados y^δ , para escolhas adequadas do parâmetro de regularização α .

Teorema 4.4.1. *Sejam $\alpha > 0$, $\{y_n\}$ e $\{x_n\}$ sequências tais que $y_n \rightarrow y^\delta$ e x_n o correspondente minimizante do funcional de Tikhonov (4.22), com y^δ substituído por y_n . Então, $\{x_n\}$ possui pelo menos uma subsequência convergente e o limite de qualquer subsequência é um minimizante de (4.22).*

Demonstração : Pelo Lema 4.3.1, a sequência x_n está bem definida. Ainda, pela definição de x_n temos que

$$\|A(x_n) - y_n\|^2 + \alpha \|x_n - x_0\|^2 \leq \|A(x) - y_n\|^2 + \alpha \|x - x_0\|^2,$$

para qualquer $x \in \mathcal{D}(A)$. Da mesma forma que no Lema 4.3.1, temos que

$$x_{n_k} \rightharpoonup \tilde{x} \quad e \quad A(x_{n_k}) \rightharpoonup A(\tilde{x}).$$

Pela semi-continuidade fraca inferior da norma

$$\|\tilde{x} - x_0\| \leq \liminf \|x_{n_k} - x_0\|$$

e

$$\|A(\tilde{x}) - y^\delta\| \leq \liminf \|A(x_{n_k}) - y_{n_k}\|. \quad (4.23)$$

Assim,

$$\begin{aligned} \|A(\tilde{x}) - y^\delta\|^2 + \alpha\|\tilde{x} - x_0\|^2 &\leq \liminf_{k \rightarrow \infty} (\|A(x_{n_k}) - y_{n_k}\|^2 + \alpha\|x_{n_k} - x_0\|^2) \\ &\leq \limsup_{k \rightarrow \infty} (\|A(x_{n_k}) - y_{n_k}\|^2 + \alpha\|x_{n_k} - x_0\|^2) \\ &\leq \lim_{k \rightarrow \infty} (\|A(x) - y_{n_k}\|^2 + \alpha\|x - x_0\|^2) \\ &= \|A(x) - y^\delta\|^2 + \alpha\|x - x_0\|^2, \end{aligned}$$

para qualquer $x \in \mathcal{D}(A)$. Isso implica que \tilde{x} é um minimizador de (4.22) e

$$\|A(\tilde{x}) - y^\delta\|^2 + \alpha\|\tilde{x} - x_0\|^2 = \lim_{k \rightarrow \infty} (\|A(x_{n_k}) - y_{n_k}\|^2 + \alpha\|x_{n_k} - x_0\|^2). \quad (4.24)$$

Assuma que $x_{n_k} \not\rightarrow \tilde{x}$. Então $c := \limsup \|x_{n_k} - x_0\| > \|\tilde{x} - x_0\|$ e existe uma subsequência $\{x_m\}$ de $\{x_{n_k}\}$ tal que $x_m \rightarrow \tilde{x}$ e $A(x_m) \rightarrow A(\tilde{x})$, com $\|x_m - x_0\| \rightarrow c$.

Logo, de (4.24)

$$\lim_{m \rightarrow \infty} \|A(x_m) - y_{n_k}\|^2 = \|A(\tilde{x}) - y^\delta\|^2 + \alpha(\|\tilde{x} - x_0\|^2 - c^2) < \|A(\tilde{x}) - y^\delta\|^2$$

o que contradiz (4.23). Logo $x_{n_k} \rightarrow \tilde{x}$. \blacksquare

O próximo resultado imita o Teorema 4.2.1, para o caso não linear.

Teorema 4.4.2. *Seja $y^\delta \in \mathcal{H}_2$, com $\|y - y^\delta\| \leq \delta$ e seja $\alpha(\delta)$ satisfazendo $\alpha(\delta) \rightarrow 0$ e $\delta^2/\alpha(\delta) \rightarrow 0$, quando $\delta \rightarrow 0$. Então, toda sequência $\{x_{\alpha_k}^{\delta_k}\}$, onde $\delta_k \rightarrow 0$ e $\alpha_k := \alpha(\delta_k)$ e $x_{\alpha_k}^{\delta_k}$ é um mini-*

mizante de (4.22), possui uma subsequência convergente. O limite de toda subsequência convergente é uma solução de norma mínima x^\dagger .

Demonstração : Seja α_k e $x_{\alpha_k}^{\delta_k}$ como acima e seja x^\dagger uma solução de norma mínima. Por definição de $x_{\alpha_k}^{\delta_k}$ temos

$$\begin{aligned} \|A(x_{\alpha_k}^{\delta_k}) - y^{\delta_k}\|^2 + \alpha_k \|x_{\alpha_k}^{\delta_k} - x_0\|^2 &\leq \|A(x^\dagger) - y^{\delta_k}\|^2 + \alpha_k \|x^\dagger - x_0\|^2 \\ &\leq \delta_k^2 + \alpha_k \|x^\dagger - x_0\|^2. \end{aligned}$$

Assim, segue da hipótese que

$$\lim_{k \rightarrow \infty} A(x_{\alpha_k}^{\delta_k}) = y \quad (4.25)$$

e

$$\limsup_{k \rightarrow \infty} \|x_{\alpha_k}^{\delta_k} - x_0\| \leq \|x^\dagger - x_0\|. \quad (4.26)$$

Portanto, a sequência $\{x_{\alpha_k}^{\delta_k}\}$ é limitada e assim, possui uma subsequência (que denotamos com o mesmo índice) tal que $\{x_{\alpha_k}^{\delta_k}\} \rightharpoonup \tilde{x}$.

Da hipótese 4.1 e de (4.25), $\tilde{x} \in \mathcal{D}(A)$ e $A(\tilde{x}) = y$.

Da semi-continuidade fraca inferior da norma e de (4.26)

$$\|\tilde{x} - x_0\| \leq \liminf_{k \rightarrow \infty} \|x_{\alpha_k}^{\delta_k} - x_0\| \leq \limsup_{k \rightarrow \infty} \|x_{\alpha_k}^{\delta_k} - x_0\| \leq \|x^\dagger - x_0\|.$$

Logo, \tilde{x} também é uma solução de norma mínima.

Da identidade

$$\|x_{\alpha_k}^{\delta_k} - \tilde{x}\|^2 = \|x_{\alpha_k}^{\delta_k} - x_0\|^2 + \|\tilde{x} - x_0\|^2 + 2\langle x_{\alpha_k}^{\delta_k} - x_0, x_0 - \tilde{x} \rangle$$

e de (4.26) obtemos que

$$\limsup_{k \rightarrow \infty} \|x_{\alpha_k}^{\delta_k} - \tilde{x}\|^2 \leq 2\|\tilde{x} - x_0\|^2 + 2 \lim_{k \rightarrow \infty} \langle x_{\alpha_k}^{\delta_k} - x_0, x_0 - \tilde{x} \rangle = 0.$$

Portanto, $\lim_{k \rightarrow \infty} x_{\alpha_k}^{\delta_k} = \tilde{x}$. ■

4.4.1 Taxas de Convergência

Para obtermos taxas, as hipóteses sobre o operador A são as mesmas feitas para obter taxas no método de Landweber iterativamente regularizado.

Teorema 4.4.3. *Seja a hipótese 2.2 satisfeita. Então para a escolha de $\alpha \sim \delta$ obtemos*

$$\|x_{\alpha}^{\delta} - x^{\dagger}\| = \mathcal{O}(\sqrt{\delta}) \quad e \quad \|A(x_{\alpha}^{\delta}) - y^{\delta}\| = \mathcal{O}(\delta). \quad (4.27)$$

Demonstração : Veja [8, Teo 10.4]. ■

4.5 Breve Apanhado Histórico do Capítulo

O método de regularização de Tikhonov tem sido investigado independentemente, para uma quantidade muito grande de contextos. Este método ficou conhecido pelas aplicações em equações integrais com os trabalhos de A. N. Tikhonov e D. L. Phillips. Por isso, alguns autores gostam de usar o termo regularização de Tikhonov-Phillips. Em dimensão finita, este método foi estudado por A. E. Hoerl, que usou uma contestualização estatística e por M. Foster, que interpretou o método como um filtro de Wiener-Kolmogorov. Segundo Hoerl, este é conhecido na literatura estatística como **ridge regression**.

Dependendo da aplicação, diferentes termos de regularização são considerados, melhorando taxas de convergência e identificando características especiais das soluções [8, 46].

- Andrey Nikolayevich Tychonoff (1906 – 1993) - matemático Soviético e Russo. Conhecido por importantes contribuições

em topologia, análise funcional, física matemática e problemas mal-postos. A grafia Tychonoff é graças a tradução dos trabalhos para o Alemão. O termo "Tikhonov" é devido a tradução para o Inglês.

- Stefan Banach (1892 - 1945) - matemático Polonês. Foi o fundador da análise funcional moderna. Seu livro, *Théorie des opérations linéaires (Theory of Linear Operations)* de 1932, foi a primeira monografia sobre teoria geral para espaços métricos lineares.
- Leonidas Alaoglu (1914 1981) - matemático Canadense-Americano. Seu mais famoso trabalho citado uma enormidade de vezes é conhecido como Teorema de Alaoglu o qual prova que a bola unitária fechada do espaço dual é compacta na topologia fraca-*. Esse teorema também é conhecido como Teorema de Banach-Alaoglu.

Apêndice A

Definições e Resultados

Aqui, postaremos algumas definições importantes e Teoremas de Análise Funcional que são utilizados com frequência nas notas. Não apresentaremos demonstrações, mas, encorajamos o leitor que as desconhece, a dar-lhes a devida atenção e realizar os exercícios complementares. As referências os guiarão a um material muito rico sobre o assunto.

A.1 Definições e Resultados Básicos em Espaços Vetoriais.

Seja \mathcal{H} um espaço vetorial sobre o corpo dos números complexos \mathbb{C} .

Definição A.1.1. *Uma **norma** em \mathcal{H} é um funcional*

$$\begin{aligned} \|\cdot\| : \mathcal{H} &\longrightarrow \mathbb{R} \\ x &\longmapsto \|x\|, \end{aligned}$$

que, para todo $u, v, w \in \mathcal{H}$ satisfaz:

$$(i) \|u\| \geq 0 \text{ e } \|u\| = 0 \text{ se, e somente se } u = 0.$$

$$(ii) \|\alpha u\| = |\alpha| \|u\|$$

$$(iii) \|u + w\| \leq \|u + v\| + \|v + w\|.$$

Ao par, $(\mathcal{H}, \|\cdot\|)$ chamamos de espaço vetorial normado. Se \mathcal{H} é completo, com relação a norma $\|\cdot\|$, chamamos \mathcal{H} de espaço de Banach.

Definição A.1.2. Um **produto interno** em \mathcal{H} é uma aplicação

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} &\longrightarrow \mathbb{C} \\ (u, v) &\longmapsto \langle u, v \rangle, \end{aligned}$$

que satisfaz os seguintes axiomas:

$$(i) \langle \alpha u + \beta w, v \rangle = \alpha \langle u, v \rangle + \beta \langle w, v \rangle,$$

$$(ii) \langle u, v \rangle = \overline{\langle v, u \rangle}, \quad \forall u, v, w \in \mathcal{H}, \forall \alpha, \beta \in \mathbb{C}.$$

Exercício A.1. Mostre que, em \mathbb{R}^n , a aplicação $\langle u, v \rangle = \sum_{j=1}^n u_j \bar{v}_j$ é um produto interno.

Exercício A.2. Seja \mathcal{H} um espaço com produto interno. Defina a aplicação

$$\begin{aligned} \|\cdot\| : \mathcal{H} &\longrightarrow \mathbb{R} \\ x &\longmapsto \|x\| := \sqrt{\langle x, x \rangle}. \end{aligned}$$

Mostre que a aplicação está bem definida. Mostre que $\|\cdot\|$ é uma norma em \mathcal{H} . Assim, $(\mathcal{H}, \|\cdot\|)$ é um espaço vetorial normado. Chamamos $(\mathcal{H}, \|\cdot\|)$ de um espaço pré-Hilbert.

Definição A.1.3. *Seja \mathcal{H} um espaço com produto interno. Se \mathcal{H} é completo como relação a norma $\|x\| := \sqrt{\langle x, x \rangle}$, então denominamos \mathcal{H} um espaço de Hilbert.*

A.1.1 Operadores Lineares

Definição A.1.4. *Seja $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ uma aplicação entre os espaços de Hilbert \mathcal{H}_1 e \mathcal{H}_2 . A é um **operador linear** se*

$$A(\alpha x + y) = \alpha Ax + Ay \quad \text{para qualquer } x, y \in \mathcal{H}_1, \text{ e } \alpha \in \mathbb{C}.$$

*O operador A é **limitado**, se existir $C > 0$ tal que*

$$\|Ax\| \leq C\|x\| \quad \forall x \in \mathcal{H}_1.$$

Definição A.1.5. *Seja A um operador linear e limitado. A **norma** de A é definida como*

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

Definição A.1.6. *Seja $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ um operador linear limitado. Dizemos que o operador A é compacto se para qualquer $S \subset \mathcal{H}_1$ subconjunto limitado, $A(S) \subset \mathcal{H}_2$ é um conjunto pré-compacto.*

Definição A.1.7. *Seja $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ um operador linear limitado. O operador adjunto de Hilbert¹, denotado por A^* , é definido por*

$$\langle Ax, y \rangle = \langle x, A^*y \rangle, \quad \forall x \in \mathcal{H}_1, y \in \mathcal{H}_2.$$

¹Existe uma definição de adjunto para operadores entre espaços de Banach. Ainda, a definição acima pode ser estendida a operadores lineares, não necessariamente limitados.

O operador $A : \mathcal{H} \rightarrow \mathcal{H}$ é **auto-adjunto** se

$$\langle Ax, y \rangle = \langle x, Ay \rangle, \quad \forall x, y \in \mathcal{H}.$$

Exercício A.3. *Seja A limitado. Mostre que o adjunto de A também é um operador limitado. O que você pode dizer sobre a norma do operador adjunto?*

Exercício A.4. *Sejam A um operador compacto e B um operador limitado. Mostre que, se os produtos AB e BA estão bem definidos, então ambos os produtos geram operadores compactos. Em particular, se A é um operador compacto, mostre que A^*A e AA^* são operadores compactos e auto-adjuntos.*

Observação: Operadores compactos e/ou auto-adjuntos possuem propriedades muito importantes e já foram amplamente estudados. Veja [28].

Definição A.1.8. *Um operador linear P é dito uma projeção ortogonal se $P^2 = P$ e P é auto-adjunto.*

O próximo Teorema é dito Teorema da Projeção.

Teorema A.1.9. *Seja \mathcal{H} um espaço de Hilbert e M um subespaço fechado de \mathcal{H} . Então, para qualquer $x \in \mathcal{H}$ existe um único $m_0 \in M$ tal que*

$$\|x - m_0\| \leq \|x - m\|, \quad \forall m \in M.$$

Demonstração: Veja [28].

Exercício A.5. *Mostre que, se M é um subespaço fechado de um espaço de Hilbert \mathcal{H} , então $\mathcal{H} = M \oplus M^\perp$. Além disso, dado qualquer $x \in \mathcal{H}$, $x = m + m^\perp$ de forma única, onde $m \in M$ e $m^\perp \in M^\perp$. Assim, fica bem definido o operador $P : \mathcal{H} \rightarrow M$ como $Px = m$. Mostre que P é um operador de projeção.*

Teorema A.1.10. [Banach-Alaoglu] *Seja \mathbb{X} um espaço de Banach. Então a bola $B[0, 1]$ é compacta na topologia fraca* de \mathbb{X} . Em particular, se \mathbb{X} é um espaço reflexivo, então a bola $B[0, 1]$ é compacta na topologia fraca.*

Demonstração : Veja textos de Análise Funcional, como [28].

Teorema A.1.11. *Todo espaço de Hilbert é reflexivo.*

Demonstração : Veja [28].

A.1.2 Transformada de Fourier

A Transformada de Fourier tem um papel importante na apresentação da Transformada de Radon no Capítulo 1. Faremos aqui um breve resumo da Análise de Fourier, que por si só, pode ser considerada como uma área da Matemática.

Denotamos por

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : |f|_{k,l} = \sup_{x \in \mathbb{R}^n} |x^k \partial^l f(x)| < \infty, \forall k, l \in \mathbb{N}\}, \quad (\text{A.1})$$

como o conjunto de todas as funções que são infinitamente diferenciáveis em \mathbb{R}^n e cujas derivadas de todas as ordens convergem para zero, mais rápido que qualquer polinômio. Este é um espaço vetorial, o qual não é normado, ao qual chamamos de Espaço de Schwartz.

Para toda função $f \in \mathcal{S}(\mathbb{R}^n)$, a Transformada de Fourier de f é definida como

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-ix \cdot \xi} dx. \quad (\text{A.2})$$

Exercício A.6. *Seja $f \in \mathcal{S}(\mathbb{R}^n)$. Mostre que $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$.*

Exercício A.7. *Mostre que $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ é uma transformação linear, limitada.*

Exercício A.8. *Seja $f \in \mathcal{S}(\mathbb{R}^n)$. Mostre que se $\mathcal{F}(f)(\xi) = 0$ para todo $\xi \in \mathbb{R}^n$, então $f = 0$. Isto é uma transformação de Fourier é injetiva.*

Portanto, esta transformação possui uma inversa, que chamamos **Transformada Inversa de Fourier** definida por

$$\check{f}(x) = \mathcal{F}^{-1}(f)(x) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(\xi) e^{i\xi \cdot x} d\xi. \quad (\text{A.3})$$

Da definição da transformada inversa de Fourier, segue que

$$\mathcal{F}(\mathcal{F}^{-1}(f)(x)) = \mathcal{F}^{-1}(\mathcal{F}(f)(x)) = f(x) \quad \forall f \in \mathcal{S}(\mathbb{R}^n).$$

Isto é, a transformada de Fourier é um isomorfismo em $\mathcal{S}(\mathbb{R}^n)$.

A Transformada de Fourier se estende como um isomorfismo isométrico ao espaço vetorial das funções mensuráveis e quadrado integráveis $L^2(\mathbb{R}^n)$, onde vale a identidade de Parseval

$$\|f\|_{L^2} = \|\hat{f}\|_{L^2}.$$

A Transformada de Fourier (A.2), também é bem definida para funções mensuráveis e módulo integráveis, isto é, para $f \in L^1(\mathbb{R}^n)$.

Exercício A.9. *Mostre que, para $f \in L^1(\mathbb{R}^n)$, a Transformada de Fourier está bem definida e que*

$$\|\hat{f}\|_{L^\infty} \leq c\|f\|_{L^1}.$$

A Transformada de Fourier possui as seguintes propriedades

Exercício A.10. *Seja $f \in \mathcal{S}(\mathbb{R})$. Prove que:*

$$1. f(x - c) = e^{-ic\xi} \hat{f}(\xi).$$

$$2. f(cx) = c^{-1} \hat{f}(c^{-1}\xi).$$

$$3. \hat{f}'(\xi) = i\xi \hat{f}(\xi).$$

$$4. \hat{f}^{(n)}(\xi) = (i\xi)^n \hat{f}(\xi).$$

Apêndice B

Sistemas de Equações Lineares

Neste apêndice, faremos uso dos conceitos de 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 inverso 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 tentar inverter esse operador.

Apresentaremos a seguir 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. O primeiro destes conceitos nos permitirá definir uma solução, com uma propriedade especial,

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 [28]

dentre todas as possíveis soluções (que podem ser infinitas), de um problema inverso. O segundo, também chamado de Teorema Espectral, permite decompor um operador (uma matriz) em uma forma mais simples, como a soma de projeções sobre subespaços adequados, e entender realmente a influência do espectro de um operador na solução de problemas inversos.

B.1 Pseudo - Inversa de Operadores Lineares

A bem da verdade, tudo o que queremos em problemas inversos é, encontrar uma maneira de aproximar a inversa de um operador (o operador inverso) por uma família de operadores bem postos e, conseqüentemente, encontrar uma aproximação (a melhor possível) para a solução do problema. Nesta seção, apresentaremos uma forma de aproximar “da melhor forma” o operador inverso.

Desenvolveremos a teoria para operadores lineares limitados que possuam imagem fechada, o que é o caso de operadores em dimensão finita (matrizes) e também de alguns operadores compactos. Assim, cobriremos uma ampla quantidade de casos interessantes. Faz jus mencionar que existem versões, dos resultados apresentadas abaixo, para operadores lineares limitados quaisquer [13].

B.1.1 Definições e Propriedades Básicas

Se o leitor não está habituado a certas definições pertinentes a teoria de operadores lineares limitados, sugerimos que faça uma breve leitura das definições e resultados contidos no Apêndice A.

Esperamos que o leitor esteja familiarizado com as definições de matrizes **Hermitiana**, **Simétricas**, **Unitárias**, **Normais**, etc.

Desta forma, podemos passar direto as definições mais gerais como as de operadores em espaços de Hilbert (veja Apêndice A). Nosso objetivo é encurtar caminho, indo direto ao Teorema Espectral² (dimensão infinita) para operadores compactos e auto-adjuntos e obter, como corolário, o Teorema da SVD (dimensão finita).

Exercício B.1. *Faça um paralelo entre as definições de matrizes Hermitiana, Simétricas, Unitárias, Normais e operadores adjuntos e auto-adjuntos encontrados no Apêndice A.*

Suponha que temos um operador linear limitado $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ entre espaços de Hilbert. Consideraremos o problema fundamental de resolver a equação linear do tipo

$$Ax = y, \quad (\text{B.1})$$

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

Exemplo B.1.1. *Exemplos da equação (B.1) são:*

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 Kernel³.

Exercício B.2. *Prove que, ambos os exemplos apresentados acima para a equação (B.1) são operadores compactos. Sugestão: Para o*

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

³Não confundir com o núcleo do operador A

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 $L^2[0, 1]$.

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

Um fato confortante é que, mesmo no caso da equação (B.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 (B.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 (veja também o Exercício ??)

Hipótese B.1. A é um operador linear, limitado e que $\mathcal{R}(A)$ é fechada em \mathcal{H}_2 .

Seja $P : \mathcal{H}_2 \rightarrow \mathcal{R}(A)$ o operador de projeção ortogonal (que está bem definido pelo Exercício B.6). Assim, $Py \in \mathcal{R}(A)$ é o vetor mais próximo de y .

Exercício B.3. Suponha a Hipótese B.1 satisfeita e P o operador de projeção ortogonal sobre $\mathcal{R}(A)$. Prove que, dado $y \in \mathcal{H}_2$, $Py - y \in (\mathcal{R}(A))^\perp$.

Definição B.1.2. Uma solução generalizada de (B.1) é qualquer solução $u \in \mathcal{H}_1$ da equação

$$Ax = Py. \quad (\text{B.2})$$

Exemplo B.1.3. Suponha que $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ e $y = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Então $\mathcal{R}(A) = \text{span}\{(1, -1)\}$ e $Py = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$. Portanto o conjunto de soluções generalizadas é dada por

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

Geometricamente, obter uma solução generalizada, dada como na Definição B.1.2, significa encontrar $u \in \mathcal{H}_1$ solução do problema de minimização

$$u = \arg \min_{x \in \mathcal{H}_1} \frac{1}{2} \|Ax - y\|^2. \quad (\text{B.3})$$

De forma geral temos:

Teorema B.1.4. *Suponha que $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ é um operador linear limitado, $y \in \mathcal{H}_2$ e a Hipótese B.1 é 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 B.3 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 (B.2). Disto e do Teorema de Pitágoras, temos que

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

Portanto

$$Au - y = Py - y \in (\mathcal{R}(A))^\perp = \mathcal{N}(A^*).$$

Assim, $A^*(Au - y) = 0$ e (iii) segue-se.

(iii) \Rightarrow (i): De (iii) satisfeito, obtemos que $Au - y \in (\mathcal{R}(A))^\perp$ e, assim,

$$0 = P(Au - y) = Au - Py. \quad \blacksquare$$

Definição B.1.5. Um vetor $u \in \mathcal{H}_1$ satisfazendo qualquer uma das sentenças do Teorema B.1.4 é chamado uma **solução de quadrados mínimos** da equação $Ax = y$.

Exercício B.4. Mostre que o conjunto de soluções de quadrados mínimos pode ser escrito como

$$\{u \in \mathcal{H}_1 : A^*Au = A^*b\}. \quad (\text{B.4})$$

Prove que este conjunto é convexo e fechado.

Uma observação importante a ser feita é a de que, sob a Hipótese B.1, uma solução de mínimos quadrados de (1.1) sempre existe $\forall b \in \mathcal{H}_2$ (veja Exercício B.5). Caso $\mathcal{N}(A) \neq 0$, então existe uma infinidade de tais soluções. De fato, se u é uma solução de mínimos quadrados e $v \in \mathcal{N}(A)$, então $u + \alpha v$ também é uma solução de mínimos quadrados.

Exercício B.5. Sobre a Hipótese B.1, prove que existe pelo menos uma solução de quadrados mínimos. Dê condições sobre o operador A para que a solução de quadrados mínimos seja única.

Exercício B.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

projecção de b sobre C . Prove ainda que a projecção de b sobre C , é o vetor de C mais próximo de b .

Estamos buscando um caminho para "inverter" o operador A , associando a cada $b \in \mathcal{H}_2$, uma única solução de quadrados mínimos. 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) Sabemos que o conjunto de soluções de quadrados mínimos é não-vazio.
- ii) Pelo Exercício B.4, o conjunto de soluções de quadrados mínimos é **convexo e fechado**.

Portanto, pelo *Teorema da Projecção* A.1.9, existe uma única solução de quadrados mínimos com norma mínima associada a cada elemento $b \in \mathcal{H}_2$. Logo, temos um caminho para definir um processo de inversão para o operador A como:

Definição B.1.6. *Seja A um operador satisfazendo a Hipótese B.1. A aplicação*

$$A^\dagger : \mathcal{H}_2 \longrightarrow \mathcal{H}_1$$

definida por

$$A^\dagger b = u,$$

*onde u é a **única** solução de quadrados mínimos de norma mínima da equação (B.1), é chamada de Inversa Generalizada⁴ de A .*

⁴Para um apanhado histórico muito interessante sobre inversa generalizada, consulte [43]

Exercício B.7. *Mostre que, se A possui uma inversa, então $A^\dagger = A^{-1}$.*

Existem outras definições de inversa generalizada de operadores, que são equivalentes a dada acima (veja [13]).

Exercício B.8. *Mostre que, se A satisfaz a Hipótese B.1, então $\mathcal{R}(A^*)$ é fechada e $\mathcal{N}(A)^\perp = \mathcal{R}(A^*)$.*

Teorema B.1.7. *Se A satisfaz a Hipótese B.1, então $\mathcal{R}(A^\dagger) = \mathcal{R}(A^*) = \mathcal{R}(A^\dagger A)$.*

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 B.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 quadrados mínimos 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 B.1.4.

Suponha 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 que $A^\dagger b$ é uma solução de quadrados mínimos 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 quadrados mínimos. Se x é qualquer

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

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

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

Corolário B.1.8. *Se A satisfaz a Hipótese B.1, então $A^\dagger : \mathcal{H}_2 \longrightarrow \mathcal{H}_1$ é linear e limitado.*

Exercício B.9. *Demonstre o Corolário B.1.8.*

Interpretação Geométrica da Inversa Generalizada

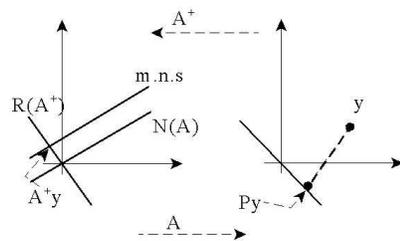


Figura B.1: Interpretação geométrica da Pseudo-Inversa.

Do ponto de vista computacional é importante ter condições mais simples para representar o operador A^\dagger . Esse é a situação se ambos \mathcal{H}_1 e \mathcal{H}_2 têm dimensão finita. Neste caso, sabemos que A possui uma representação matricial e encontrar A^\dagger reduz-se a calcular a inversa generalizada de uma matriz. Em qualquer forma, temos:

Teorema B.1.9. *Suponha que A satisfaça a Hipótese B.1. Então,*

$$A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger.$$

Demonstração : Faremos a demonstração da primeira das igualdades, a segunda é similar e, assim, um bom exercício.

Seja $y \in \mathcal{H}_2$. Pelo Teorema da Projeção $y = y_1 + y_2 \in \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ e assim, $A^*y = A^*y_1 \in \mathcal{R}(A^*A)$. Portanto, se $b \in \mathcal{H}_2$, então

$$A^*A(A^*A)^\dagger A^*b = P_{\mathcal{R}(A^*A)}A^*b = A^*b.$$

Logo, $(A^*A)^\dagger A^*b$ é uma solução de quadrados mínimos (satisfaz as equações normais). Consequentemente,

$$(A^*A)^\dagger A^*b = A^\dagger b + v,$$

para algum $v \in \mathcal{N}(A)$. Como $(A^*A)^\dagger A^*b \in \mathcal{R}((A^*A)^\dagger)$, segue do Teorema B.1.7 que $(A^*A)^\dagger A^*b \in \mathcal{R}((A^*A)) = \mathcal{R}(A^*) = \mathcal{N}(A)^\perp$. Portanto, $(A^*A)^\dagger A^*b = A^\dagger b$. ■

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

Teorema B.1.10. *A inversa generalizada A^\dagger possui gráfico $Gr(A^\dagger)$ fechado. Consequentemente, A^\dagger é contínua se, e só se, $\mathcal{R}(A)$ é fechada.*

Demonstração : A demonstração é um pouco truncada e foge ao

escopo destas notas, pois usa conceitos fortes de Análise Funcional em particular o Teorema do Gráfico Fechado [28]. Interessados na demonstração podem consultar [8]. ■

Observação: O Teorema B.1.10 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.

Exercício B.10. Prove que, se A é um operador linear entre espaços de dimensão finita, então $\mathcal{R}(A)$ é fechada.

Exercício B.11. Suponha que A satisfaça a Hipótese B.1. Prove que, $A^\dagger : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ é o único operador linear limitado satisfazendo

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

Esta é a definição de Moore para Inversa Generalizada. (Sugestão: Consulte [13])

B.2 A Decomposição em Valores Singulares

Um dos principais resultados da Álgebra Linear é o *Teorema de Decomposição em Valores Singulares (SVD)*. Este teorema nos permite escrever uma matriz qualquer, como uma soma de matrizes de projeção de posto 1. Mais geralmente, o Teorema de SVD vale para operadores lineares em espaços de Hilbert de dimensão infinita que possuem a propriedade de ser auto-adjuntos⁵. Neste caso, o

⁵Um operador A entre espaços de Hilbert é dito ser Auto-Adjunto se $\mathcal{D}(A^*) = \mathcal{D}(A)$ e $\langle x, Ay \rangle = \langle Ax, y \rangle, \forall x, y \in \mathcal{D}(A)$

Teorema de SVD é conhecido como *Teorema Espectral* (veja [28]). 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.

Começaremos a entrar em mais detalhes a partir de agora. Esperamos que o leitor esteja familiarizado com o conceito de autovalores e autovetores da Álgebra Linear e com o conceito de espectro e resolvente⁶ para operadores lineares. Seguem algumas referências importantes para os que querem se aprofundar no assunto [12, 28, 36, 48].

Observação: 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.

Exercício B.12. *Prove que, uma matriz quadrada A possui no máximo n autovalores. Dê um exemplo de uma matriz que não possui autovalores.*

Exercício B.13. *Prove que, considerando o espaço vetorial sobre o corpo dos números complexos, uma matriz quadrada $n \times n$ A possui n autovalores.*

Exercício B.14. *Dê um exemplo, em dimensão infinita, de um operador linear que não possui autovalores.*

Nosso ponto de partida aqui é uma versão simplificada do Teorema SVD, o qual faremos a demonstração. Formulações mais gerais podem ser encontradas em [28, 48, 36, 12].

Teorema B.2.1. [*Diagonalização*] *Seja A uma matriz quadrada de ordem $n \times n$ com um conjunto de n autovetores L.I. Então, A é similar a uma matriz diagonal ou diagonalizável.*

⁶Operadores lineares em dimensão infinita podem possuir elementos no espectro que não são autovalores [28]

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

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

Como S é invertível, temos

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

Exercício B.15. *Mostre que, dado um conjunto L.I. 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 B.16. *Justifique, de maneira adequada, que a matriz S no Teorema acima é de fato invertível.*

Corolário B.2.2. *Seja A uma matriz com de ordem $n \times n$ que possui n autovalores distintos. Então, A é diagonalizável.*

Exercício B.17. *Prove o Corolário B.2.2.*

Pergunta: Será que toda matriz quadrada é diagonalizável? Não, pelo menos, no sentido do Teorema B.2.1. O contra-exemplo é a matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Exercício B.18. *Mostre que a matriz A acima não é diagonalizável no sentido do Teorema B.2.1.*

O Teorema B.2.1 é a versão mais simples do Teorema de Decomposição em valores singulares. Passaremos agora para uma versão mais geral. Para tal, necessitaremos de alguns resultados que tomarão algumas linhas dessas notas.

Teorema B.2.3. *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 B.2.4. [*Teorema Espectral⁷ (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 : Esse é um dos problemas mais importantes da Análise. Não faremos a demonstração, pois foge a nossas pretensões. Para a demonstração veja [28]. ■

Teorema B.2.5. [*Teorema Espectral (A compacto)*]

Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ um operador linear e compacto. Então, existem conjuntos ortonormais (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

$$Ax = \sum_{j=1}^m \sigma_j \langle x, e_j \rangle f_j, \quad x \in \mathcal{H}_1. \quad (\text{B.5})$$

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 B.2.4 existe um conjunto

⁷O Teorema Espectral como enunciado também é conhecido como Teorema de Hilbert-Schmidt. Existem várias versões deste Teorema (Veja [28] e referências)

ortogonal $\{e_1, \dots, e_m\}$ de \mathcal{H}_1 tal que $A^* A e_j = \lambda_j e_j$, onde $0 \leq \lambda_j \in \mathcal{R}$. Defina $\sigma_j = \sqrt{\lambda_j}$ e $f_j = \frac{1}{\sigma_j} A e_j$ (para $\sigma_j > 0$). Um cálculo simples mostra que $\{f_1, \dots, f_m\}$ é um conjunto ortonormal e que a equação (B.5) é satisfeita. ■

Definição B.2.6. Os valores $\sigma_1, \dots, \sigma_m$ são chamados de **valores espectrais** de A . A tripla (σ_j, e_j, f_j) chamamos de **sistema singular** de A .

Exercício B.19. Mostre que o conjunto $\{f_1, \dots, f_m\}$ definido como no Teorema B.2.5 é ortonormal.

Exercício B.20. Mostre que se A^* é um operador linear compacto, então

$$A^* y = \sum_{j=1}^m \sigma_j \langle y, f_j \rangle e_j, \quad x \in \mathcal{H}_1. \quad (\text{B.6})$$

Exercício B.21. Mostre que se A^* é um operador linear compacto, então

$$A^\dagger y = \sum_{j=1}^m \sigma_j^{-1} \langle y, f_j \rangle f_j, \quad y \in \mathcal{D}(A^\dagger). \quad (\text{B.7})$$

Corolário B.2.7. [Teorema espectral em dimensão finita - SVD.] Seja $A \in \mathbb{R}^{m,n}$ com $n \leq m$. Então, existem matrizes unitárias $U \in M_{m \times m}$, $V \in M_{n \times n}$ e uma matriz diagonal com entradas não-negativas $\Sigma := \text{diag}\{\sigma_1, \dots, \sigma_n\}$ tais que

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

Passos para demonstrar o teorema SVD.

Exercício B.22. Mostre que todo operador linear cuja imagem possui dimensão finita é compacto. Consequentemente, toda matriz é um operador linear compacto.

Exercício B.23. *Mostre que, se A é uma matriz, então AA^* e A^*A são operadores compactos e auto-adjuntos.*

Exercício B.24. *Demonstre o Corolário B.2.7.*

B.2.1 Funções de Operadores: 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 *Funções Filtro* que são a peça chave para o entendimento dos Métodos de Regularização dos capítulos anteriores.

Por simplicidade, daremos somente a ideia intuitiva em dimensão finita. Para interessados em detalhes mais aprofundados, recomendamos [28] e livros de Análise Funcional em geral.

Exercício B.25. *Seja A uma matriz como no Teorema B.2.1. Mostre que, $A^2 = S \text{diag}(\lambda_1^2, \dots, \lambda_n^2) S^{-1}$. Use indução para mostrar que $A^p = S \text{diag}(\lambda_1^p, \dots, \lambda_n^p) S^{-1}$, para qualquer $p \in \mathbb{N}$.*

Sejam $t \in [0, T]$ e $g(t) = a_0 + a_1 t + \dots + a_n t^n$ um polinômio de ordem n em t .

Definição B.2.8. *Seja A um operador linear limitado ($\|A\| \in [0, T]$), definimos um polinômio do operador A por*

$$g(A) = a_0 + a_1 A + \dots + a_n A^n$$

onde $g(t) = a_0 + a_1 t + \dots + a_n t^n$.

Exercício B.26. *Mostre que, o polinômio $g(A)$ está bem definido, como um operador linear, para A linear limitado.*

Exercício B.27. *Mostre que, se A satisfaz as hipóteses do Teorema B.2.1, então*

$$g(A) = Sg(\text{diag}(\lambda_1, \dots, \lambda_n))S^{-1}.$$

O que você pode dizer de uma função contínua f aplicada em A ?

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 ver [28].

Exemplo B.2.9. *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 B.2.1.

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

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

que converge uniformemente $\forall t \in \mathbb{R}$.

Usando o Exercício B.25, temos que

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{A^j}{j!} &= I + \frac{A}{1!} + \frac{A^2}{2!} + \dots + \frac{A^n}{n!} + \dots = I + S \frac{\text{diag}(\lambda_1, \dots, \lambda_n)}{1!} S^{-1} \\ &+ S \frac{\text{diag}(\lambda_1^2, \dots, \lambda_n^2)}{2!} S^{-1} + \dots + S \frac{\text{diag}(\lambda_1^k, \dots, \lambda_n^k)}{k!} S^{-1} + \dots \\ &= S \text{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n), \dots) S^{-1} =: \exp(A). \end{aligned}$$

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

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

$$A^{-1} = V\Sigma^{-1}U^T, \quad (\text{B.8})$$

onde $\Sigma^{-1} = \text{diag}\{\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_n}\}$.

Observação: Note que, a inversão de uma matriz pode ser pensada como a função $f(t) = t^{-1}$ aplicada a matriz.

De fato, prova-se que este resultado não é mera coincidência. O Teorema que garante tal situação é

Teorema B.2.10. [Teorema da Aplicação Espectral.] *Seja A um operador linear limitado e $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função contínua. Denotemos por $\Sigma(A)$ o espectro de A . Então,*

$$\Sigma(f(A)) = f(\Sigma(A)).$$

Em particular, se A é compacto, então

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

O Teorema também vale se f for contínua à direita ou à esquerda.

Demonstração: Veja [28]. ■

Esse resultado é extremamente importante no entendimento, para a construção das estratégias de regularização (veja os capítulos anteriores) e para entender a relação existente entre mal-condicionamento de um problema com os respectivos valores espectrais do operador associado.

B.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

$$A^{-1}y = \sum_{j=1}^m \frac{1}{\sigma_j} \langle y, f_j \rangle e_j, \quad y \in \mathcal{R}(A). \quad (\text{B.9})$$

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 nos dados, denotados por y^δ , na direção de um autovetor associado ao autovalor σ_m , para m grande, implicam numa grande variação na solução, $x^\delta = A^{-1}y^\delta$.

Exercício B.28. *Prove que, se A é um operador linear compacto e possui inversa limitada, então a dimensão do espaço vetorial é finita.*

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

Considere 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 $\frac{1}{n}$.

Suponha que a matriz A tenha sua decomposição espectral dada por (B.8), onde os valores singulares são⁸ $\sigma_j = \mathcal{O}(\frac{1}{j})$, $j = 1, 2, \dots, n$. Logo, A é invertível.

Da equação (B.8) temos que a solução com ruídos é dada por

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

⁸ $\mathcal{O}(a)$ significa que $\frac{\mathcal{O}(a)}{a} = \text{constante}$.

Usando o fato de que as matrizes U e V são unitárias e que todas as normas são equivalentes (dimensão finita), temos a seguinte estimativa para a solução

$$\begin{aligned} \|x - x^\delta\|^2 &= \|V\Sigma^{-1}U^T y - V\Sigma^{-1}U^T y^\delta\|^2 & (\text{B.10}) \\ &= \sum_{j=1}^n (\mathcal{O}(j)(y_j - y_j^\delta))^2 = \left(\frac{\mathcal{O}(n)}{n}\right)^2 = (\mathcal{O}(1))^2. \end{aligned}$$

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

Exercício B.29. *Seja A um operador compacto. Considere o problema de determinar x na equação $Ax = y^\delta$. Você usaria a estratégia de minimizar o resíduo $\|Ax - y^\delta\|^2$ para solucionar o problema? Justifique adequadamente sua resposta. Sugestão: Olhe para as equações normais.*

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** [2, 3, 8].

Teorema B.2.11. *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:*

$$a) \ y \in \mathcal{R}(A),$$

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

Demonstração : $a) \Rightarrow b)$ De $y \in \mathcal{R}(A) \subset \overline{\mathcal{R}(A)}$. Seja $x \in \mathcal{H}_1$ tal que $Ax = y$. Segue do Teorema B.2.5 que $A^* f_j = \sigma_j e_j$. Assim,

usando a desigualdade de Bessel,

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

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

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

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

$$\|z\|^2 = \|y\|^2 - \sum_{j=0}^{\infty} |\langle y, f_j \rangle|^2 \langle z, f_j \rangle = 0, \quad \forall j \in \mathbb{N} \quad (\text{B.11})$$

$$\text{e} \quad A^* z = 0.$$

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

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

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

Exercício B.30. Complete os detalhes da demonstração do Teorema

B.2.11.

Observação: Note que o Teorema de Picard B.2.11, sugere que uma tentativa de solução para a equação $Ax = y$ é

$$x = \sum_{j=0}^{\infty} \sigma_j^{-1} \langle y, f_j \rangle e_j. \quad (\text{B.12})$$

Suponha que temos somente o dado perturbado $y^\delta = y + \delta f_k$. Substituindo na equação (B.12), obtemos como solução perturbada $x^\delta = x + \delta \sigma_k^{-1} e_k$. Assim,

$$\|x - x^\delta\|^2 = \left\| \frac{\delta}{\sigma_k} e_k \right\|^2 = \frac{\delta^2}{\sigma_k^2}. \quad (\text{B.13})$$

Como $\sigma_j \rightarrow 0$, segue que a estimativa (B.13) pode ser muito ruim, mostrando o efeito de mal-condicionamento causado pelos valores singulares de um operador compacto.

Apêndice C

O Conceito de Regularização

É um fato que, muitos dos *problemas inversos* são mal postos, no sentido que, a presença de ruídos nos dados levam a soluções inapropriadas para o problema. Os problemas de instabilidade podem ser controlados por meio de incorporação de informações *a priori*, por meio de transformações que estabilizam o problema, ou pelo uso apropriado de métodos numéricos. Essas técnicas são chamadas de **métodos de regularização**.

Desta forma, 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 por

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

enquanto). Em termos gerais, uma *estratégia ou método de regularização* consiste em aproximar uma solução x^\dagger (que pode ser a *solução generalizada* como a solução de norma mínima²) de um problema mal posto (1.1), por uma família (a um parâmetro α) de problemas bem postos.

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: de uma estratégia para aproximar o operador inverso A^{-1} de maneira a evitar o mal condicionamento; de uma regra para escolha do parâmetro 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. Além disso, depende também do conceito de solução que estamos considerando e da topologia em que esse conceito de solução está imerso (veja [2, 3, 8, 9, 14, 16, 21, 27, 52]).

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?

²Uma **solução de norma mínima** é, dentre todas as possíveis soluções do problema, a que possui menor norma. No caso de existirem infinitas, usa-se o infimo.

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

Para responder as questões (i) e (ii) acima, apresentamos nos Capítulos 2, 3 e 4, alguns métodos de regularização, divididos em duas classes:

Métodos iterativos de Regularização:

Sobre certas hipóteses, estratégias iterativas combinadas com uma maneira apropriada de escolher o critério de parada, implicam em maneiras estáveis de aproximar uma solução de (1.1).

Nessas notas, em especial nos Capítulos 2 e 3, estamos interessados em técnicas iterativas da forma

$$x_{k+1}^\delta = x_k^\delta + G_k(x_k^\delta, y^\delta), \quad k \in \mathbb{N} \quad (\text{C.1})$$

para várias escolhas de G_k .

Método de regularização de Tikhonov: É um dos métodos mais populares e bem conhecidos para a solução de problemas inversos. Este consiste em encontrar aproximações estáveis de uma solução de (1.1) minimizando o funcional

$$J_\alpha(x) = \|A(p)x - y^\delta\|^2 + \alpha\|x - x_0\|^2, \quad (\text{C.2})$$

onde x_0 é tipicamente alguma informação *a priori* da solução e $\alpha > 0$ é o chamado *parâmetro de regularização*. Trataremos da regularização de Tikhonov no Capítulo 4.

Para responder a (iii), ou seja, para assegurar que um método de regularização aplicado ao problema inverso (1.1) converge a uma solução e para expressar essa convergência em termos de taxas, são necessárias 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 [8, 2, 3, 14, 27].

Trataremos de responder (iii) no momento apropriado, durante os Capítulos 2, 3 e 4.

Definição C.0.12. *Sejam $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ um operador linear e limitado e $\alpha_0 \in (0, +\infty)$. Para todo $\alpha \in (0, \alpha_0)$, seja*

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

*um operador contínuo (não necessariamente linear). A família de operadores $\{R_\alpha\}$ é chamada de uma **regularização** ou **uma família de operador 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*

$$\limsup_{\delta \rightarrow 0} \{\|R_\alpha y^\delta - A^\dagger y\| : y^\delta \in \mathcal{H}_2, \|y - y^\delta\| \leq \delta\} = 0 \quad (\text{C.3})$$

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

Observação: Note 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 C.0.13. *Uma estratégia de regularização (R_α, α) é dita **convergente** se $x_\alpha := R_\alpha y$ converge para x^\dagger .*

C.1 Resultados de Convergência

Para responder as questões (i) e (ii) acima, são apresentados nos Capítulos 2, 3 e 4 alguns métodos de regularização, divididos em duas classes: iterativos (por exemplo o Método tipo gradiente e tipo Newton) e contínuos (por exemplo o Método de Tikhonov). Estes respondem, pelo menos em parte, as primeiras duas questões.

No nosso contexto, podemos dizer o seguinte:

Teorema C.1.1. *Seja $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ linear e limitado.*

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

$$\|x\| \leq C_1 \tau^{\frac{1}{2}}. \quad (\text{C.4})$$

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

$$\|x\| \leq C_2 \tau^{\frac{2}{3}}. \quad (\text{C.5})$$

Demonstração : *i)* De $x \in \mathcal{R}(A^*)$, 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 $x \in \mathcal{R}(A^*A)$, 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 .

Exercício C.1. Um operador B é dito ser semi-definido positivo se $\langle z, Bz \rangle \geq 0 \forall z \in \mathcal{D}(B)$. Prove que se A é um operador linear limitado entre espaços de Hilbert, então A^*A e AA^* são operadores lineares positivos.

Exemplo C.1.2 (Diferenciação nos dados). Considere o operador

$$\begin{aligned} A : L^2[0, 1] &\longrightarrow L^2[0, 1] \\ x &\longmapsto (Ax)(t) = \int_0^t x(s) ds = y(t), \quad t \in [0, 1]. \end{aligned} \quad (\text{C.6})$$

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:

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

Como, claramente, o operador adjunto é dado por

$$\begin{aligned} A^* : L^2[0, 1] &\longrightarrow L^2[0, 1] \\ y &\longmapsto (A^*y)(s) = - \int_s^1 y(r) dr. \end{aligned} \quad (\text{C.7})$$

Assim, se $\|Ax\| \leq \tau$, então

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

Exercício C.2. Prove que o operador adjunto do operador A definido pela equação (C.6) é o operador dado pela equação (C.7).

Exercício C.3. *Suponha A um operador linear limitado com $\mathcal{R}(A)$ 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 B.1.10)*

Observação: 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.

C.2 Escolha do Parâmetro de Regularização

Existem basicamente duas formas de escolha do parâmetro de regularização: *a-priori*, ($\alpha = \alpha(\delta)$) ou *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 C.2.1. *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

$$\limsup_{\delta \rightarrow 0} \{ \|R_{\alpha(y^\delta)} y^\delta - A^\dagger y\| : y^\delta \in \mathcal{H}_2, \|y - y^\delta\| \leq \delta \} = 0, \quad (\text{C.8})$$

e, portanto, $R_{\alpha(y)}y = A^\dagger y$ para todo $y \in \mathcal{D}(A^\dagger)$. Logo, de (C.8) segue que para qualquer sequência $\{y_n\} \in \mathcal{D}(A^\dagger)$ que converge para $y \in \mathcal{D}(A^\dagger)$,

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

e, assim, A^\dagger é limitado no $\mathcal{D}(A^\dagger)$. Pelo Teorema B.1.10 que $\mathcal{D}(A^\dagger) = \mathcal{H}_2$. ■

Exercício C.4. *Justifique a última afirmação na Proposição C.2.1.*

A Proposição C.2.1 nos ensina que uma escolha de parâmetros em um método de regularização convergente para problemas mal postos deve, obrigatoriamente, levar em conta o nível de ruídos δ .

C.2.1 Escolha *a priori*

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 C.2.2. *Para todo $\alpha > 0$, suponha que R_α um operador contínuo. Então a família R_α é uma regularização para A^\dagger se*

$$R_\alpha \longrightarrow A^\dagger \quad \text{pontualmente no } \mathcal{D}(A^\dagger) \quad \text{quando } \alpha \longrightarrow 0. \quad (\text{C.9})$$

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 : Fixe $y \in \mathcal{D}(A^\dagger)$ qualquer. Pela hipótese, existe um função monótona $\gamma : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ 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 $\lim_{\delta \rightarrow 0} \rho^{-1}(\delta) = 0$.

Defina $\alpha := \gamma(\rho^{-1}(\delta))$. Note que α é monotona 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. ■

Exercício C.5. *Mostre que, se uma família de operadores lineares converge uniformemente para um operador linear, então este limite é um operador contínuo. Use isto para justificar porquê não podemos requerer a convergência uniforme na Proposição C.2.2. Mostre que, no caso em que $\mathcal{R}(A)$ não é fechada, então $\|R_\alpha\| \xrightarrow{\alpha \rightarrow 0} +\infty$.*

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

Suponha que $\mathcal{R}(A)$ não é fechada, assim, A^\dagger é não limitada (pelo Teorema B.1.10). 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

$$\begin{aligned} \|R_{\alpha(\delta)}y^\delta - A^\dagger y\| &\leq \|x_{\alpha(\delta)} - A^\dagger y\| + \|R_{\alpha(\delta)}y^\delta - x_{\alpha(\delta)}\| \\ &= \|x_{\alpha(\delta)} - A^\dagger y\| + \|R_{\alpha(\delta)}y^\delta - R_{\alpha(\delta)}y\| \quad (\text{C.10}) \\ &\leq \|x_{\alpha(\delta)} - A^\dagger y\| + \delta \|R_{\alpha(\delta)}\|. \end{aligned}$$

Na estimativa (C.10) 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 C.5). A figura abaixo ilustra esta situação.

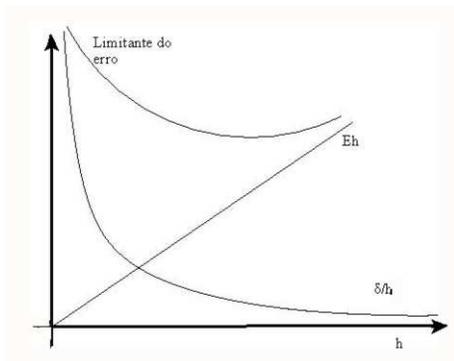


Figura C.1: Estimativa fundamental.

O caso em que temos igualdade na equação (C.10) é, sem sombra de dúvidas, o **pior caso**. 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 gostaríamos a solução do problema original.

Algumas técnicas de escolha *a priori* para o parâmetro α são bem conhecidas e amplamente usadas. Uma delas é chamada de **curva L**. Não entraremos em detalhes aqui. Para interessados sugerimos [3, 8].

Exercício C.6. *Faça uma comparação entre a estimativa (C.10) e a estimativa (1.7). Qual o papel de h em (1.7)?*

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 (C.10).

Uma maneira de construir uma família de regularização de maneira adequada será apresentada logo mais na Seção C.3 em termos dos valores espectrais. Essa técnica está baseada na construção das chamadas *funções filtro* [2, 3, 27] ou, neste caso em especial, nas funções de truncamento.

Da estimativa (C.10) segue que:

Proposição C.2.3. *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*

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

Exercício C.7. *Demonstre a Proposição C.2.3.*

C.2.2 Escolha *a posteriori*

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

$$\|Ax_\alpha^\delta - y^\delta\| \leq \tau\delta \quad (\text{C.11})$$

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

Uma motivação heurística para tal escolha é a seguinte:

Observação: 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 (C.10), 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* (C.11) seja satisfeita. Note 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 C.2.4. *Um método de regularização (R_α, α) , onde $\alpha := (\delta, y^\delta)$ é escolhido de acordo com o princípio da discrepância (C.11) é convergente $\forall y \in \mathcal{R}(A)$.*

Demonstração : Veja [8, Teorema 4.17]. ■

Métodos onde a escolha do parâmetro de regularização é feita *a posteriori* são os métodos iterativos estudados nos Capítulos 2 e 3.

C.3 Regularização por Truncamento dos Valores Singulares

Nesta seção 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 [3, 8, 46]. Para entrarmos em detalhes deste caso, teríamos que nos ater com mais fidelidade a Teoria Espectral [28] de operadores lineares e a uma versão do Teorema Espectral mais geral que o enunciado nessas notas. Interessados podem consultar [3, 8, 46] para detalhes.

Vamos ao setup do 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 $\mathcal{R}(A)$, 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 B.2.11, temos que uma possibilidade de solução seria

$$x^\delta = \sum_{j=1}^{\infty} \sigma_j^{-2} \langle A^*y^\delta, e_j \rangle e_j. \quad (\text{C.12})$$

Vimos na Observação B.2.2 que usar a equação (C.12) não é uma boa alternativa para calcular uma aproximação para a solução x^\dagger do problema inverso acima, haja visto 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,

$$\begin{aligned} x_\alpha^\delta &:= f_\alpha(A^*A)A^*y^\delta = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle A^*y^\delta, e_j \rangle e_j = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle y^\delta, Ae_j \rangle e_j \\ &= \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-2} \langle y^\delta, \sigma_j f_j \rangle e_j = \sum_{\substack{j=1 \\ \sigma_j^2 \geq \alpha}}^{\infty} \sigma_j^{-1} \langle y^\delta, f_j \rangle e_j. \end{aligned} \tag{C.13}$$

A definição de x_α^δ como na equação (C.13) pode ser vista como uma versão *truncada* da expansão em valores singulares (B.7).

Definição C.3.1. *O método dado pela equação (C.13) é chamado de expansão truncada em valores singulares.*

Exercício C.8. *Mostre que x_α^δ dado em (C.13) 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*

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

Observação: Note que, 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 .

Observação: 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 C.3.2. *O método de expansão truncada em valores singulares é um método de regularização.*

Demonstração: Basta notar que A_α satisfaz a Proposição C.2.2.

■

Exercício C.9. *Complete os detalhes da demonstração do Teorema C.3.2.*

Exercício C.10 (Taxas de convergência). *Seja A um operador linear compacto com sistema singular (σ_j, e_j, f_j) . Suponha que a solução x^\dagger de $Ax = y$ satisfaça a condição de fonte*

$$x^\dagger \in \mathcal{R}((A^*A)^\nu). \quad (\text{C.14})$$

i) Mostre que a condição de fonte (C.14) é satisfeita se e só se

$$\sum_{j=1}^{\infty} \frac{|\langle y, f_j \rangle|^2}{\sigma_j^{2+4\nu}} < \infty \quad (\text{C.15})$$

ii) Suponha que temos somente dados perturbados y^δ com $\|y - y^\delta\| \leq \delta$. Dê uma estimativa da taxa de convergência de x_α^δ para x^\dagger em função dos valores singulares.

Índice

- Abel
 - equação integral de , 30
- absorção, 21
- Alaoglu, 109
- Banach, 109
- bem-posto, 13
- Brauer, 83
- Calderón, 43
- Cauchy, 68, 82, 83
- Cauchy-Schwarz
 - desigualdade de, 26, 82
- coeficiente de absorção, 21
- coeficientes
 - de atenuação, 31
- condição de cone, 91
- condição do cone tangente, 55
- condições de fonte, 53
- Cormack, 43
- critério de parada, 50, 79
- critério de Picard, 136
- Cromak, 41
- curva L, 100
- diferenças simétricas, 17
- Dirichlet, 42
 - problema de, 35
- equação normal, 45, 87, 96, 121
- Erastóstenes, 42
- Eratostenes, 40
- erro
 - de aproximação, 50
 - nos dados, 50
- espaço
 - com produto interno, 112
 - de Banach, 111
 - de Hilbert, 112
 - de Schwartz, 114
 - pré-Hilbert, 111
 - vetorial normado, 111
- espaços vetoriais, 110
- espectro, 48
- estabilidade, 140
- estimativa de erro, 18
- estimativa fundamental, 50
- Fourier, 24
 - transformada de, 24
 - transformada inversa, 28
- Fréchet, 84
- Fréchet diferenciável, 69, 85
- função, 21
 - característica de um conjunto, 36
 - filtro, 98
 - mensurável e módulo integrável, 115

- mensurável e quadrado integrável, 115
 - potencial, 35
 - radial, 25
- funcional
 - coercivo, 100
 - de Tikhonov, 99, 104
- Gauss, 40, 42, 45, 82
- Gauss-Newton, 87, 92, 93
- Hadamard, 13, 42
- Hilbert, 12, 42
 - transformada de, 29
 - espaço de, 12
- hiperplano, 24
- Hounsfield, 41, 43
- identidade
 - de Parseval, 115
- integral
 - de linha, 23
 - de superfície, 24
- inversa generalizada, 117, 123
- iteração
 - ART, 77
 - de Landweber-Kaczmarz, 79
 - de Newton, 85
 - do gradiente, 70
- iteração de Landweber
 - iterativamente regularizada, 63
- Kaczmarz, 84
- Landweber, 47
 - iteração de: caso linear, 47
 - iteração de: caso não-linear, 54
- Levenberg, 94
- método
 - de regularização iterativo, 44
 - ART, 77
 - de Gauss-Newton, 87
 - de Gauss-Newton iterativamente regularizado, 92
 - de Landweber iterativamente regularizado, 63
 - de Landweber-Kaczmarz, 78
 - de Levenberg-Marquardt, 88
 - de máxima descida, 68
 - de mínimos quadrados, 45
 - de Newton, 45, 92
 - de Newton-Raphson, 92
 - de regularização de Landweber, 47
 - de regularização de Tikhonov, 95, 108
 - do gradiente, 68
 - do tipo gradiente, 45
 - passo de máxima descida, 68
 - steepest descent, 68
 - tipo Kaczmarz, 76
 - tipo Newton, 85, 86
- métodos de regularização, 139
- mal-posto, 13
- Marquardt, 94
- Maxwell
 - equações de, 34
- Neumann, 43
 - dados de, 36
- Newton, 45, 83
- norma, 110
- operador, 12
 - adjunto, 112
 - auto-adjunto, 113
 - compacto, 112
 - Dirichlet-to-Newmann, 37
 - integral, 119

- limitado, 112
- linear, 112
- Neumann-to-Dirichlet, 37
- projeção ortogonal, 113
- parâmetro
 - de regularização, 100, 104
- Platão, 40, 42
- ponto fixo
 - iteração de, 46
- precisão, 140
- princípio da discrepância, 51, 57, 64, 74, 89, 91
- problema
 - de condutividade inversa, 36
 - de condutividade inversa, 34
- problemas diretos, 10
- problemas inversos, 10
 - de identificação, 11
 - de reconstrução, 11
- produto interno, 111
- pseudo-inversa, 48, 117
- Radon, 23, 40, 43
 - transformada inversa de, 28
 - transformada de, 22
- Raphson, 93
- região de confiança, 88
- regularização
 - de Tikhonov, 95, 98, 103
- ruídos, 12
- Schwarz, 84
- sistema singular, 131
- sistemas de equações, 76
- solução
 - de mínimos quadrados com norma mínima, 48
 - de norma mínima, 140
 - de quadrados mínimos, 122
 - generalizada, 120
- Teorema
 - da Aplicação Espectral, 96, 133, 134
 - da aplicação espectral, 48
 - da decomposição em valores singulares, 117, 127
 - da diagonalização, 128
 - da Projeção, 113
 - de Alaoglu - ver Teorema de Banach-Alaoglu, 109
 - de Ascoli-Arzelá, 120
 - de Banach-Alaoglu, 114
 - de Fubini, 26
 - de Pitágoras, 121
 - do Gráfico Fechado, 73, 98
 - do gráfico fechado, 127
 - do ponto fixo de Brauer, 62
 - espectral, 119, 130
- termo de penalização, 100
- Tikhonov, 108
- tomografia, 19
 - acústica, 19
 - axial computadorizada, 43
 - computadorizada foto-acústica, 80
 - elétrica por impedância, 19, 34
 - por raio-X, 19
 - axial computadorizada, 20
 - caso discreto, 31
 - problema inverso, 22
- transformada
 - de Fourier, 114
 - inversa de Fourier, 115
- valores espectrais, 131

Bibliografia

- [1] G. Anger, *Inverse problems in differential equations*, Plenum Press, New York, 1990.
- [2] J. Baumeister, *Stable solution of inverse problems*, Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1987.
- [3] J. Baumeister and A. Leitão, *Topics in inverse problems*, Publicações Matemáticas do IMPA. [IMPA Mathematical Publications], Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2005, , 25° Colóquio Brasileiro de Matemática. [25th Brazilian Mathematics Colloquium].
- [4] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*, Institute of Physics Publishing, Bristol, 1998.
- [5] T.M. Buzug, *Computer Tomography - From Photom Statistics to Modern Cone-Beam CT*, Springer, Berlin Heidelberg, 2008.
- [6] Alberto P. Calderón, *On an inverse boundary value problem*, Comput. Appl. Math. **25** (2006), no. 2-3, 133–138.
- [7] A. De Cezaro and A. Leitão, *Introdução aos Problemas Inversos Lineares*, Lectures Notes: IV Bienal da So-

- cidade Brasileira de Matemática - Maringá - PR, online: <http://www.dma.uem.br/bienalsbm>, 2008, pp. 1–60.
- [8] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
- [9] Heinz W. Engl, *Inverse problems*, Aportaciones Matemáticas: Textos [Mathematical Contributions: Texts], vol. 8, Sociedad Matemática Mexicana, México, 1995, With an introduction in Spanish by Roberto Martínez Villa.
- [10] C. L. Epstein, *Introduction to the Mathematics of Medical Imaging*, Pearson Prentice Hall, Upper Saddle River, NJ, 2003.
- [11] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.
- [12] G. H. Golub and C. F. Van Loan, *Matrix computations*, third ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996.
- [13] C. W. Groetsch, *Generalized inverses of linear operators: representation and approximation*, Marcel Dekker Inc., New York, 1977, Monographs and Textbooks in Pure and Applied Mathematics, No. 37.
- [14] ———, *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Research Notes in Mathematics, vol. 105, Pitman (Advanced Publishing Program), Boston, MA, 1984.

- [15] ———, *Inverse problems in the mathematical sciences*, Vieweg Mathematics for Scientists and Engineers, Friedr. Vieweg & Sohn, Braunschweig, 1993.
- [16] ———, *Inverse problems*, Mathematical Association of America, Washington, DC, 1999, Activities for undergraduates.
- [17] ———, *Stable approximate evaluation of unbounded operators*, Lecture Notes in Mathematics, vol. 1894, Springer-Verlag, Berlin, 2007.
- [18] F. A. Grünbaum, *Some mathematical problems suggested by limited angle tomography*, Inverse problems (New York, 1983), SIAM-AMS Proc., vol. 14, Amer. Math. Soc., Providence, RI, 1984, pp. 65–77.
- [19] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Dover Publications, New York, 1953.
- [20] P. C. Hansen, *Rank-deficient and discrete ill-posed problems*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998, Numerical aspects of linear inversion.
- [21] B. Hofmann, *Regularization for applied inverse and ill-posed problems*, Teubner-Texte zur Mathematik [Teubner Texts in Mathematics], vol. 85, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1986, A numerical approach, With German, French and Russian summaries.
- [22] V. Isakov, *Inverse source problems*, Mathematical Surveys and Monographs, vol. 34, American Mathematical Society, Providence, RI, 1990.

- [23] ———, *Inverse problems for partial differential equations*, second ed., Applied Mathematical Sciences, vol. 127, Springer, New York, 2006.
- [24] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, Applied Mathematical Sciences, vol. 160, Springer-Verlag, New York, 2005.
- [25] B. Kaltenbacher, A. Neubauer, and 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.
- [26] J. B. Keller, *Inverse problems*, Amer. Math. Monthly **83** (1976), no. 2, 107–118.
- [27] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.
- [28] E. Kreyszig, *Introductory functional analysis with applications*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1989.
- [29] L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math. **73** (1951), 615–624.
- [30] M. M. Lavrentiev, A. V. Avdeev, M. M. Lavrentiev, Jr., and V. I. Priimenko, *Inverse problems of mathematical physics*, Inverse and Ill-posed Problems Series, VSP, Utrecht, 2003.
- [31] E. L. Lima, *Curso de análise. Vol. 1*, Projeto Euclides [Euclid Project], vol. 1, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1976.

- [32] ———, *Espaços métricos*, Projeto Euclides [Euclid Project], vol. 4, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1977.
- [33] ———, *Curso de análise. Vol. 2*, Projeto Euclides [Euclid Project], vol. 13, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1981.
- [34] A. K. Louis, *Inverse problems in medicine*, Applications of mathematics in industry and technology (Siena, 1988), Teubner, Stuttgart, 1989, pp. 277–287.
- [35] ———, *Application of the approximate inverse to 3D X-ray CT and ultrasound tomography*, Inverse problems in medical imaging and nondestructive testing (Oberwolfach, 1996), Springer, Vienna, 1997, pp. 120–133.
- [36] C. Meyer, *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000, With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.).
- [37] F. Natterer, *The mathematics of computerized tomography*, Classics in Applied Mathematics, vol. 32, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001, Reprint of the 1986 original.
- [38] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.
- [39] Platão, *A república: Coleção a obra - prima de cada autor*, Martin Claret, São Paulo SP, 2002.

- [40] A. I. Prilepko, D. G. Orlovsky, and I. A. Vasin, *Methods for solving inverse problems in mathematical physics*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 231, Marcel Dekker Inc., New York, 2000.
- [41] J. Radon, *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten [on the determination of functions by their integral values along certain manifolds]*, 75 years of Radon transform (Vienna, 1992), Conf. Proc. Lecture Notes Math. Phys., IV, Int. Press, Cambridge, MA, 1994, pp. 324–339.
- [42] A. G. Ramm and A. I. Katsevich, *The Radon transform and local tomography*, CRC Press, Boca Raton, FL, 1996.
- [43] W. T. Reid, *Generalized inverses of differential and integral operators*, Proc. Sympos. Theory and Application of Generalized Inverses of Matrices (Lubbock, Texas, 1968), Texas Tech. Press, Lubbock, Tex., 1968, pp. 1–25.
- [44] A. Sá Barreto, *Introdução as transformadas de Radon*, Lectures Notes: Symposium on Spectral and Scattering Theory, realizado em Recife e Serrambi, de 11 a 22 de Agosto de 2003, online: <http://www.math.purdue.edu/sabarre/papers.html>, 2003, pp. 1–22.
- [45] M. Salo, *Calderón problem*, Lectures Notes, Springer, 2008, pp. 1–59.
- [46] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*, Applied Mathematical Sciences, vol. 167, Springer, New York, 2009.

- [47] T. Schuster, *The method of approximate inverse: theory and applications*, Lecture Notes in Mathematics, vol. 1906, Springer, Berlin, 2007.
- [48] G. Strang, *Linear algebra and its applications*, second ed., Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1980.
- [49] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.
- [50] G. Uhlmann, *Electrical impedance tomography and Calderón's problems*, Lectures Notes: CIMPA - Summer School on Inverse Problems and Applications- Santiago - Chile, Jan. 2010, online: <http://indico.cmm.uchile.cl>, 2010, pp. 1–42.
- [51] C. R. Vogel, *Computational methods for inverse problems*, Frontiers in Applied Mathematics, vol. 23, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, With a foreword by H. T. Banks.
- [52] J. P. Zubelli, *An introduction to inverse problems. Examples, methods and questions*, 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium], Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999, Appendix A by the author and Luis Orlando Castellano Pérez.