



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA

Joel Conceição Rabelo

**MÉTODOS TIPO TIKHONOV COM MINIMIZAÇÃO INEXATA E O MÉTODO
TIKHONOV KACZMARZ PARA RESOLVER SISTEMAS DE EQUAÇÕES
LINEARES MAL POSTOS**

Florianópolis
2020

Joel Conceição Rabelo

**MÉTODOS TIPO TIKHONOV COM MINIMIZAÇÃO INEXATA E O MÉTODO
TIKHONOV KACZMARZ PARA RESOLVER SISTEMAS DE EQUAÇÕES
LINEARES MAL POSTOS**

Tese submetida ao Programa de Pós-Graduação
em Matemática Aplicada da Universidade Federal
de Santa Catarina para a obtenção do título de dou-
tor em Matemática Aplicada.

Orientador: Prof. Dr. Antônio Carlos Gardel Leitão
Coorientador: Prof. Dr. Fábio Júnior Margotti

Florianópolis
2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Rabelo, Joel Conceição

Métodos tipo Tikhonov com minimização inexata e o método Tikhonov Kaczmarz para resolver sistemas de equações lineares mal postos / Joel Conceição Rabelo ; orientador, Antônio Carlos Gardel Leitão, coorientador, Fábio Júnior Margotti, 2021.

72 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Ciências Físicas e Matemáticas, Programa de Pós-Graduação em Matemática Pura e Aplicada, Florianópolis, 2021.

Inclui referências.

1. Matemática Pura e Aplicada. 2. Problemas mal postos. . 3. Método de Tikhonov Iterado. . 4. Minimização inexata. . 5. Método Kaczmarz.. I. Leitão, Antônio Carlos Gardel . II. Margotti, Fábio Júnior. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Matemática Pura e Aplicada. IV. Título.

Joel Conceição Rabelo

**MÉTODOS TIPO TIKHONOV COM MINIMIZAÇÃO INEXATA E O MÉTODO
TIKHONOV KACZMARZ PARA RESOLVER SISTEMAS DE EQUAÇÕES
LINEARES MAL POSTOS**

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Dr. Fábio Júnior Margotti
Universidade Federal de Santa Catarina- UFSC

Prof. Dr. Helcio Rangel Barreto Orlande
Universidade Federal do Rio de Janeiro-UFRJ

Prof. Dr. Haroldo Fraga de Campos Velho
Instituto Nacional de Pesquisas Espaciais - INPE

Prof. Dr. Uri M. Ascher
The University of British Columbia

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Matemática Aplicada.

Coordenação do Programa de
Pós-Graduação

Prof. Dr. Antônio Carlos Gardel Leitão
Orientador

Florianópolis, 2020.

Este trabalho é dedicado aos meus queridos pais,
Raimundo Nonato Lima Rabelo e Maria de Jesus
Conceição Rabelo(In Memoriam).

AGRADECIMENTOS

- A Deus, pela força e permissão concedida para a realização deste trabalho.
- A minha família, meus pais e a meus irmãos Iolanda, Orlando, Manoel, Maria e Jesmo pelo apoio e motivação para chegar até aqui.
- A minha esposa Renata pelo apoio dado durante essa caminhada.
- Aos professores, F. Margotti e A. Leitão por terem orientado este trabalho, pelo apoio, pelos conselhos dados no desenvolvimento deste trabalho e pelas boas conversas.
- A banca examinadora composta pelos professores Uri Ascher, Haroldo Fraga Velho, Helcio Rangel por terem aceito participar da avaliação deste trabalho e pelas devidas sugestões.
- A CAPES.

RESUMO

Na primeira parte deste trabalho, propomos e estudamos as propriedades de três distintos algoritmos para obter soluções aproximadas estáveis para sistemas de equações mal postas modeladas por operadores lineares atuando entre espaços de Hilbert. Desenvolvemos uma versão inexata do método de Tikhonov de um passo penalizado com funcional uniformemente convexo, bem como uma versão inexata para os métodos de Tikhonov iterado. Para o caso de métodos de um passo, propomos dois algoritmos distintos, um baseado em escolha *a priori* e outro baseado na escolha *a posteriori* do parâmetro de regularização. As propriedades de convergência e estabilidade são provadas, bem como taxas ótimas de convergência (sob condição de fonte apropriada). Estabelecemos ainda uma variante do método de Tikhonov iterado com a escolha *a posteriori* da sequência dos parâmetros de penalização. Para este algoritmo, provamos estabilidade para dados com ruídos e a propriedade de regularização.

Na segunda parte deste trabalho propomos e analisamos um método tipo Tikhonov Iterado Kaczmarz relaxado não estacionário (TIKr). Generalizamos para o TIKr critérios propostos em [6] para o método iterado de Tikhonov (iT). O objetivo é obter uma estratégia eficiente para a escolha dos multiplicadores de Lagrange neste método. Estabelecemos para o método TIKr estabilidade e regularização. Apresentamos modelos de aplicações para duas aplicações, a saber: um problema de desfoque de imagem e o problema do potencial inverso. Os resultados numéricos obtidos validam a eficiência do método proposto.

Palavras-chave: Problemas mal postos. Método de Tikhonov Iterado. Minimização inexata. Otimização Convexa. Distância de Bregman. Método Kaczmarz.

ABSTRACT

In the first part of this work, we propose and study the properties of three distinct algorithms to obtain stable approximate solutions for ill-posed equations modeled by linear operators acting between Hilbert spaces. We have developed an inexact version of the Tikhonov method penalized by a uniformly convex functional, as well as an inexact version for the iterated Tikhonov method. For one-step methods, we propose two distinct algorithms, one based on *a priori* choice and another based on *a posteriori* choice of the regularization parameters. Convergence and stability properties are established, as well as optimal convergence rates (under appropriate source conditions). We also established a variant of the iterated Tikhonov method with a posteriori choice of the sequence of regularization parameters. For this last algorithm, we prove stability for noisy data and the regularization property.

In the second part of this work we propose and analyze a Kaczmarz version of the relaxed nonstationary iterated-Tikhonov method (TIKr). We generalized for TIKr the criteria proposed in [6] for the iterated Tikhonov (iT) method. The goal is obtain an efficient strategy for choosing Lagrange multipliers in this method. We established for the TIKr method stability and regularization properties. We present application models for two applications, namely: an image blurring problem and the inverse potential problem. The numerical results obtained validate the efficiency of the proposed method.

Keywords: Ill-posed problems. Iterated-Tikhonov method. Inexact minimization. Convex optimization. Bregman distances. Kaczmarz method.

SUMÁRIO

1	INTRODUÇÃO	9
2	DEFINIÇÕES E RESULTADOS BÁSICOS	14
2.1	CONVEXIDADE E DISTÂNCIA DE BREGMAN	14
2.2	O MÉTODO HÍBRIDO PROXIMAL EXTRAGRADIENTE	17
2.3	RESULTADOS PRELIMINARES	19
3	MÉTODOS DE TIKHONOV DE UM PASSO COM MINIMIZAÇÃO INEXATA	25
3.1	REGRA A PRIORI DA ESCOLHA DOS PARÂMETROS DE REGULARIZAÇÃO	26
3.2	REGRA A POSTERIORI DA ESCOLHA DOS PARÂMETROS DE REGULARIZAÇÃO	26
4	MÉTODO DE TIKHONOV ITERADO COM MINIMIZAÇÃO INEXATA	30
4.1	APRESENTAÇÃO DO MÉTODO	30
4.2	ANÁLISE DE CONVERGÊNCIA	32
4.3	OPÇÕES DE MÉTODOS PARA A ITERAÇÃO INTERNA	41
5	MÉTODO DE TIKHONOV ITERADO KACZMARZ	44
5.1	APRESENTAÇÃO DO MÉTODO	44
5.2	ANÁLISE DE CONVERGÊNCIA	50
6	CONCLUSÕES E TRABALHOS FUTUROS	57
	APÊNDICE A – APLICAÇÕES	60
A.1	DESFOQUE DE IMAGEM	60
A.2	PROBLEMA DO POTENCIAL INVERSO	62
	BIBLIOGRAFIA	65

1 INTRODUÇÃO

Nesse trabalho apresentamos quatro métodos: dois métodos Tikhonov de um passo com minimização inexata, o método Tikhonov iterado com minimização inexata e uma versão relaxada do método de Tikhonov iterado Kaczmarz.

Começamos esta seção apresentando alguns exemplos básicos de problemas inversos, para mais detalhes e exemplos veja [17]. Em seguida, estabeleceremos quando um problema inverso é mal posto e, por fim, discutiremos um pouco sobre os métodos presentes neste trabalho.

Exemplo 1. *Encontrar um polinômio p de grau n com raízes x_1, \dots, x_n . Esse é um problema inverso do problema direto: Encontrar as raízes x_1, \dots, x_n de um certo polinômio p . Esse problema inverso admite infinitas soluções.*

Exemplo 2. *Dada uma matriz real e simétrica $A_{n \times n}$ e n números reais $\lambda_1, \dots, \lambda_n$ encontre uma matriz $D_{n \times n}$ tal que $A + D$ tenha autovalores $\lambda_1, \dots, \lambda_n$. Esse é um problema inverso do problema direto: Computar os autovalores da matriz $A + D$.*

Segundo Hadamard [64], um problema é bem posto quando:

1. O problema admite uma solução (existência);
2. O problema admite no máximo uma solução (unicidade);
3. A solução depende continuamente dos dados (estabilidade).

A existência e a unicidade dependem apenas da natureza algébrica dos espaços e do operador envolvido, i.e, se é injetivo ou sobrejetivo, enquanto que a estabilidade depende da topologia dos espaços, i.e, se o inverso do operador é contínuo. Os problemas inversos para os quais pelo menos um dos itens acima não acontece são chamados de mal postos.

O problema que desejamos resolver consiste em determinar um ponto desconhecido $x \in X$ a partir de um dado $b \in Y$, onde X e Y são espaços de Hilbert. Vamos assumir que os dados são obtidos por uma medida indireta descrita por uma equação mal posta, a saber

$$Ax = b, \tag{1}$$

onde $A : X \rightarrow Y$ é um operador linear limitado, sendo que a inversa $A^{-1} : \mathcal{R}(A) \rightarrow X$ não existe ou, caso exista, é descontínua. Vamos assumir ainda que na prática não seja conhecido o valor exato de b e sim apenas uma aproximação $b^\delta \in Y$, satisfazendo

$$\|b - b^\delta\| \leq \delta, \tag{2}$$

onde $\delta > 0$, o nível de ruído (erro cometido nas medidas) é conhecido.

A principal dificuldade em resolver o problema (1) está em obter para cada parâmetro $\alpha > 0$ um operador linear e limitado $R_\alpha : Y \rightarrow X$ de forma que o operador $R_\alpha A$ convirja pontualmente para a identidade quando $\alpha \rightarrow 0$.

Uma maneira de se obter esses operadores é a *regularização de Tikhonov*, uma ferramenta bastante conhecida para produzir soluções estáveis do problema inverso usando (2). A formulação mais clássica do método de Tikhonov para resolver (1) consiste em encontrar um minimizador $x_\alpha \in X$ do funcional convexo

$$T_\alpha(x) = \frac{1}{2} \|Ax - b^\delta\|^2 + \alpha \frac{1}{2} \|x\|^2, \quad (3)$$

onde o parâmetro $\alpha > 0$ é responsável por controlar o nível de regularização. Este parâmetro deve ser escolhido em função do nível de ruído δ , i.e, $\alpha = \alpha(\delta)$. Neste caso, dizemos que a escolha do parâmetro é *a priori*; se, além disso, α depende também de b^δ então dizemos que a escolha do parâmetro é *a posteriori*.

Uma estratégia *a priori* dos parâmetros, por exemplo, consiste em escolher $\alpha = \alpha(\delta)$ de modo que

$$\alpha \rightarrow 0 \text{ e } \frac{\delta^2}{\alpha} \rightarrow 0 \text{ quando } \delta \rightarrow 0. \quad (4)$$

Com essa regra, é possível provar (veja seção 2.3 e [9]) a propriedade de regularização

$$x_{\alpha(\delta)} \rightarrow x^\dagger \text{ quando } \delta \rightarrow 0, \quad (5)$$

onde $x^\dagger \in s(b) := \{x \in X : Ax = b\}$ é chamado de *solução de norma mínima*, caracterizado por

$$\|x^\dagger\| \leq \|x\|, \text{ para todo } x \in s(b).$$

Além disso, sob a condição de fonte $x^\dagger \in R(A^*)$ e escolhendo $\alpha = \alpha(\delta)$ é possível obter a taxa de convergência

$$\|x_{\alpha(\delta)} - x^\dagger\| = O(\delta^{1/2}). \quad (6)$$

Sob a condição de fonte $x^\dagger \in R(A^*)$ a taxa de convergência dada em (6) é considerada ótima, segundo a Definição 3.17 de [9]. Não menos importante, uma escolha *a posteriori* dos parâmetros pode ser feita de modo que

$$c_1 \delta \leq \|Ax_{\alpha(\delta)} - b^\delta\| \leq c_2 \delta, \quad (7)$$

com $1 < c_1 \leq c_2$ constantes fixadas. A escolha *a posteriori* implica em (5), veja a seção 2.3 e [17]. E ainda, sob a condição de fonte $x^\dagger \in R(A^*)$, obtemos a taxa de convergência (6).

Nas últimas décadas, a teoria da regularização de Tikhonov tem sido bastante aprimorada. Entre as modificações mais importantes estão a substituição do funcional

de penalização $P(x) = \frac{1}{2} \|x\|^2$ em (3) por um funcional convexo mais geral $P : X \rightarrow \overline{\mathbb{R}}$. Essa é uma teoria bem desenvolvida e a maioria dos resultados são bem conhecidos, consulte a seção 2.3. Outra variação consiste no uso de regras para a escolha do parâmetro α , veja [10, 27, 17]. Para resultados sobre a troca dos espaços de Hilbert X e Y por espaços de Banach e aplicações em problemas inversos não lineares consulte [58].

Os métodos do tipo Tikhonov apresentados acima são conhecidos como métodos de um passo, isso porque o minimizador x_α do funcional (3) é usado como uma aproximação de uma solução do problema inverso (1). Em contrapartida o *método de Tikhonov iterado*, na sua formulação mais clássica, é definido pelo procedimento de minimizar sucessivamente os funcionais de Tikhonov dados por

$$T_k(x) = \frac{1}{2} \|Ax - b^\delta\|^2 + \alpha_k \frac{1}{2} \|x - x_k\|^2, \quad (8)$$

em que a iteração do método é dada por $x_{k+1} := \arg \min_{x \in X} T_k(x)$, começando a partir de um ponto inicial $x_0 \in X$. A iteração deste método termina num índice $k_\delta \in \mathbb{N}$. O mais importante nos métodos de Tikhonov é a propriedade de regularização (5), que pode ser provada sob condições exigidas na sequência de parâmetros α_k .

A escolha da sequência de parâmetros α_k é de grande importância na performance do método de Tikhonov iterado. Por exemplo, para um cenário *a priori* podemos escolher a sequência constante $\alpha_k = c$ (método Tikhonov iterado estacionário), a sequência geométrica $\alpha_k = \alpha_0 r^k$, com $0 < r < 1$ (método Tikhonov iterado geométrico) ou uma sequência α_k satisfazendo $\sum_{k=0}^{\infty} \alpha_k^{-1} = \infty$, veja [28, 27, 12, 7, 13]. O desempenho deste método melhora significativamente se uma escolha *a posteriori* da sequência α_k é usada, como feito em [6, 19, 21], onde α_k é computado *a posteriori* de forma que

$$(1 - \eta_0) \delta + \eta_0 \|Ax_k - b^\delta\| \leq \|Ax_{k+1} - b^\delta\| \leq (1 - \eta_1) \delta + \eta_1 \|Ax_k - b^\delta\|, \quad (9)$$

com $0 \leq \eta_0 \leq \eta_1 \leq 1$.

Métodos onde o α_k é computado inspirado em (9) são chamados de métodos com critério de *intervalo relaxado* (*range-relaxed*). Esse critério foi introduzido em [6] e cada passo do método é calculado projetando a iteração anterior sobre um conjunto convexo fechado apropriado. Convergência fraca foi provada em [6] no caso de dados sem ruídos, $\delta = 0$. Extensões de [6] para operadores lineares em espaços de Banach podem ser encontradas em [19], incluindo convergência forte no caso sem ruídos e também a propriedade de regularização no caso de dados com ruídos. A análise do problema não linear sob o critério intervalo relaxado é abordada em [20]. O critério de intervalo relaxado será usado nos métodos iterativos deste trabalho, veja Capítulos 4 e 5.

Na tentativa de diminuir o esforço computacional no cálculo do minimizador exato do funcional de Tikhonov, propomos, nesse trabalho, a adaptação de quatro métodos de regularização. Os três primeiros métodos consistem em versões inexatas do método de Tikhonov de um passo (com escolha dos parâmetros dados *a priori* por (4) e *a posteriori* por (7)) e Tikhonov iterado com a estratégia do intervalo relaxado (9). Todos esses métodos geram aproximações estáveis para o problema (1). O artigo que contém a análise desses métodos citados anteriormente, [1], foi publicado na revista *Inverse Problems*. A vantagem está no esforço exigido em computar, em cada passo, apenas uma aproximação do minimizador ao invés de calcular o próprio minimizador exato. A precisão das aproximações dos minimizadores deve ser cuidadosamente controlada, a fim de manter a convergência desejada e resultados de estabilidade do método original de Tikhonov. Em nossa análise, exigimos apenas uma propriedade genérica do método usado na iteração interna, o qual aproxima o minimizador exato.

No quarto e último método deste trabalho estudamos um método tipo *Tikhonov Iterado Kaczmarz relaxado* não estacionário (TIKr) para obter aproximações de soluções regularizadas de sistemas de equações lineares mal postas, o mesmo foi realizado em colaboração com A. Leitão, R. Boiger e R. Filippozzi, [2]. Este é um método do tipo Kaczmarz [54], que tem como objetivo resolver um problema tipo (1), porém, para N blocos, com $N \geq 1$. Cada passo é definido como no método de Tikhonov iterado [37, Sec. 1.2] que é aplicado de maneira cíclica. O multiplicador de Lagrange é escolhido de forma que a próxima iteração tenha resíduo em um certo *intervalo*, similarmente ao que foi feito em [6].

Nesse cenário o problema inverso que nos interessa consiste em descobrir um quantidade $x \in X$ a partir de um conjunto de dados $(b_0, \dots, b_{N-1}) \in Y^N$, onde X e Y são espaços de Hilbert e $N \geq 1$. Na prática conhecemos apenas $b_i^\delta \in Y$ satisfazendo

$$\|b_i^\delta - b_i\| \leq \delta_i, \quad i = 0, \dots, N-1, \quad (10)$$

onde $\delta_i > 0$ é o nível de ruído (conhecido). Os dados disponíveis b_i^δ são obtidos através de medidas indiretas no parâmetro x . Dessa forma o processo é descrito pelo sistema de equações mal postas

$$A_i x = b_i, \quad i = 0, \dots, N-1, \quad (11)$$

onde $A_i : X \rightarrow Y$ é um operador linear limitado, cuja inversa $A_i^{-1} : R(A_i) \rightarrow X$ não existe ou é descontínua. Existe uma vasta literatura que trata do problema (11), por exemplo [46, 52, 34, 56, 57, 33, 17, 9, 60, 16].

As principais contribuições deste trabalho são:

- Propor e analisar versões inexatas do método de Tikhonov de um passo com critérios *a priori* e *a posteriori*, bem como estabelecer taxas de convergência segundo condições de fonte apropriadas, veja Capítulo 3.

- Propor e analisar uma versão inexata do método de Tikhonov iterado com critério *intervalo relaxado*, provando estabilidade e regularização do método, veja Capítulo 4.
- Propor e analisar uma versão do método Tikhonov iterado Kaczmarz com critério *intervalo relaxado*, provando convergência para dados exatos ($\delta = 0$), estabilidade e regularização do método no caso de dados com ruídos, veja Capítulo 5.

Vale ressaltar que os resultados apresentados neste trabalho, frutos de [1] e [2], são originais e que os mesmos foram aceitos para publicação na revista *Inverse Problems*.

Este trabalho está organizado da seguinte maneira: no Capítulo 2 apresentaremos definições e resultados básicos que darão suporte a este trabalho, bem como resultados já conhecidos na literatura sobre Métodos de Tikhonov de um passo. No Capítulo 3 apresentaremos nossa variante do Método de Tikhonov de um passo, a saber o Método de Tikhonov de um passo com minimização inexata. Fazemos a análise de convergência e estabelecemos ainda, sob condições adequadas, taxas de convergência tanto para a escolha *a priori* dos parâmetros de regularização quanto para a escolha *a posteriori*. No Capítulo 4 além de apresentar uma variante para o Método de Tikhonov iterado, estabelecemos a análise de convergência do método proposto e citamos exemplos que se encaixam na estrutura do nosso algoritmo. No Capítulo 5 abordaremos uma variante do Método Kaczmarz, nele vamos estabelecer um critério relaxado para a escolha dos multiplicadores de Lagrange, bem como a análise de convergência do método. Em seguida no Capítulo 6, estabelecemos as conclusões e citamos possíveis extensões dos métodos analisados neste trabalhos. Por fim, no Apêndice A fornecemos o modelo de duas aplicações (o problema de imagem desfocada e o problema do potencial inverso) para o método apresentado no Capítulo 5.

2 DEFINIÇÕES E RESULTADOS BÁSICOS

Nesta seção, apresentamos alguns resultados básicos já existentes na literatura que darão suporte teórico a este trabalho. Vamos apresentar resultados de análise convexa (veja [5]) e falar um pouco sobre o método híbrido extragradiente proximal (HPE), introduzido por [30], método este que serviu de piloto nas análises teóricas deste trabalho.

2.1 CONVEXIDADE E DISTÂNCIA DE BREGMAN

Neste trabalho X denota um espaço de Hilbert real com produto interno $\langle \cdot, \cdot \rangle$. A norma associada vamos denotar por $\|\cdot\|$. Seja $f : X \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ uma função, o *domínio efetivo* de f é o conjunto $\text{dom}(f) = \{x \in X : f(x) < \infty\}$. A função f é dita *própria* se $\text{dom}(f) \neq \emptyset$.

Definição 2.1.1. *Seja $f : X \rightarrow \overline{\mathbb{R}}$ uma função própria. Então f é uniformemente convexa se existe uma função contínua e estritamente crescente $\varphi : [0, \infty) \rightarrow [0, \infty)$ tal que $\varphi(0) = 0$ e*

$$f(\lambda x + (1 - \lambda)y) + \lambda(1 - \lambda)\varphi(\|x - y\|) \leq \lambda f(x) + (1 - \lambda)f(y), \quad (12)$$

para todo $\lambda \in (0, 1)$ e para todo $x, y \in \text{dom}(f)$. A função φ é chamada *módulo de convexidade de f* .

Se na Definição 2.1.1, φ é identicamente nula então f é dita convexa. Se $\varphi(\cdot) = (\beta/2)(\cdot)^2$ para algum $\beta > 0$, então f é dita *fortemente convexa* com constante β . Note ainda que f uniformemente convexa implica que f é estritamente convexa (quando em (12) $\varphi \equiv 0$ e a desigualdade é estrita) que por sua vez implica que f é convexa. Dizemos que f é *semicontínua inferiormente* (s.c.i) se para alguma sequência x_k em X tal que $x_k \rightarrow x$, a desigualdade

$$f(x) \leq \liminf f(x_k),$$

é verificada. Além disso, f é dita *fracamente semicontínua inferiormente* (f.s.c.i) se a propriedade acima é verdade quando trocamos $x_k \rightarrow x$, por $x_k \rightharpoonup x$. Essa duas noções de semicontinuidade são equivalentes para funções convexas, veja [5, Teorema 9.1].

Definição 2.1.2. *Seja $f : X \rightarrow \overline{\mathbb{R}}$ uma função própria. O subdiferencial de f é um operador ponto conjunto $\partial f : X \rightarrow 2^X$ definido por*

$$\partial f(x) := \{u \in X : f(x) + \langle u, y - x \rangle \leq f(y), \text{ para todo } y \in X\}.$$

O *domínio efetivo* do ∂f é o conjunto $\text{dom}(\partial f) := \{x \in X : \partial f(x) \neq \emptyset\}$.

Quando $\partial f(x) \neq \emptyset$, f é dita subdiferenciável em x . Podemos ver ainda que $\text{dom}(\partial f) \subset \text{dom}(f)$, pois como f é própria $f(x) = \infty$ implica que $\partial f(x) = \emptyset$. Segue ainda da definição de subdiferencial a condição necessária e suficiente de minimalidade

$$0 \in \partial f(x) \iff f(x) \leq f(y), \text{ para todo } y \in X. \quad (13)$$

Se $f, g : X \rightarrow \overline{\mathbb{R}}$ são duas funções convexas, próprias e s.c.i e existe um elemento $x \in \text{dom}(f) \cap \text{dom}(g)$ de forma que f ou g seja contínua em x , então

$$\partial(f + g)(x) = \partial f(x) + \partial g(x), \text{ para todo } x \in X.$$

Para funções reais, no caso em que f é diferenciável, $\partial f(x) = \{\nabla f(x)\}$. Além do mais, ∂f é contínuo se f é Fréchet diferenciável.

Definição 2.1.3. *Sejam $f : X \rightarrow \overline{\mathbb{R}}$ uma função convexa, própria e s.c.i, $x, y \in X$ com $y \in \text{dom}(\partial f)$ e $v \in \partial f(y)$. Com relação a f vamos definir a distância de Bregman entre x e y na direção de v como sendo*

$$D_v(x, y) := f(x) - f(y) - \langle v, x - y \rangle.$$

Segue direto da definição de subdiferencial que $D_v(x, y) \geq 0$. A função $D_v(\cdot, y)$ é convexa, e se f é estritamente convexa a função $D_v(\cdot, y)$ também é estritamente convexa. Se $x \in \text{dom}(\partial f)$ e f é s.c.i, então $D_v(\cdot, y)$ é sempre subdiferenciável neste ponto e

$$\partial D_v(\cdot, y)(x) = \partial f(x) - v. \quad (14)$$

O lema a seguir traz a chamada *identidade dos três pontos*, sua prova não é difícil e será de grande valor na análise do nosso método.

Lema 2.1.4. *Sejam $f : X \rightarrow \overline{\mathbb{R}}$ convexa, $x \in X$, $x_1, x_2 \in \text{dom}(\partial f)$, $\xi_1 \in \partial f(x_1)$ e $\xi_2 \in \partial f(x_2)$. Então*

$$D_{\xi_1}(x, x_1) - D_{\xi_2}(x, x_2) = -D_{\xi_2}(x_1, x_2) + \langle \xi_1 - \xi_2, x_1 - x \rangle, \quad (15)$$

Demonstração. Veja [21]. □

Proposição 2.1.5. *Suponha que $f : X \rightarrow \overline{\mathbb{R}}$ seja própria e uniformemente convexa com módulo φ . Para cada $x \in X$, $y \in \text{intdom}(f)$ e $v \in \partial f(y)$ temos que*

$$D_v(x, y) \geq \varphi(\|x - y\|). \quad (16)$$

Demonstração. Da definição de ∂f , dado $v \in \partial f(y)$, para todo $x \in X$ e para todo $\lambda \in (0, 1)$ segue que,

$$f(\lambda x + (1 - \lambda)y) \geq f(y) + \langle v, (\lambda x + (1 - \lambda)y) - y \rangle = f(y) + \lambda \langle v, x - y \rangle. \quad (17)$$

Sendo f uniformemente convexa com módulo φ , temos que,

$f(\lambda x + (1-\lambda)y) + \lambda(1-\lambda)\varphi(\|x-y\|) \leq \lambda f(x) + (1-\lambda)f(y)$. Essa última desigualdade juntamente com (17) nos dá

$$f(y) + \lambda \langle \nu, x-y \rangle + \lambda(1-\lambda)\varphi(\|x-y\|) \leq \lambda f(x) + (1-\lambda)f(y)$$

Assim, $(1-\lambda)\varphi(\|x-y\|) \leq f(x) - f(y) - \langle \nu, x-y \rangle = D_\nu(x, y)$. Fazendo $\lambda \rightarrow 0^+$, temos que $D_\nu(x, y) \geq \varphi(\|x-y\|)$. \square

Lema 2.1.6. *Suponha que $f : X \rightarrow \overline{\mathbb{R}}$ seja própria e uniformemente convexa com módulo φ . Então, para todo $\xi \in \partial f(x)$ e $\nu \in \partial f(y)$, temos que*

$$\langle \xi - \nu, x-y \rangle \geq 2\varphi(\|x-y\|). \quad (18)$$

Consequentemente, $\langle \xi - \nu, x-y \rangle \geq 0$, para todo $\xi \in \partial f(x)$ e $\nu \in \partial f(y)$.

Demonstração. Segue da proposição 2.1.5 que $D_\nu(x, y) \geq \varphi(\|x-y\|)$ assim como $D_\xi(y, x) \geq \varphi(\|x-y\|)$. Somando essas duas últimas desigualdades temos (18). \square

Observação 2.1.7. *Se f é fortemente convexa, i.e., $\varphi(\cdot) = (\beta/2)(\cdot)^2$ então, a desigualdade (16) se torna $D_\nu(x, y) \geq (\beta/2)\|x-y\|^2$ e em vez de (18) temos $\langle \xi - \nu, x-y \rangle \geq \beta\|x-y\|^2$, para todo $\xi \in \partial f(x)$ e $\nu \in \partial f(y)$.*

Dizemos que uma função $f : X \rightarrow \overline{\mathbb{R}}$ tem a propriedade de Kadec se para toda sequência x_k em X tal que $x_k \rightarrow x$ e $f(x_k) \rightarrow f(x)$ então $x_k \rightarrow x$. O próximo Teorema mostra que toda função própria, uniformemente convexa e s.c.i. satisfaz tal propriedade.

Teorema 2.1.8 (Kadec). *Seja $f : \rightarrow \overline{\mathbb{R}}$ própria, uniformemente convexa e s.c.i. Se qualquer sequência x_n é tal que $x_n \rightarrow \bar{x}$ e $f(x_n) \rightarrow f(\bar{x}) < \infty$, então $x_n \rightarrow \bar{x}$.*

Demonstração. Suponha que $x_n \not\rightarrow \bar{x}$, assim a menos de uma subsequência existe $\varepsilon > 0$ tal que $\|x_n - \bar{x}\| \geq \varepsilon$, para todo n . Sendo f uniformemente convexa, temos

$$f\left(\frac{x_n + \bar{x}}{2}\right) + \mu \leq \frac{f(x_n) + f(\bar{x})}{2}, \quad (19)$$

onde $0 < \mu := \frac{1}{4}\varphi(\varepsilon) \leq \frac{1}{4}\varphi(\|x_n - \bar{x}\|)$. Note que $\frac{x_n + \bar{x}}{2} \rightarrow \bar{x}$. Uma vez que $f(x_n) \rightarrow f(\bar{x})$ e f é s.c.i, por (19) temos que $f(\bar{x}) \leq \liminf f\left(\frac{x_n + \bar{x}}{2}\right) \leq \limsup f\left(\frac{x_n + \bar{x}}{2}\right) \leq f(\bar{x}) - \mu$. O que leva a uma contradição. \square

2.2 O MÉTODO HÍBRIDO PROXIMAL EXTRAGRADIENTE

Vamos aqui discutir um pouco sobre o método híbrido proximal extragradiente (HPE) apresentado por [30]. Este método é uma poderosa ferramenta que conjuga etapas do algoritmo do ponto proximal com o método extragradiente para encontrar $x \in X$ tal que

$$0 \in T(x), \quad (20)$$

onde $T : X \rightarrow X$ é um operador monótono maximal. Nesse método um alargamento no operador T é usado, o que permite um relaxamento na tolerância. Cada passo interno consiste em, para cada $x \in X$ resolver o subproblema

$$0 \in T(\cdot) + (\cdot - x). \quad (21)$$

Em 1997 a definição de alargamento de um operador foi introduzida por [32].

Definição 2.2.1. Um operador $T : X \rightarrow 2^X$ é dito monótono se

$$\langle x - y, u - v \rangle \geq 0, \quad \forall (x, u), (y, v) \in G(T),$$

onde $G(T) = \{(x, s) ; s \in T(x)\}$ é o gráfico de T .

Definição 2.2.2. Um operador $T : X \rightarrow 2^X$ é monótono maximal se não existe um operador monótono $B : X \rightarrow 2^X$ tal que $G(B)$ contenha o $G(T)$, i.e para todo $(x, u) \in X \times X$,

$$(x, u) \in G(T) \Leftrightarrow \langle x - y, u - v \rangle \geq 0, \quad \forall (y, v) \in G(T).$$

Definição 2.2.3. Seja T um operador monótono maximal. Dado $\varepsilon \geq 0$, o alargamento de T é o operador $T^\varepsilon : X \rightarrow 2^X$ definido por

$$T^\varepsilon(x) = \{v \in X; \langle w - v, z - x \rangle \geq -\varepsilon, \forall (z, w) \in G(T)\}.$$

Definição 2.2.4. Dados $x \in X$, $c > 0$ e $\sigma \in [0, 1)$. Dizemos que o par $(y, v) \in X \times X$ é uma solução aproximada com erro de tolerância σ do problema $0 \in T(\cdot) + (\cdot - x)$, se para cada $\varepsilon \geq 0$ tivermos que,

$$v \in T^\varepsilon(y) \text{ e } \|cv + y - x\|^2 + 2c\varepsilon \leq \sigma^2 \|y - x\|^2.$$

Lema 2.2.5. Sejam $x \in X$, $c > 0$. Suponha que (y, v) é uma solução aproximada com erro de tolerância σ do problema (21). Suponha que z é uma solução exata de (21), então $\|z - y\| \leq \sigma \|y - x\|$.

Demonstração. Veja [30]. □

Lema 2.2.6. *Sejam $x \in X$, $c > 0$ e $0 \in T(x^*)$. Suponha que (y, v) é uma solução aproximada com erro de tolerância σ do problema (21). Se $x^+ = x - cv$, então*

$$\|x^* - x\|^2 - \|x^* - x^+\|^2 \geq (1 - \sigma^2)\|y - x\|^2.$$

Demonstração. Veja [30]. □

No que segue apresentamos o algoritmo híbrido extragradiente proximal (HPE).

- 0] inicie com $x_0 \in X$ e $\sigma \in [0, 1)$;
 1] Escolha $c_k > 0$ e encontre $(y_k, v_k, \varepsilon_k) \in X \times X \times \mathbb{R}_+$ tal que $v_k \in T^{\varepsilon_k}(y_k)$ e $\|c_k v_k + y_k - x_k\|^2 + 2c_k \varepsilon_k \leq \sigma^2 \|y_k - x_k\|^2$
 2] Defina $x_{k+1} = x_k - c_k v_k$.

Algoritmo 1 – Algoritmo HPE.

A fim de assegurar a convergência do Algoritmo HPE duas condições são necessárias, uma é que o problema (20) tenha solução e outra é que a sequência c_k seja limitada longe do zero.

Observação 2.2.7. *Os Lemas 2.2.5 e 2.2.6 podem ser adaptados para receberem as sequências envolvidas no Algoritmo 1, para isso basta fazer*

$$x \leftarrow x_k, \quad y \leftarrow y_k \quad \text{and} \quad x^+ \leftarrow x_{k+1}.$$

Portanto pelo Lema (2.2.6) temos que $(1 - \sigma^2)\|y_k - x_k\|^2 \leq \|x^ - x_k\|^2 - \|x^* - x_{k+1}\|^2$, isso implica que $(1 - \sigma^2) \sum_{k=0}^l \|y_k - x_k\|^2 \leq \|x^* - x_0\|^2 - \|x^* - x_{l+1}\|^2 \leq \|x^* - x_0\|^2$. Portanto*

$$\sum_{k=0}^{\infty} \|y_k - x_k\|^2 < \infty, \quad \text{consequentemente} \quad \|y_k - x_k\| \rightarrow 0 \quad \text{quando} \quad k \rightarrow \infty.$$

Sobre a hipótese de limitação inferior na sequência de escalares c_k e que $T^{-1}(0) \neq \emptyset$, i.e, existe algum $x \in X$ tal que $0 \in T(x)$, Solodov e Svaiter mostraram que a sequência gerada pela iteração acima converge fracamente para uma solução de (20). Se além dessas hipóteses supormos ainda que T^{-1} é Lipschitz contínuo em torno de 0, então temos convergência linear.

Teorema 2.2.8. *Suponha que $c_k \geq \bar{c} > 0$ para todo $k \geq 0$ e que $T^{-1}(0) \neq \emptyset$. Então a sequência x_k gerada pelo Algoritmo 1 converge fracamente para um elemento de $T^{-1}(0)$.*

Demonstração. Veja [30]. □

Teorema 2.2.9. *Suponha as hipóteses do Teorema 2.2.8. Se T^{-1} é Lipschitz contínuo em torno de 0 então a sequência x_k converge linearmente para um elemento de $T^{-1}(0)$.*

Demonstração. Seja z_k uma solução do subproblema $0 \in T(\cdot) + (\cdot - x_k)$, então $u_k := \frac{-1}{c_k}(z_k - x_k) \in T(z_k)$.

Segue do Lema 2.2.5 que $\|z_k - y_k\| \leq \sigma\|y_k - x_k\|$. Consequentemente, $\|z_k - x_k\| \leq \|z_k - y_k\| + \|y_k - x_k\| \leq (1 + \sigma)\|y_k - x_k\|$.

Assim, segue da Observação 2.2.7 que, $\|u_k\| = \frac{1}{c_k}\|z_k - x_k\| \leq \frac{1}{c}(1 + \sigma)\|y_k - x_k\|$, dessa forma $\|u_k\| \rightarrow 0$ quando $k \rightarrow \infty$,

Seja L a constante de Lipschitz de T^{-1} e defina $\kappa := L/\bar{c}$, assim, temos que $\|z_k - x^*\| \leq L\|u_k\| \leq \kappa(1 + \sigma)\|y_k - x_k\|$. Por outro lado,

$\|x_k - x^*\| \leq \|x_k - z_k\| + \|z_k - x^*\| \leq (1 + \sigma)\|y_k - x_k\| + \kappa(1 + \sigma)\|y_k - x_k\| = (1 + \kappa)(1 + \sigma)\|y_k - x_k\|$. Portanto segue do Lema 2.2.6 que

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \|x_k - x^*\|^2 - (1 - \sigma^2)\|y_k - x_k\|^2 \\ &\leq \|x_k - x^*\|^2 - \frac{1 - \sigma^2}{(1 + \kappa)^2(1 + \sigma)^2}\|x_k - x^*\|^2 \\ &= \left(1 - \frac{1 - \sigma^2}{(1 + \kappa)^2(1 + \sigma)^2}\right)\|x_k - x^*\|^2. \end{aligned}$$

□

2.3 RESULTADOS PRELIMINARES

Neste capítulo apresentamos alguns dos resultados que podem ser encontrados na clássica teoria da regularização no contexto de métodos tipo Tikhonov. A penalização com norma Euclidiana dada em (3) é substituída por uma função mais geral $P : X \rightarrow [0, +\infty]$ de modo que P seja própria, s.c.i e uniformemente convexa com módulo φ onde $\lim_{t \rightarrow \infty} \varphi(t) = \infty$. Numa situação comum por exemplo, se $P(\cdot) = \|\cdot\|^2$ seu módulo de convexidade satisfaz essas propriedades. Os resultados apresentados nesta seção não são novos e versões semelhantes (ou até mais gerais) podem ser encontradas na literatura. Apresentamos todas as provas para comodidade do leitor e também porque queremos usar essas estratégias para apresentar os resultados do Capítulo 3 de modo mais claro.

No que segue, vamos assumir que X e Y sejam espaços de Hilbert reais, $A : X \rightarrow Y$ é um operador linear e limitado, $b \in R(A)$ é um vetor fixado e $b^\delta \in Y$ é um vetor satisfazendo (2). Para cada $\alpha > 0$ vamos definir o funcional de Tikhonov $T_\alpha : X \rightarrow \overline{\mathbb{R}}$ como sendo

$$T_\alpha(x) = \frac{1}{2}\|Ax - b^\delta\|^2 + \alpha P(x). \quad (22)$$

Lema 2.3.1. Para cada $\alpha > 0$, existe um único minimizador do funcional de Tikhonov T_α definido em (22).

Demonstração. Fixe $\alpha > 0$ e seja x_k uma sequência em X tal que $T_\alpha(x_k) \rightarrow \inf T_\alpha(x)$. Então a sequência $T_\alpha(x_k)$ é limitada em k . Agora usando $f = P$, $\lambda = 1/2$, $x = x_k$ e $y = 0$

em (12), obtemos

$$P\left(\frac{1}{2}x_k\right) + \frac{1}{4}\varphi(\|x_k\|) \leq \frac{1}{2}P(x_k) + \frac{1}{2}P(0). \quad (23)$$

Uma vez que $P\left(\frac{1}{2}x_k\right) \geq 0$, segue que

$$\|x_k\| \leq \varphi^{-1}\left(\frac{2}{\alpha}T_\alpha(x_k) + 2P(0)\right),$$

pois $\lim_{t \rightarrow \infty} \varphi(t) = \infty$. Com isso provamos que x_k é limitada. Consequentemente, essa sequência admite uma subsequência fracamente convergente, digamos, $x_{k_j} \rightharpoonup \bar{x} \in X$. Agora, como P é f.s.c.i., a desigualdade $P(\bar{x}) \leq \liminf P(x_{k_j})$ é verdadeira. Analogamente, $\|A\bar{x} - b^\delta\|^2 \leq \liminf \|Ax_{k_j} - b^\delta\|^2$, e portanto

$$T_\alpha(\bar{x}) \leq \liminf T_\alpha(x_{k_j}) = \lim T_\alpha(x_{k_j}) = \inf_{x \in X} T_\alpha(x).$$

Dessa forma \bar{x} é um minimizador de T_α . A unicidade do minimizador segue da convexidade estrita de P . \square

Observação 2.3.2. *O Lema 2.3.1 garante, em particular, a existência de um único minimizador de (50), pois a distância de Bregman $D_{\varepsilon_k}(\cdot, x_k^\delta)$ define uma função estritamente convexa.*

Vamos denotar o único minimizador de T_α por x_α . Segue de (13) que

$$A^*Ax_\alpha + \alpha\partial P(x_\alpha) \ni A^*b^\delta. \quad (24)$$

Agora, definimos $s(b) := \{x \in X : Ax = b\}$ como o conjunto de todas as soluções do problema inverso (1).

Lema 2.3.3. *Existe uma única P -solução mínima (P -SM), definida como sendo o único vetor $x^\dagger \in s(b)$ satisfazendo*

$$P(x^\dagger) \leq P(x), \text{ para todo } x \in s(b). \quad (25)$$

Demonstração. Uma vez que P é própria e uniformemente convexa e o conjunto $s(b)$ é não vazio, fechado e convexo, o problema de minimizar $P(x)$ sujeito a $x \in s(b)$ possui solução única x^\dagger , que satisfaz (25). A prova segue a direção do que fizemos na prova do Lema 2.3.1. \square

Uma estimativa que usaremos mais tarde é a seguinte,

$$\begin{aligned} \alpha\left(P(x_\alpha) - P(x^\dagger)\right) &\leq T_\alpha(x_\alpha) - \alpha P(x^\dagger) \leq T_\alpha(x^\dagger) - \alpha P(x^\dagger) \\ &= \frac{1}{2}\|b - b^\delta\|^2 \leq \frac{1}{2}\delta^2. \end{aligned} \quad (26)$$

O que implica, $P(x_\alpha) - P(x^\dagger) \leq \delta^2/2\alpha$.

Proposição 2.3.4. *Assuma que a escolha a priori dos parâmetros de regularização (4) é realizada. Então, $x_{\alpha(\delta)} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$, onde x^\dagger é a P-SM. Adicionalmente, se $\alpha = o(\delta)$ e vale a condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ então*

$$D_\xi(x_{\alpha(\delta)}, x^\dagger) = \mathcal{O}(\delta), \quad (27)$$

onde $\xi \in \partial P(x^\dagger) \cap R(A^*)$. Além do mais, se P fortemente convexa, então

$$\|x_{\alpha(\delta)} - x^\dagger\| = \mathcal{O}(\delta^{1/2}), \quad (28)$$

a qual é a taxa de convergência ótima sob a condição de fonte assumida.

Demonstração. Primeiro vamos mostrar o resultado de convergência. Considere δ_n uma sequência de números positivos convergindo para zero. Vamos fixar uma subsequência de δ_n , a qual será denotada novamente por δ_n . Defina $\alpha_n := \alpha(\delta_n)$, $b_n := b^{\delta_n}$ e $x_n := x_{\alpha_n}$. Segue do Lema 2.3.1 que x_n é o único minimizador de T_n , onde

$$T_n(x) = \frac{1}{2} \|Ax - b_n\|^2 + \alpha_n P(x).$$

De modo análogo a (26), obtemos

$$P(x_n) \leq \frac{\delta_n^2}{2\alpha_n} + P(x^\dagger), \quad (29)$$

assim segue de (4) que $P(x_n)$ é limitada. Como feito na prova do Lema 2.3.1, obtemos uma subsequência fracamente convergente $x_{n_k} \rightharpoonup \bar{x} \in X$. Além disso,

$$\begin{aligned} \|Ax_{n_k} - b\| &\leq \|b - b_{n_k}\| + \|Ax_{n_k} - b_{n_k}\| \leq \delta_{n_k} + \sqrt{2T_{n_k}(x_{n_k})} \\ &\leq \delta_{n_k} + \sqrt{2T_{n_k}(x^\dagger)} \leq \delta_{n_k} + \sqrt{\delta_{n_k}^2 + 2\alpha_{n_k} P(x^\dagger)} \rightarrow 0, \end{aligned} \quad (30)$$

quando $k \rightarrow \infty$. Então $Ax_{n_k} \rightarrow b$, como $Ax_{n_k} \rightharpoonup A\bar{x}$, concluímos que $A\bar{x} = b$. Sendo P f.s.c.i, segue de (29) e de (4) que

$$P(\bar{x}) \leq \liminf P(x_{n_k}) \leq \limsup P(x_{n_k}) \leq P(x^\dagger).$$

Isso prova que $\bar{x} = x^\dagger$, e também implica que $P(x_{n_k}) \rightarrow P(x^\dagger) < \infty$. Como P é uniformemente convexa, P satisfaz a propriedade de Kadec, e como $x_{n_k} \rightharpoonup x^\dagger$, segue que $x_{n_k} \rightarrow x^\dagger$ quando $k \rightarrow \infty$.

O raciocínio acima mostra que qualquer subsequência de δ_n tem ela própria uma subsequência δ_{n_k} tal que $x_{n_k} \rightarrow x^\dagger$. Isso mostra que $x_n \rightarrow x^\dagger$. Finalmente, como $\delta_n \rightarrow 0$ é arbitrária, segue que $x_{\alpha(\delta)} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$.

Agora vamos provar as taxas (27) e (28). Assuma a condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ bem como $\alpha = o(\delta)$. Assim existe um $\xi \in \partial P(x^\dagger)$ satisfazendo $\xi = A^*w$ tal

que $\|w\| \leq E$. Observe que

$$\begin{aligned} 0 &\in A^*(Ax_\alpha - b^\delta) + \alpha \partial P(x_\alpha) \\ 0 &= A^*(Ax^\dagger - b), \end{aligned}$$

de onde temos que

$$A^*A(x_\alpha - x^\dagger) \in A^*(b^\delta - b) - \alpha \partial P(x_\alpha), \quad (31)$$

multiplicando (31) por $x_\alpha - x^\dagger$, nós obtemos

$$\|A(x_\alpha - x^\dagger)\|^2 = \langle b^\delta - b, A(x_\alpha - x^\dagger) \rangle - \alpha \langle y_\alpha, x_\alpha - x^\dagger \rangle,$$

para um elemento adequado $y_\alpha \in \partial P(x_\alpha)$. Como $\xi \in \partial P(x^\dagger)$, segue a partir do Lema 2.1.6 que,

$$\begin{aligned} \|A(x_\alpha - x^\dagger)\|^2 &= \langle b^\delta - b, A(x_\alpha - x^\dagger) \rangle - \alpha \langle y_\alpha - \xi, x_\alpha - x^\dagger \rangle + \alpha \langle \xi, x^\dagger - x_\alpha \rangle \\ &\leq \|b^\delta - b\| \cdot \|A(x_\alpha - x^\dagger)\| + 0 + \alpha \langle w, A(x^\dagger - x_\alpha) \rangle \\ &\leq \delta \|A(x_\alpha - x^\dagger)\| + \alpha E \|A(x^\dagger - x_\alpha)\|. \end{aligned}$$

Então,

$$\|A(x_\alpha - x^\dagger)\| \leq \delta + \alpha E, \quad (32)$$

e por (26) temos que

$$\begin{aligned} D_\xi(x_\alpha, x^\dagger) &= P(x_\alpha) - P(x^\dagger) - \langle \xi, x_\alpha - x^\dagger \rangle \leq \frac{\delta^2}{2\alpha} + \langle \xi, x^\dagger - x_\alpha \rangle \\ &\leq \frac{\delta^2}{2\alpha} + E \|A(x^\dagger - x_\alpha)\|. \end{aligned}$$

Portanto, como $\alpha = o(\delta)$, obtemos de (32) que $D_\xi(x_{\alpha(\delta)}, x^\dagger) = \mathcal{O}(\delta)$. Finalmente, (28) segue da Observação 2.1.7. \square

A fim de obter resultados semelhantes para a escolha a posteriori dos parâmetros (7), vamos precisar provar primeiro que essa regra está bem definida. Observe que podemos assumir que $\|b^\delta\| > \delta$ para todo $\delta > 0$. Caso contrário, uma solução trivial do problema inverso (1) seria $x = 0$, porque nesse caso, o vetor $b = Ax = 0$ satisfaz $\|b^\delta - b\| = \|b^\delta\| \leq \delta$. Note que, a desigualdade $\|b^\delta\| > \delta$ implica a partir de (2) que $b \neq 0$.

Lema 2.3.5. *Considere $1 < c_1 \leq c_2$ constantes fixadas. Então, para qualquer $\delta > 0$ suficientemente pequeno, existe $\alpha > 0$ tal que (7) é satisfeito.*

Demonstração. Fixe $\delta > 0$. Primeiro vamos mostrar que a função $g : (0, \infty) \rightarrow X$, $\alpha \mapsto x_\alpha$, é contínua. De fato, observe que

$$\alpha P(x_\alpha) \leq T_\alpha(x_\alpha) \leq T_\alpha(0) = \|b^\delta\|^2 + \alpha P(0). \quad (33)$$

Escolha $0 < \gamma \neq \alpha$. Agora use (24) para α e γ , em seguida subtraia as duas inclusões e multiplique tudo por $x_\alpha - x_\gamma$. Isso vai nos dar

$$\|A(x_\alpha - x_\gamma)\|^2 + \alpha \langle \xi_\alpha - \xi_\gamma, x_\alpha - x_\gamma \rangle + (\alpha - \gamma) \langle \xi_\gamma, x_\alpha - x_\gamma \rangle = 0, \quad (34)$$

onde $\xi_\alpha \in \partial P(x_\alpha)$ e $\xi_\gamma \in \partial P(x_\gamma)$. Tendo em vista o Lema 2.1.6, segue que

$$2\alpha\varphi(\|x_\alpha - x_\gamma\|) \leq \alpha \langle \xi_\alpha - \xi_\gamma, x_\alpha - x_\gamma \rangle \leq (\gamma - \alpha) \langle \xi_\gamma, x_\alpha - x_\gamma \rangle.$$

Agora, se $\alpha < \gamma$, de (33) e da definição de subdiferencial, temos

$$2\alpha\varphi(\|x_\alpha - x_\gamma\|) \leq (\gamma - \alpha) (P(x_\alpha) - P(x_\gamma)) \leq (\gamma - \alpha) P(x_\alpha) \leq \frac{\gamma - \alpha}{\alpha} (\|b^\delta\|^2 + \alpha P(0)).$$

Por outro lado, se $\gamma < \alpha$, podemos proceder de modo similar para obter

$$2\gamma\varphi(\|x_\alpha - x_\gamma\|) \leq (\alpha - \gamma) \langle \xi_\alpha, x_\gamma - x_\alpha \rangle \leq (\alpha - \gamma) P(x_\gamma) \leq \frac{\alpha - \gamma}{\gamma} (\|b^\delta\|^2 + \gamma P(0)).$$

Portanto, em qualquer dos casos temos

$$\varphi(\|x_\alpha - x_\gamma\|) \leq \left[\frac{\|b^\delta\|^2}{2 \min\{\alpha^2, \gamma^2\}} + \frac{P(0)}{2 \min\{\alpha, \gamma\}} \right] |\gamma - \alpha|,$$

isso implica que $\lim_{\gamma \rightarrow \alpha} \|x_\alpha - x_\gamma\| = 0$. Consequentemente g é contínua em α . Isso por sua vez mostra que a função $\alpha \mapsto \|Ax_\alpha - b^\delta\|$ é contínua.

A ideia que usaremos abaixo já conhecida na literatura pode ser encontrada em [22, Lema 30]. Note que

$$\lim_{\alpha \rightarrow 0} \|Ax_\alpha - b^\delta\| = \inf_{x \in X} \|Ax - b^\delta\| \leq \|Ax^\dagger - b^\delta\| \leq \delta < c_2\delta, \quad (35)$$

e

$$\lim_{\alpha \rightarrow \infty} \|Ax_\alpha - b^\delta\| = \|b^\delta\|.$$

$\|b^\delta\| \rightarrow \|b\| > 0$ quando $\delta \rightarrow 0$, consequentemente

$$\lim_{\delta \rightarrow 0^+} \frac{\|b^\delta\|}{\delta} = \infty,$$

e portanto,

$$\lim_{\alpha \rightarrow \infty} \|Ax_\alpha - b^\delta\| > c_1\delta \quad (36)$$

para todo $\delta > 0$ suficientemente pequeno. O fato de (7) ser satisfeito segue de (35), (36) e da continuidade da função $\alpha \mapsto \|Ax_\alpha - b^\delta\|$. \square

Proposição 2.3.6. *Se vale a escolha a posteriori parâmetros (7), então, $x_{\alpha(\delta)} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$, onde x^\dagger é uma P-SM. Adicionalmente, se a condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ é satisfeita, então (27) também vale. Além disso, se P é fortemente convexa, então (28) vale.*

Demonstração. Primeiro vamos mostrar que $x_{\alpha(\delta)} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$. De (7) segue que

$$\alpha P(x_\alpha) + \frac{1}{2}\delta^2 \leq \alpha P(x_\alpha) + \frac{1}{2}c_1^2\delta^2 \leq T_\alpha(x_\alpha) \leq T_\alpha(x^\dagger) \leq \alpha P(x^\dagger) + \frac{1}{2}\delta^2,$$

isso implica que

$$P(x_\alpha) \leq P(x^\dagger), \quad (37)$$

para todo $\delta > 0$. Escolha $\delta_n \rightarrow 0$ e defina $\alpha_n := \alpha(\delta_n)$, $b_n := b^{\delta_n}$ e $x_n := x_{\alpha_n}$. Como $P(x_n)$ é limitada, a sequência x_n também é (veja a prova do Lema 2.3.1). A partir daqui a prova segue os passos da prova que fizemos na Proposição 2.3.4. A única modificação vem em (30), que passa a ser

$$\|Ax_{n_k} - b\| \leq \|b - b_{n_k}\| + \|Ax_{n_k} - b_{n_k}\| \leq \delta_{n_k} + c_2\delta_{n_k} \rightarrow 0,$$

quando $k \rightarrow \infty$.

Agora assumindo que $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ vamos provar (27). Seja $\xi \in \partial P(x^\dagger) \cap R(A^*)$. Assim, $\xi = A^*w$, para algum $w \in Y$ tal que $\|w\| \leq E$, onde $E > 0$. Usando (37) obtemos

$$\begin{aligned} D_\xi(x_\alpha, x^\dagger) &= P(x_\alpha) - P(x^\dagger) - \langle \xi, x_\alpha - x^\dagger \rangle \\ &\leq \langle \xi, x^\dagger - x_\alpha \rangle = \langle w, b - Ax_\alpha \rangle \\ &= \langle w, b - b^\delta \rangle + \langle w, b^\delta - Ax_\alpha \rangle \leq E(1 + c_2)\delta. \end{aligned}$$

Portanto, $D_\xi(x_\alpha, x^\dagger) = \mathcal{O}(\delta)$. A taxa de convergência dada por (28) segue da Observação 2.1.7. \square

3 MÉTODOS DE TIKHONOV DE UM PASSO COM MINIMIZAÇÃO INEXATA

Neste capítulo vamos apresentar um método iterativo para produzir iterações internas a fim de obter uma aproximação para x_α . A seguir vamos apresentar condições suficientes para obter resultados de convergência como em (5) e de taxa de convergência dado em (6). Vamos fazer isso tanto para o caso a priori, com a escolha dos parâmetro de regularização feita em (4), quanto para o caso a posteriori com a escolha dos parâmetros realizada em (7). Nesse caso o minimizador exato x_α do funcional T_α é substituído por uma aproximação.

Vamos considerar o funcional de Tikhonov T_α como foi definido em (22) e $P : X \rightarrow [0, \infty]$ e uma função própria, s.c.i. e uniformemente convexa com módulo de convexidade φ . Para cada $\alpha > 0$, o funcional T_α tem um único minimizador (veja Lema 2.3.1) que será denotado por x_α . Os chamados Métodos de Tikhonov de um passo são aqueles métodos onde o minimizador x_α aproxima a solução do problema inverso (1). O parâmetro de regularização α depende do nível de ruído δ e pode ser escolhido a priori (independente de b^δ) ou a posteriori (dependendo de b^δ).

Vamos dar início a nossa análise exigindo uma condição sobre a sequência x_k gerada na iteração interna.

Hipótese 3.0.1. Assuma que para cada $\alpha > 0$, e $x_0 \in X$, uma sequência x_k satisfazendo

$$\|x_k - x_\alpha\| \leq \Lambda_\alpha \|x_{k-1} - x_\alpha\|, \text{ para todo } k \in \mathbb{N} \quad (38)$$

é conhecida, onde $0 < \Lambda_\alpha < 1$.

Observação 3.0.2. A Hipótese 3.0.1 é bastante geral. Na seção 4.3, nós apresentamos uma lista de métodos iterativos em otimização que podem ser aplicados ao funcional de Tikhonov (22) a fim de gerar uma sequência satisfazendo (38). Essa lista de métodos inclui os conhecidos ADMM, Método do ponto proximal, o Método Forward-Backward, entre outros, veja a seção 4.3.

A partir de (38), temos que

$$\|x_k - x_\alpha\| \leq \Lambda_\alpha \|x_{k-1} - x_\alpha\| \leq \Lambda_\alpha^2 \|x_{k-2} - x_\alpha\| \leq \dots \leq \Lambda_\alpha^k \|x_0 - x_\alpha\|.$$

Portanto,

$$\lim_{k \rightarrow \infty} \|x_k - x_\alpha\| \leq \|x_0 - x_\alpha\| \lim_{k \rightarrow \infty} \Lambda_\alpha^k = 0,$$

o que significa que a sequência (x_k) gerada pelas iterações internas converge para o minimizador exato de T_α quando $k \rightarrow \infty$.

Primeiro vamos analisar o método de Tikhonov combinando minimização inexata com a estratégia de escolha a priori do parâmetro de regularização (4).

3.1 REGRA A PRIORI DA ESCOLHA DOS PARÂMETROS DE REGULARIZAÇÃO

Inicialmente escolhemos um critério apropriado para encerrar a iteração interna do nosso método, a saber

$$k_\delta := \lfloor \log_{\Lambda_\alpha} (f(\delta)) \rfloor, \quad (39)$$

onde $f : (0, \infty) \rightarrow (0, 1)$ é uma função tal que

$$\lim_{\delta \rightarrow 0^+} f(\delta) = 0. \quad (40)$$

Aqui $\lfloor x \rfloor$ denota o menor inteiro maior ou igual do que x , para cada $x \in \mathbb{R}$.

Teorema 3.1.1. *Suponha que a Hipótese 3.0.1 acontece e que a escolha a priori dos parâmetros de regularização (4) é usada. Se a iteração interna termina com k_δ definido em (39), então a convergência dada em (5) vale com x_{k_δ} no lugar de $x_{\alpha(\delta)}$, isto é, $x_{k_\delta} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$.*

Demonstração. Fixe um elemento $x_0 \in X$. Da definição de k_δ segue que $k_\delta \geq \log_{\Lambda_\alpha} (f(\delta))$ onde $0 < \Lambda_\alpha < 1$. Consequentemente $\Lambda_\alpha^{k_\delta} \leq f(\delta)$. Portanto,

$$\begin{aligned} \|x_{k_\delta} - x^\dagger\| &\leq \|x_{k_\delta} - x_\alpha\| + \|x_\alpha - x^\dagger\| \\ &\leq \Lambda_\alpha^{k_\delta} \|x_0 - x_\alpha\| + \|x_\alpha - x^\dagger\| \\ &\leq f(\delta) \|x_0 - x_\alpha\| + \|x_\alpha - x^\dagger\| \\ &\leq f(\delta) \|x_0 - x^\dagger\| + 2\|x_\alpha - x^\dagger\|. \end{aligned}$$

A conclusão agora é clara tendo em vista que $x_\alpha \rightarrow x^\dagger$ quando $\delta \rightarrow 0$ (veja a Proposição 2.3.4) e que $f(\delta) \rightarrow 0$ quando $\delta \rightarrow 0$. \square

Teorema 3.1.2. *Assuma todas as hipóteses do Teorema 3.1.1, além disso que P seja fortemente convexa. Escolha $f(\delta) = \mathcal{O}(\sqrt{\delta})$ e $\alpha = o(\delta)$. Se a condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ vale, então a taxa de convergência*

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

é alcançada.

Demonstração. Segue da Proposição 2.3.4 que $\|x_\alpha - x^\dagger\| = \mathcal{O}(\sqrt{\delta})$. Agora, pelo Teorema 3.1.1, temos que $\|x_{k_\delta} - x^\dagger\| \leq f(\delta)\|x_0 - x^\dagger\| + 2\|x_\alpha - x^\dagger\|$, assim o resultado vem da escolha de $f(\delta) = \mathcal{O}(\sqrt{\delta})$. \square

3.2 REGRA A POSTERIORI DA ESCOLHA DOS PARÂMETROS DE REGULARIZAÇÃO

Nesse momento vamos analisar o método de Tikhonov de um passo com minimização inexata combinado com a escolha a posteriori dos parâmetros de regularização

(7). Porém, nesse caso testar (7) se torna bem mais complicada do que para o caso do minimizador exato. Pois, o que temos é apenas uma aproximação do minimizador exato, uma vez que se trata de minimização inexata. Dessa forma, a condição (7) deve ser substituída por

$$c_3\delta \leq \|Ax_{k_\delta} - b^\delta\| \leq c_4\delta, \quad (41)$$

onde c_3 e c_4 são constantes positivas conhecidas. Neste caso, o número de parada das iterações internas $k_\delta \in \mathbb{N}$ passa a ser definido por

$$k_\delta := \lfloor \log_{\wedge_\alpha} (\min \{1, \alpha\} f(\delta)) \rfloor, \quad (42)$$

onde $f : (0, \infty) \rightarrow (0, 1)$ satisfaz (40).

O próximo resultado nos permite concluir que o critério (41) é satisfeito para qualquer $\delta > 0$ pequeno suficiente.

Lema 3.2.1. *Considere as constantes $1 < c_3 < c_1 \leq c_2 < c_4$. Então existe $\delta_0 > 0$ tal que, para cada $0 < \delta \leq \delta_0$, existe um parâmetro $\alpha(\delta) > 0$ com $x_{\alpha(\delta)}$ satisfazendo (7). Além disso, a família $x_{\alpha(\delta)}$ com $0 < \delta \leq \delta_0$ é limitada. Fixe $x_0 \in X$, assuma a Hipótese 3.0.1 e que a função f em (42) satisfaz $f(\delta) \leq c_5\delta$, para todo $0 < \delta \leq \delta_0$, com*

$$c_5 = \frac{\min \{c_1 - c_3, c_4 - c_2\}}{C \|A\|} > 0, \quad (43)$$

e $C > 0$ uma cota superior para $\|x_0 - x_{\alpha(\delta)}\|$. Então, para cada $0 < \delta \leq \delta_0$, deve existir $\alpha > 0$ de modo que (41) é satisfeito.

Demonstração. Pelo Lema 2.3.5, existe $\delta_0 > 0$ tal que, para cada $0 < \delta \leq \delta_0$, existe um parâmetro $\alpha(\delta) > 0$ com $x_{\alpha(\delta)}$ satisfazendo (7). Segue da Proposição 2.3.6, que a família $x_{\alpha(\delta)}$ com $0 < \delta \leq \delta_0$ é convergente, e portanto é limitada. Fixe $0 < \delta \leq \delta_0$. Então,

$$\begin{aligned} \left| \|Ax_{k_\delta} - b^\delta\| - \|Ax_{\alpha(\delta)} - b^\delta\| \right| &\leq \|A\| \|x_{k_\delta} - x_{\alpha(\delta)}\| \\ &\leq \|A\| \wedge_\alpha^{k_\delta} \|x_0 - x_{\alpha(\delta)}\| \\ &\leq C \|A\| f(\delta) \leq \min \{c_1 - c_3, c_4 - c_2\} \delta, \end{aligned}$$

tendo em vista que vale (7), as desigualdades $c_3 < c_1 \leq c_2 < c_4$ implicam (41). \square

No Lema 3.2.1 provamos essencialmente que, sob condições adequadas,

$$\|Ax_{\alpha(\delta)} - b^\delta\| = o(\delta) \implies \|Ax_{k_\delta} - b^\delta\| = o(\delta), \quad (44)$$

isso é suficiente para provar que a regra (41) está bem definida. Nosso objetivo principal agora é provar que $x_{k_\delta} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$, e para isso é necessário provar a recíproca de (44). Note que não temos conhecimento de que x_{k_δ} com $0 < \delta \leq \delta_0$ é limitada.

dessa forma, a demonstração não deve seguir os passos do que fizemos no Lema 3.2.1.

Suponha que P seja fortemente convexa com constante $\beta > 0$. Sejam $\bar{x} \in X$ o minimizador de P em X e $\alpha > 0$ fixado. A existência de \bar{x} é garantida com base no Lema 2.3.3. Considere $\delta > 0$. Se $\delta < \|Ax_\alpha - b^\delta\|$, então

$$\alpha P(x_\alpha) + \frac{1}{2}\delta^2 < T_\alpha(x_\alpha) \leq T_\alpha(x^\dagger) \leq \alpha P(x^\dagger) + \frac{1}{2}\delta^2,$$

o que implica $P(x_\alpha) \leq P(x^\dagger)$. Consequentemente, $\|x_\alpha\| \leq 2\sqrt{\frac{1}{\beta}(P(x^\dagger) + P(0))} := \omega$, veja (23). Logo temos que $\|x_\alpha - \bar{x}\| \leq \omega + \|\bar{x}\|$. Por outro lado, se $\|Ax_\alpha - b^\delta\| \leq \delta$, segue da definição de x_α que $\xi_\alpha = -\frac{1}{\alpha}A^*(Ax_\alpha - b^\delta) \in \partial P(x_\alpha)$, e pela Observação 2.1.7 temos

$$\begin{aligned} \frac{\beta}{2} \|\bar{x} - x_\alpha\|^2 &\leq D_{\xi_\alpha}(\bar{x}, x_\alpha) = P(\bar{x}) - P(x_\alpha) - \langle \xi_\alpha, \bar{x} - x_\alpha \rangle \\ &\leq \langle \xi_\alpha, x_\alpha - \bar{x} \rangle = -\frac{1}{\alpha} \langle Ax_\alpha - b^\delta, A(x_\alpha - \bar{x}) \rangle \\ &\leq \frac{\|A\|}{\alpha} \|Ax_\alpha - b^\delta\| \|x_\alpha - \bar{x}\| \leq \frac{\|A\|}{\alpha} \delta \|x_\alpha - \bar{x}\|. \end{aligned}$$

Assim, $\|x_\alpha - \bar{x}\| \leq 2\frac{\|A\|}{\beta\alpha}\delta$. Em todo caso teremos,

$$\|A\| \wedge_\alpha^{k_\delta} \|x_\alpha - \bar{x}\| \leq \max \left\{ f(\delta) \|A\| (\omega + \|\bar{x}\|), 2\alpha f(\delta) \frac{\|A\|^2}{\beta\alpha} \delta \right\}.$$

Fixe $\epsilon > 0$. Uma vez que $f(\delta) \rightarrow 0$ quando $\delta \rightarrow 0$, podemos obter $2f(\delta) \frac{\|A\|^2}{\beta} < \frac{\epsilon}{2}$, para $\delta > 0$ suficientemente pequeno. Por outro lado, se $f(\delta) \leq \frac{\epsilon}{2\|A\|(\omega + \|\bar{x}\|)}\delta$, temos que

$$\|A\| \wedge_\alpha^{k_\delta} \|x_\alpha - \bar{x}\| \leq \max \left\{ \frac{\epsilon}{2}\delta, \frac{\epsilon}{2}\delta \right\} = \frac{\epsilon}{2}\delta. \quad (45)$$

Dessa forma, definindo

$$c_6 = \min \left\{ \frac{\epsilon}{2\|A\|(\omega + \|\bar{x}\|)}, \frac{\epsilon}{2\|A\|\|x_0 - \bar{x}\|} \right\} > 0, \quad (46)$$

para cada $\delta > 0$ suficientemente pequeno e $f(\delta) \leq c_6\delta$, temos que

$$\begin{aligned} \left| \|Ax_{k_\delta} - b^\delta\| - \|Ax_\alpha - b^\delta\| \right| &\leq \|A\| \|x_{k_\delta} - x_\alpha\| \leq \|A\| \wedge_\alpha^{k_\delta} (\|x_0 - \bar{x}\| + \|x_\alpha - \bar{x}\|) \\ &\leq f(\delta) \|A\| \|x_0 - \bar{x}\| + \|A\| \wedge_\alpha^{k_\delta} \|x_\alpha - \bar{x}\| \leq \epsilon\delta. \end{aligned} \quad (47)$$

Teorema 3.2.2. *Suponha que a Hipótese 3.0.1 e que P seja fortemente convexa. Fixe $x_0 \in X$, $1 < c_3 < c_4$ e defina k_δ como em (42). Então existe uma constante $\hat{c} > 0$ tal que, se $f(\delta) \leq \hat{c}\delta$ para todo $\delta > 0$ suficientemente pequeno, a escolha a posteriori do parâmetro de regularização, como em (41), está bem definida e além disso*

$$\lim_{\delta \rightarrow 0} x_{k_\delta} = x^\dagger.$$

Demonstração. Defina $\epsilon = \min \{c_3 - 1, c_4 - c_3\} / 2 > 0$. Com isso podemos definir $c_1 = c_3 + \epsilon$ e $c_2 = c_4 - \epsilon$, o que implica $1 < c_3 < c_1 \leq c_2 < c_4$. Pelo Lema 3.2.1, se $f(\delta) \leq c_5\delta$, com c_5 como em (43), então deve existir $\delta_0 > 0$ de modo que a regra (41) está bem definida para todo $0 < \delta \leq \delta_0$.

Fixe $0 < \delta \leq \delta_0$ e seja $\alpha > 0$ dado de modo que (41) valha. Defina $\hat{c} = \min \{c_5, c_6\} > 0$, onde c_6 é definido como em (46). Agora vamos provar que vale a recíproca de (44). Para ser mais claro, vamos mostrar que se δ_0 é pequeno o suficiente e $f(\delta) \leq \hat{c}\delta$ para todo $0 < \delta \leq \delta_0$, o minimizador exato de T_α, x_α , satisfaz

$$c_7\delta \leq \|Ax_\alpha - b^\delta\| \leq c_8\delta, \quad (48)$$

onde $c_7 = c_3 - \epsilon$ e $c_8 = c_4 + \epsilon$.¹ De fato, a partir de (47) segue que

$$\left| \|Ax_{k_\delta} - b^\delta\| - \|Ax_\alpha - b^\delta\| \right| \leq \epsilon\delta = \min \{c_3 - c_7, c_8 - c_4\} \delta,$$

sempre que $\delta_0 > 0$ é suficientemente pequeno. Isso junto com (41), implica (48).

Agora, da Proposição 2.3.6 concluímos que $x_\alpha \rightarrow x^\dagger$ quando $\delta \rightarrow 0$. Em particular, x_α é limitada. De (40) temos que,

$$\|x_{k_\delta} - x_\alpha\| \leq \Lambda_\alpha^{k_\delta} \|x_0 - x_\alpha\| \leq f(\delta) \|x_0 - x_\alpha\| \rightarrow 0, \quad (49)$$

quando $\delta \rightarrow 0$. Portanto, $x_{k_\delta} \rightarrow x^\dagger$ quando $\delta \rightarrow 0$, o que encerra a prova. \square

Teorema 3.2.3. *Suponha que a Hipótese 3.0.1 e que P seja fortemente convexa. Fixe $x_0 \in X$, $1 < c_3 < c_4$ e defina k_δ como em (42), com f satisfazendo $f(\delta) \leq \hat{c}\delta$ para todo $\delta > 0$ suficientemente pequeno onde $\hat{c} > 0$ é a constante do Teorema 3.2.2. Se a condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ é satisfeita, então temos a taxa ótima de convergência*

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

Demonstração. Das desigualdades (48), da condição de fonte $\partial P(x^\dagger) \cap R(A^*) \neq \emptyset$ e da Proposição 2.3.6 obtemos $\|x_\alpha - x^\dagger\| = \mathcal{O}(\sqrt{\delta})$. Esse fato, junto com (49) e $f(\delta) \leq \hat{c}\delta$, nos permite concluir o resultado a partir de uma desigualdade triangular. \square

¹ Observe que $1 < c_7 \leq c_8$.

4 MÉTODO DE TIKHONOV ITERADO COM MINIMIZAÇÃO INEXATA

Neste Capítulo vamos apresentar uma versão adaptada do algoritmo apresentado em [6], este por sua vez é uma versão modificada do método de Tikhonov iterado. Em [19] é feita uma generalização para espaços de Banach do algoritmo apresentado em [6]. Em [6] é proposto um método iterativo baseado no método iterativo não estacionário de Tikhonov para resolver problemas inversos lineares com a escolha a posteriori dos parâmetros de regularização.

4.1 APRESENTAÇÃO DO MÉTODO

Seja $P : X \rightarrow \overline{\mathbb{R}}$ uma função uniformemente convexa. A $(k + 1)$ -ésima iteração do método apresentado em [19] consiste em computar o parâmetro de regularização α_k^δ de modo que o único minimizador (veja a Observação 2.3.2) $x_{\alpha_k^\delta}$ do funcional de Tikhonov

$$T_{\alpha_k^\delta}^\delta(x) = \frac{1}{2} \|Ax - b^\delta\|^2 + \alpha_k^\delta D_{\xi_k^\delta}(x, x_k^\delta), \quad (50)$$

satisfaça

$$(1 - \eta_0)\delta + \eta_0 \|Ax_k^\delta - b^\delta\| \leq \|Ax_{\alpha_k^\delta} - b^\delta\| \leq (1 - \eta_1)\delta + \eta_1 \|Ax_k^\delta - b^\delta\|. \quad (51)$$

Aqui, $0 < \eta_0 < \eta_1 < 1$ são constantes fixadas, $\xi_k^\delta \in \partial P(x_k^\delta)$ e $D_{\xi_k^\delta}(\cdot, x_k^\delta)$ é a distância de Bregman associada a P . O passo seguinte é dado por $x_{k+1}^\delta := x_{\alpha_k^\delta}$. Em [19] é provado que sob condições adequadas, esse algoritmo gera um método de regularização. A iteração desse método acaba após k_δ passos, onde

$$k_\delta := \min \{k \in \mathbb{N} : \|Ax_k - b^\delta\| \leq \tau\delta, \text{ para algum } \tau > 1\}. \quad (52)$$

O critério de parada acima é chamado de princípio da discrepância. Esse algoritmo tem se mostrado eficiente, sem falar que (51) nos fornece um decaimento geométrico do resíduo. Mais precisamente, enquanto o princípio da discrepância não for atingido, temos

$$\|Ax_k^\delta - b^\delta\| - \delta \leq \eta_1^k \left(\|Ax_0^\delta - b^\delta\| - \delta \right). \quad (53)$$

A dificuldade maior na implementação do método em [19] está no fato de que a sequência dos parâmetros de regularização α_k^δ dado em (51) é definida de forma implícita. Esse fato nos leva a duas situações; Uma é que o cálculo dos parâmetros α_k^δ pode exigir a minimização de muitos funcionais em (50), se tornando caro computacionalmente, a outra é que devemos assumir que a minimização de (50) seja feita de forma exata, o que se torna impossível na prática. O primeiro problema ainda pode ser remediado por estratégias apresentadas em [19]. O segundo fato gerado pela minimização inexata é mais sério, porque aqui ao testarmos a desigualdade (51) com

um minimizador aproximado, as chances de termos imprecisão no cálculo de α_k^δ são grandes.

Nesta seção vamos apresentar uma adaptação do algoritmo estudado em [6]. Aqui o minimizador do funcional de Tikhonov (50) é computado de forma inexata em cada iteração. Para fazer isso, vamos recorrer a uma iteração interna para cada etapa do método de Tikhonov iterado a fim de aproximar o minimizador de (50).

Vamos definir

$$\tilde{x}_{k+1}^\delta := \arg \min_{x \in X} T_{\alpha_k^\delta}^\delta(x), \quad (54)$$

como o minimizador exato do funcional dado em (50). Vamos também precisar de uma versão modificada da Hipótese 3.0.1.

Hipótese 4.1.1. Assuma que para cada $x_k^\delta \in X$, cada número real $\alpha_k^\delta > 0$ e cada $z_{k,0}^\delta \in X$, uma sequência $z_{k,n}^\delta$ satisfazendo

$$\|z_{k,n}^\delta - \tilde{x}_{k+1}^\delta\| \leq \Lambda_k \|z_{k,n-1}^\delta - \tilde{x}_{k+1}^\delta\|, \text{ para todo } n \in \mathbb{N},$$

é conhecida, onde $0 < \Lambda_k < 1$.

Nosso algoritmo vai funcionar com uma iteração externa e uma interna: A iteração externa representa a versão inexata do método de Tikhonov iterado não estacionário, onde x_{k+1}^δ é o minimizador inexato de (50). Já a iteração interna consiste em gerar a sequência $z_{k,n}^\delta$ (que fornece os minimizadores inexatos) satisfazendo a Hipótese 4.1.1 normalmente através de um método de otimização convexa (veja a Seção 4.3).

O algoritmo vai começar com $x_0^\delta \in X$. Assumindo que a k -ésima iteração externa x_k^δ é conhecida, escolhemos $\xi_k^\delta \in \partial P(x_k^\delta)$ e então a partir de $z_{k,0}^\delta := x_k^\delta$ começamos a iteração interna gerando a sequência $z_{k,n}^\delta$ satisfazendo a Hipótese 4.1.1. O critério de parada da iteração interna será dado por

$$n_k^\delta := \left\lfloor \log_{\Lambda_k} \left(\min\{1, \alpha_k^\delta\} f(\delta) \right) \right\rfloor, \quad (55)$$

onde f é como em (40). A atualização da iteração externa vem dada por $x_{k+1}^\delta := z_{k,n_k^\delta}^\delta$.

A sequência dos parâmetros de regularização α_k^δ é escolhida de forma similar a estratégia a posteriori (51): fixe $0 < \eta < \bar{\eta} < 1$ e escolha $\alpha_k^\delta > 0$ de modo que a desigualdade

$$(1 - \eta)\delta + \eta \|Ax_k^\delta - b^\delta\| \leq \|Az_{k,n_k^\delta}^\delta - b^\delta\| \leq (1 - \bar{\eta})\delta + \bar{\eta} \|Ax_k^\delta - b^\delta\| \quad (56)$$

seja satisfeita. Por fim a iteração externa termina com k_δ estabelecido em (52). Dessa forma, $x_{k_\delta}^\delta$ aproxima a solução do problema inverso (1).

Segundo a Definição 2.1.2, observe que, fixada a iteração x_k^δ com $k < k_\delta$, existe, em geral, mais de uma opção para se escolher $\xi_k^\delta \in \partial P(x_k^\delta)$. Em certos algoritmos ξ_k^δ é escolhido a partir de uma regra específica, veja [19]. No passo [2d] do Algoritmo 2

abaixo, ξ_k^δ é escolhido de forma aleatória. A partir do Teorema 4.2.7 assumimos que P é Fréchet diferenciável e nesse caso ξ_k^δ se torna unicamente definido.

Nossa estratégia do método de Tikhonov com minimização inexata está representada pelo algoritmo abaixo.

INPUT: $0 < \eta < \bar{\eta} < 1$, $\tau > 1$, $x_0^\delta \in X$, $\xi_0^\delta \in \partial P(x_0^\delta)$, $\delta > 0$ e b^δ como em (2).
 OUTPUT: x_k^δ satisfazendo $\|Ax_k^\delta - b^\delta\| \leq \tau\delta$.

[1] $k = 0$;
 [2] Enquanto $\|Ax_k^\delta - b^\delta\| > \tau\delta$, faça
 [2a] $z_{k,0}^\delta = x_k^\delta$;
 [2b] Encontre $\alpha_k^\delta > 0$ tal que (56) seja satisfeito;
 [2c] $x_{k+1}^\delta = z_{k,n_k}^\delta$;
 [2d] Escolha $\xi_{k+1}^\delta \in \partial P(x_{k+1}^\delta)$;
 [2e] $k = k + 1$.

Algoritmo 2 – Método de Tikhonov Iterado com Minimização Inexata

4.2 ANÁLISE DE CONVERGÊNCIA

Nosso primeiro desafio para estabelecer a análise de convergência passa por mostrar que o método está bem definido e que o mesmo acaba após um número finito de passos. Vamos inicialmente provar que a iteração interna está bem definida, o que é equivalente a provar que se o princípio da discrepância não for atingindo no passo k , então existe um parâmetro $\alpha_k^\delta > 0$ tal que z_{k,n_k}^δ satisfaz (56). No decorrer da análise vamos sempre assumir que o ponto inicial $x_0^\delta = x_0$ e o vetor $\xi_0^\delta = \xi_0$ são sempre os mesmos independente do nível de ruídos $\delta > 0$. Além disso, vamos também assumir que P é fortemente convexa com módulo β .

Lema 4.2.1. Fixe as constantes $\eta < \eta_0 \leq \eta_1 < \bar{\eta}$. Se a função f em (55) satisfaz $f(\delta) \leq \hat{c}$, para todo $\delta > 0$, onde

$$\hat{c} = \frac{\beta \left(1 - \frac{1}{\tau}\right)}{\|A\|^2} \min \{ \eta_0 - \eta, \bar{\eta} - \eta_1 \}, \quad (57)$$

e τ dado em (52), então o passo [2b] do Algoritmo 2 está bem definido.

Demonstração. Suponha que o princípio da discrepância não tenha sido atingido até a iteração k . Então, $\|Ax_k^\delta - b^\delta\| > \tau\delta > \delta$. Uma vez que $0 < \eta_0 \leq \eta_1 < 1$, segue que o intervalo

$$\left[(1 - \eta_0)\delta + \eta_0 \|Ax_k^\delta - b^\delta\|, (1 - \eta_1)\delta + \eta_1 \|Ax_k^\delta - b^\delta\| \right] \quad (58)$$

é não vazio. Segue de [19, Lema 3.2], que existe $\alpha_k^\delta > 0$ tal que o minimizador exato \tilde{x}_{k+1}^δ de (50) tem resíduo $\|A\tilde{x}_{k+1}^\delta - b^\delta\|$ pertencente ao intervalo (58). Agora, a partir de (54) e de (50) temos

$$0 = A^*(A\tilde{x}_{k+1}^\delta - b^\delta) + \alpha_k^\delta(\tilde{\xi}_{k+1}^\delta - \xi_k^\delta), \quad (59)$$

onde $\tilde{\xi}_{k+1}^\delta \in \partial P(\tilde{x}_{k+1}^\delta)$. Multiplicando essa última igualdade por $x_k^\delta - \tilde{x}_{k+1}^\delta$ e usando a Observação 2.1.7, temos

$$\begin{aligned} \alpha_k^\delta \beta \|x_k^\delta - \tilde{x}_{k+1}^\delta\|^2 &\leq \alpha_k^\delta \langle \xi_k^\delta - \tilde{\xi}_{k+1}^\delta, x_k^\delta - \tilde{x}_{k+1}^\delta \rangle = \langle A^*(A\tilde{x}_{k+1}^\delta - b^\delta), x_k^\delta - \tilde{x}_{k+1}^\delta \rangle \\ &\leq \|A\| \|A\tilde{x}_{k+1}^\delta - b^\delta\| \|x_k^\delta - \tilde{x}_{k+1}^\delta\| \leq \|A\| \|Ax_k^\delta - b^\delta\| \|x_k^\delta - \tilde{x}_{k+1}^\delta\|, \end{aligned}$$

a última desigualdade acima segue do fato de $T_{\alpha_k^\delta}^\delta(\tilde{x}_{k+1}^\delta) \leq T_{\alpha_k^\delta}^\delta(x_k^\delta)$. Então,

$$\begin{aligned} \|x_k^\delta - \tilde{x}_{k+1}^\delta\| &\leq \frac{\|A\|}{\beta \alpha_k^\delta} \left(\frac{\|Ax_k^\delta - b^\delta\|}{\|Ax_k^\delta - b^\delta\| - \delta} \right) (\|Ax_k^\delta - b^\delta\| - \delta) \\ &\leq \frac{\|A\|}{\beta \left(1 - \frac{1}{\tau}\right) \alpha_k^\delta} (\|Ax_k^\delta - b^\delta\| - \delta), \end{aligned}$$

a última desigualdade acima é devido a $\delta < \frac{1}{\tau} \|Ax_k^\delta - b^\delta\|$. Dessa forma, tendo em vista a Hipótese 4.1.1, (55) e (57), segue que

$$\begin{aligned} \left| \|Az_{k,n_k}^\delta - b^\delta\| - \|A\tilde{x}_{k+1}^\delta - b^\delta\| \right| &\leq \|A\| \|z_{k,n_k}^\delta - \tilde{x}_{k+1}^\delta\| \leq \|A\| \wedge_k^{n_k} \|z_{k,0}^\delta - \tilde{x}_{k+1}^\delta\| \\ &\leq \|A\| \alpha_k^\delta f(\delta) \|x_k^\delta - \tilde{x}_{k+1}^\delta\| \\ &\leq f(\delta) \frac{\|A\|^2}{\beta \left(1 - \frac{1}{\tau}\right)} (\|Ax_k^\delta - b^\delta\| - \delta) \\ &\leq \min \{ \eta_0 - \eta, \bar{\eta} - \eta_1 \} (\|Ax_k^\delta - b^\delta\| - \delta). \end{aligned}$$

Uma vez que $\|A\tilde{x}_{k+1}^\delta - b^\delta\|$ pertence ao intervalo (58), a norma $\|Az_{k,n_k}^\delta - b^\delta\|$ pertence ao intervalo

$$\left[(1 - \eta)\delta + \eta \|Ax_k^\delta - b^\delta\|, (1 - \bar{\eta})\delta + \bar{\eta} \|Ax_k^\delta - b^\delta\| \right],$$

e isso completa a prova. \square

Observação 4.2.2. Uma vez que f em (55) satisfaz (40), a condição $f(\delta) \leq \hat{c}$ do Lema 4.2.1 é satisfeita para qualquer $\delta > 0$ pequeno o suficiente. Note também que em (55) podemos trocar f por $\min \{ f(\delta), \hat{c} \}$, assim ainda teremos $f(\delta) \leq \hat{c}$, para todo $\delta > 0$.

O próximo resultado nos diz que o Algoritmo 2 termina após um número finito de iterações.

Proposição 4.2.3. *O índice de parada k_δ definido em (52) é finito para todo $\delta > 0$ suficientemente pequeno.*

Demonstração. De fato, fixe $\delta > 0$. Segue do passo [2b] do Algoritmo 2 que

$$\|Ax_{k+1}^\delta - b^\delta\| - \delta = \|Az_{k,n_k}^\delta - b^\delta\| - \delta \leq \bar{\eta} \left(\|Ax_k^\delta - b^\delta\| - \delta \right)$$

para $0 \leq k < k_\delta$. Assim,

$$\|Ax_k^\delta - b^\delta\| - \delta \leq \bar{\eta}^k \left(\|Ax_0 - b^\delta\| - \delta \right), \text{ para todo } k \leq k_\delta. \quad (60)$$

Por outro lado, $\bar{\eta}^{\bar{k}} \left(\|Ax_0 - b^\delta\| - \delta \right) \leq (\tau - 1) \delta$ para \bar{k} grande o bastante, portanto $\|Ax_{\bar{k}}^\delta - b^\delta\| \leq \tau \delta$. \square

Corolário 4.2.4. *Seja x_k^δ a sequência gerada pelo Algoritmo 2 com $\delta > 0$ suficientemente pequeno. Então,*

$$k_\delta \leq |\ln \bar{\eta}|^{-1} \ln \left(\frac{\|Ax_0 - b^\delta\| - \delta}{(\tau - 1) \delta} \right) + 1, \quad (61)$$

para $k = 0, \dots, k_\delta$.

Demonstração. A partir de (52) segue que $\tau \delta < \|Ax_k^\delta - b^\delta\|$, para $k = 0, \dots, k_\delta - 1$. Essa desigualdade (para $k = k_\delta - 1$), junto com a desigualdade (60) na Proposição 4.2.3 nos dá

$$(\tau - 1) \delta < \|Ax_{k_\delta - 1}^\delta - b^\delta\| - \delta \leq \bar{\eta}^{k_\delta - 1} (\|Ax_0^\delta - b^\delta\| - \delta),$$

isso implica (61). \square

Vamos agora nos encaminhar para o estudo da propriedade de estabilidade da sequência gerada pelo Algoritmo 2. Como observado anteriormente, dentre as possibilidades para se escolher o passo seguinte, x_{k+1}^δ , vamos chamar de sucessor de x_k^δ qualquer uma dessas possibilidades.

Vamos introduzir o conceito de sequência sem ruídos no sentido apresentado em [19, Teorema 4.6]. Para o caso de dados exatos ($\delta = 0$) a minimização do funcional (50) no Algoritmo 2 é realizada de forma exata ($n_k = \infty$), nesse caso possivelmente teremos $k_\delta = \infty$.

Definição 4.2.5. *Uma sequência sem ruídos é uma sequência x_k em X tal que*

1. $x_0 = x_0^\delta$ e $\xi_0 = \xi_0^\delta$;
2. para cada $k \in \mathbb{N}$, existe um número $\alpha_k > 0$, tal que o minimizador exato, x_{α_k} , do funcional de Tikhonov

$$T_{\alpha_k}(x) = \frac{1}{2} \|Ax - b\|^2 + \alpha_k D_{\xi_k}(x, x_k), \quad (62)$$

satisfaz a desigualdade

$$\eta_2 \|Ax_k - b\| \leq \|Ax_{\alpha_k} - b\| \leq \eta_3 \|Ax_k - b\|, \quad (63)$$

onde $0 < \eta_2 \leq \eta_3 < 1$ são constantes pré-fixadas.

3. $x_{k+1} = x_{\alpha_k}$ e $\xi_{k+1} = \xi_k - \frac{1}{\alpha_k} A^* (Ax_{k+1} - b)$, para todo $k \in \mathbb{N}$.

Observação 4.2.6. Se (x_k) é uma sequência sem ruídos, vamos chamar x_{k+1} do sucessor de x_k para cada $k \geq 0$. Note que esse conceito de sequência sem ruídos coincide com a sequência analisada em [19, Teorema 4.6]. Uma consequência desse teorema é que $D_{\xi_k}(x^\dagger, x_k) \rightarrow 0$ quando $k \rightarrow \infty$, onde x^\dagger é a única solução do problema inverso (1) satisfazendo

$$D_{\xi_0}(x^\dagger, x_0) \leq D_{\xi_0}(x^*, x_0), \text{ para todo } x^* \text{ solução de (1).}$$

Em [19, Lema 4.5], é provado a existência e unicidade de x^\dagger . Observe que, uma vez que P é uniformemente convexa, teremos $x_k \rightarrow x^\dagger$, quando $k \rightarrow \infty$.

Na última Definição 4.2.5, o vetor x_{k+1} é o minimizador exato de (62). Portanto, segue que

$$0 \in \partial T_{\alpha_k}(x_{k+1}) = A^* (Ax_{k+1} - b) + \alpha_k (\partial P(x_{k+1}) - \xi_k),$$

consequentemente, $\xi_{k+1} = \xi_k - \frac{1}{\alpha_k} A^* (Ax_{k+1} - b) \in \partial P(x_{k+1})$.

No próximo teorema, fixado $k \in \mathbb{N}$ vamos observar o comportamento da família $(x_k^\delta)_{\delta > 0}$ para δ suficientemente pequeno.

Teorema 4.2.7 (Estabilidade). *Seja δ_j uma sequência de números reais positivos convergindo para zero e assumamos que as sequências $x_k^{\delta_j}$ estão fixadas para cada $j \in \mathbb{N}$ com $0 \leq k \leq k_{\delta_j}$ onde $x_{k+1}^{\delta_j}$ é o sucessor de $x_k^{\delta_j}$ para $k = 0, \dots, k_{\delta_j} - 1$. Além disso, assumamos que P é Fréchet diferenciável e as constantes η_0 e η_1 no Lema 4.2.1 sejam dadas por*

$$\eta_0 = \eta + \epsilon_0 \text{ e } \eta_1 = \bar{\eta} - \epsilon_1,$$

onde $0 < \epsilon_0 < \min\{\eta, \bar{\eta} - \eta\}/2$ e $0 < \epsilon_1 < \min\{1 - \bar{\eta}, \bar{\eta} - \eta\}/2$. Se x_0 não é solução de (1), então vai existir uma sequência sem ruídos x_k , e constantes η_2 e η_3 na Definição 4.2.5 dadas por

$$\eta_2 = \eta - \epsilon_0 \text{ e } \eta_3 = \bar{\eta} + \epsilon_1, \quad (64)$$

tal que, para todo $k \in \mathbb{N}$ fixado, existe uma subsequência δ_{j_m} (dependendo de k) de δ_j satisfazendo

$$x_\ell^{\delta_{j_m}} \rightarrow x_\ell, \text{ e } \xi_\ell^{\delta_{j_m}} \rightarrow \xi_\ell, \text{ quando } m \rightarrow \infty, \text{ para } \ell = 0, \dots, k.$$

Demonstração. Note que, de acordo com as definições de ϵ_0 e ϵ_1 , temos

$$(\eta_0, \eta_1) \subset (\eta, \bar{\eta}) \subset (\eta_2, \eta_3) \subset (0, 1).$$

Vamos usar um argumento de indução: uma vez que $x_0^\delta = x_0$ e $\xi_0^\delta = \xi_0$ para cada $\delta > 0$, o resultado é óbvio para $k = 0$. O argumento consiste na escolha sucessiva de uma subsequência da atual subsequência, vamos aqui denotar qualquer subsequência de δ_j também por δ_j . Suponha que o resultado seja válido para algum $k \in \mathbb{N}$. Então, existe uma subsequência δ_j satisfazendo

$$x_\ell^{\delta_j} \rightarrow x_\ell, \text{ e } \xi_\ell^{\delta_j} \rightarrow \xi_\ell, \text{ quando } j \rightarrow \infty, \text{ para } \ell = 0, \dots, k, \quad (65)$$

onde x_ℓ é o sucessor de $x_{\ell-1}$ e $\xi_\ell = \xi_{\ell-1} - \frac{1}{\alpha_{\ell-1}} A^* (Ax_\ell - b)$, para $\ell = 1, \dots, k$.

Agora vamos construir um sucessor de x_k , i.e., vamos encontrar $\alpha_k > 0$, definir $x_{k+1} = \arg \min T_{\alpha_k}(x)$ (veja (62)) e mostrar que (63) é satisfeito para $x_{\alpha_k} = x_{k+1}$. Em seguida, vamos provar que para alguma subsequência temos que $x_{k+1}^{\delta_j} \rightarrow x_{k+1}$, e isso encerrará a prova.

Seja

$$\alpha_k := \liminf_{j \rightarrow \infty} \alpha_k^{\delta_j}, \quad (66)$$

vamos provar que $\alpha_k > 0$. De fato, suponha que $\alpha_k = 0$. Segue de (65) que $D_{\xi_k^\delta}(x^\dagger, x_k^{\delta_j})$ é limitada em j , assim existe uma constante $M > 0$ tal que $D_{\xi_k^\delta}(x^\dagger, x_k^{\delta_j}) \leq M$, para todo $j \in \mathbb{N}$. Agora, observe que

$$0 \leq T_{\alpha_k^{\delta_j}}^{\delta_j}(x_{k+1}^{\delta_j}) \leq T_{\alpha_k^{\delta_j}}^{\delta_j}(x^\dagger) \leq \delta_j^2 + \alpha_k^{\delta_j} D_{\xi_k^\delta}(x^\dagger, x_k^{\delta_j}) \leq \delta_j^2 + \alpha_k^{\delta_j} M,$$

isso implica

$$\liminf_{j \rightarrow \infty} \|Ax_{k+1}^{\delta_j} - b^{\delta_j}\| \leq \liminf_{j \rightarrow \infty} \sqrt{2T_{\alpha_k^{\delta_j}}^{\delta_j}(x_{k+1}^{\delta_j})} = 0. \quad (67)$$

Provaremos a seguir que

$$\eta_2(\|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j) \leq \|Ax_{k+1}^{\delta_j} - b^{\delta_j}\| - \delta_j, \quad (68)$$

para cada $j \in \mathbb{N}$. Conforme a prova do Lema 4.2.1,

$$\begin{aligned} \left| \|Ax_{k+1}^{\delta_j} - b^{\delta_j}\| - \|Ax_{k+1}^{\delta_j} - b^{\delta_j}\| \right| &\leq \min \{ \eta_0 - \eta, \bar{\eta} - \eta_1 \} \left(\|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \\ &= \min \{ \epsilon_0, \epsilon_1 \} \left(\|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \\ &= \min \{ \eta - \eta_2, \eta_3 - \bar{\eta} \} \left(\|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j \right). \end{aligned} \quad (69)$$

Agora, a norma $\|Ax_{k+1}^{\delta_j} - b^{\delta_j}\|$ pertence ao intervalo

$$\left[(1 - \eta)\delta_j + \eta \|Ax_k^{\delta_j} - b^{\delta_j}\|, (1 - \bar{\eta})\delta_j + \bar{\eta} \|Ax_k^{\delta_j} - b^{\delta_j}\| \right],$$

pois $x_{k+1}^{\delta_j}$ é um sucessor de $x_k^{\delta_j}$. Então, tendo em vista (69), obtemos

$$(1 - \eta_2)\delta_j + \eta_2\|Ax_k^{\delta_j} - b^{\delta_j}\| \leq \|A\tilde{x}_{k+1}^{\delta_j} - b^{\delta_j}\| \leq (1 - \eta_3)\delta_j + \eta_3\|Ax_k^{\delta_j} - b^{\delta_j}\|. \quad (70)$$

Isso prova (68).

Segue do passo [2b] do Algoritmo 2 que

$$\eta^k \left(\|Ax_0^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \leq \|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j. \quad (71)$$

Agora (68) junto com (71) contradiz o fato de x_0 não ser solução de (1), pois

$$\begin{aligned} 0 &\leq \eta_2 \eta^k \|Ax_0 - b\| = \eta_2 \liminf_{j \rightarrow \infty} \eta^k \left(\|Ax_0^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \\ &\leq \eta_2 \liminf_{j \rightarrow \infty} \left(\|Ax_k^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \leq \liminf_{j \rightarrow \infty} \left(\|A\tilde{x}_{k+1}^{\delta_j} - b^{\delta_j}\| - \delta_j \right) \stackrel{(67)}{=} 0, \end{aligned} \quad (72)$$

e isso implica que $\|Ax_0 - b\| = 0$. Assim, temos $\alpha_k > 0$. Escolhendo uma subsequência se necessário, podemos assumir que $\lim_{j \rightarrow \infty} \alpha_k^{\delta_j} = \alpha_k$. Isso nos permite definir $x_{k+1} := \arg \min_{x \in X} T_{\alpha_k}(x)$.

Em seguida vamos provar que

$$\tilde{x}_{k+1}^{\delta_j} \rightarrow x_{k+1}, \text{ quando } j \rightarrow \infty. \quad (73)$$

De fato, uma vez que $\tilde{x}_{k+1}^{\delta_j}$ e x_{k+1} são os minimizadores de (50) e (62) respectivamente, temos

$$0 = A^*(A\tilde{x}_{k+1}^{\delta_j} - b^{\delta_j}) + \alpha_k^{\delta_j}(\tilde{\xi}_{k+1}^{\delta_j} - \xi_k^{\delta_j}) \quad (74)$$

e

$$0 = A^*(Ax_{k+1} - b) + \alpha_k(\xi_{k+1} - \xi_k), \quad (75)$$

com $\tilde{\xi}_{k+1}^{\delta_j}$ o único elemento de $\partial P(\tilde{x}_{k+1}^{\delta_j})$ (pois P é Fréchet diferenciável). Subtraindo a equação (75) de (74) obtemos

$$\begin{aligned} 0 &= A^*A(\tilde{x}_{k+1}^{\delta_j} - x_{k+1}) + A^*(b - b^{\delta_j}) + \alpha_k^{\delta_j}(\tilde{\xi}_{k+1}^{\delta_j} - \xi_{k+1}) + \xi_{k+1}(\alpha_k^{\delta_j} - \alpha_k) \\ &\quad + \alpha_k \xi_k - \alpha_k^{\delta_j} \xi_k^{\delta_j}. \end{aligned} \quad (76)$$

Agora, multiplicando (76) por $\tilde{x}_{k+1}^{\delta_j} - x_{k+1}$ segue,

$$\begin{aligned} \alpha_k^{\delta_j} \langle \tilde{\xi}_{k+1}^{\delta_j} - \xi_{k+1}, \tilde{x}_{k+1}^{\delta_j} - x_{k+1} \rangle &= -\|A(\tilde{x}_{k+1}^{\delta_j} - x_{k+1})\|^2 + \langle b^{\delta_j} - b, A(\tilde{x}_{k+1}^{\delta_j} - x_{k+1}) \rangle \\ &\quad + (\alpha_k - \alpha_k^{\delta_j}) \langle \xi_{k+1}, \tilde{x}_{k+1}^{\delta_j} - x_{k+1} \rangle + \langle \alpha_k^{\delta_j} \xi_k^{\delta_j} - \alpha_k \xi_k, \tilde{x}_{k+1}^{\delta_j} - x_{k+1} \rangle. \end{aligned}$$

Segue da Observação 2.1.7 que,

$$\begin{aligned} \alpha_k^{\delta_j} \beta \|\tilde{x}_{k+1}^{\delta_j} - x_{k+1}\|^2 &\leq \alpha_k^{\delta_j} \langle \tilde{\xi}_{k+1}^{\delta_j} - \xi_{k+1}, \tilde{x}_{k+1}^{\delta_j} - x_{k+1} \rangle \leq \delta_j \|A\| \|\tilde{x}_{k+1}^{\delta_j} - x_{k+1}\| \\ &\quad + |\alpha_k^{\delta_j} - \alpha_k| \|\xi_{k+1}\| \|\tilde{x}_{k+1}^{\delta_j} - x_{k+1}\| + \|\alpha_k^{\delta_j} \xi_k^{\delta_j} - \alpha_k \xi_k\| \|\tilde{x}_{k+1}^{\delta_j} - x_{k+1}\|, \end{aligned}$$

isso implica,

$$\|\tilde{x}_{k+1}^{\delta_j} - x_{k+1}\| \leq \frac{1}{\beta \alpha_k^{\delta_j}} \left[\delta_j \|A\| + |\alpha_k^{\delta_j} - \alpha_k| \|\xi_{k+1}\| + \|\alpha_k^{\delta_j} \xi_k^{\delta_j} - \alpha_k \xi_k\| \right],$$

a partir dessa última desigualdade, de (65) e (66) segue (73).

De (73), (70) e de (65), segue que

$$\eta_2 \|Ax_k - b\| \leq \|Ax_{k+1} - b\| \leq \eta_3 \|Ax_k - b\|,$$

isso prova que x_{k+1} é um sucessor de x_k . Por fim vamos provar que $x_{k+1}^{\delta_j} \rightarrow x_{k+1}$ e $\xi_{k+1}^{\delta_j} \rightarrow \xi_{k+1}$, quando $j \rightarrow \infty$. Aos passos da prova do Lema 4.2.1, segue que

$$\begin{aligned} \|x_{k+1}^{\delta_j} - \tilde{x}_{k+1}^{\delta_j}\| &= \|z_{k, n_k^{\delta_j}}^{\delta_j} - \tilde{x}_{k+1}^{\delta_j}\| \leq \Lambda_k^{\eta_k^{\delta_j}} \|x_k^{\delta_j} - \tilde{x}_{k+1}^{\delta_j}\| \\ &\leq \alpha_k^{\delta_j} f(\delta_j) \frac{\|A\|}{\beta \alpha_k^{\delta_j}} \|Ax_k^{\delta_j} - b^{\delta_j}\| \rightarrow 0, \end{aligned}$$

quando $j \rightarrow \infty$. Portanto, isso junto com (73), nos permite concluir que $x_{k+1}^{\delta_j} \rightarrow x_{k+1}$, quando $j \rightarrow \infty$. Finalmente, como P é Fréchet diferenciável, $\xi_{k+1}^{\delta_j} = \partial P(x_{k+1}^{\delta_j}) \rightarrow \partial P(x_{k+1}) = \xi_{k+1}$, quando $j \rightarrow \infty$. \square

Como consequência do Teorema 4.2.7 obtemos o seguinte resultado.

Corolário 4.2.8. *Seja δ_j uma seqüência positiva convergindo para zero e assumamos que as seqüências $x_k^{\delta_j}$ estão fixadas para cada $j \in \mathbb{N}$, com $0 \leq k \leq k_{\delta_j}$ onde $x_{k+1}^{\delta_j}$ é um sucessor de $x_k^{\delta_j}$ para $k = 0, \dots, k_{\delta_j} - 1$.*

Além disso, assumamos que $k_{\delta_j} \rightarrow \infty$, quando $j \rightarrow \infty$. Então existe uma seqüência sem ruídos x_k e uma subsequência δ_{j_m} tal que

$$\|x_k^{\delta_{j_m}} - x_k\| \leq \frac{1}{k}, \quad \forall k \in \mathbb{N}, \quad \forall m \geq k. \quad (77)$$

Demonstração. Considere x_k a seqüência sem ruídos construída no Teorema 4.2.7. Usaremos indução para provar (77). Pelo Teorema 4.2.7, existe uma subsequência, denota apenas por δ_j , tal que $x_1^{\delta_j} \rightarrow x_1$, quando $j \rightarrow \infty$. Escolha $j_1 \in \mathbb{N}$ de modo que

$$j \geq j_1 \implies \|x_1^{\delta_j} - x_1\| \leq 1.$$

Agora, uma vez que a subsequência corrente é positiva e converge a zero, aplicando o Teorema 4.2.7 novamente, podemos extrair outra subsequência, também denotada por δ_j tal que $x_2^{\delta_j} \rightarrow x_2$, quando $j \rightarrow \infty$. Fixe $j_2 > j_1$ de modo que

$$j \geq j_2 \implies \|x_2^{\delta_j} - x_2\| \leq \frac{1}{2}.$$

Procedendo dessa maneira obtemos uma seqüência $j_1 < j_2 < \dots < j_m < \dots$, tal que (77) acontece. \square

Observação 4.2.9. A partir do Corolário 4.2.8, podemos obter uma subsequência δ_{j_m} tal que a sequência $x_k^{\delta_{j_m}}$ é uniformemente limitada em k e m . Isso é verdade pois, para a sequência sem ruídos x_k do Corolário 4.2.8 temos que $x_k \rightarrow x^\dagger$, quando $k \rightarrow \infty$, veja a Observação 4.2.6. Assim, existe $C > 0$ tal que $\|x_k - x^\dagger\| \leq C$, para todo $k \in \mathbb{N}$. Então, para a subsequência δ_{j_m} dada no Corolário 4.2.8, temos

$$\|x_k^{\delta_{j_m}} - x^\dagger\| \leq \|x_k^{\delta_{j_m}} - x_k\| + \|x_k - x^\dagger\| \leq \frac{1}{k} + C \leq 1 + C,$$

para todo $k \in \mathbb{N}$ e para $m \geq k$. Com isso, $\lim_{m \rightarrow \infty} \|x_k^{\delta_{j_m}} - x^\dagger\| \leq 1 + C$, para todo $k \in \mathbb{N}$.

Então, existe $N \in \mathbb{N}$ tal que $m \geq N$ implica que $\|x_k^{\delta_{j_m}} - x^\dagger\| \leq 2 + C$, para todo $k \in \mathbb{N}$. Considerando a nova subsequência δ_{j_m} , a partir de $m = N$, temos que $\|x_k^{\delta_{j_m}} - x^\dagger\| \leq 2 + C$, para todo $k, m \in \mathbb{N}$.

Proposição 4.2.10 (Monotonia). Suponha que P é Fréchet diferenciável com gradiente L -Lipschitz. Assuma que a função f em (55) satisfaz $f(\delta) \leq \widehat{C}\delta$, para cada $\delta > 0$ suficientemente pequeno, onde

$$\widehat{C} := \frac{\beta\eta_2(1-1/\tau)}{L\|A\|_c} > 0, \quad (78)$$

com η_2 como em (64) e $c > 0$ uma cota superior para $\|x_k^\delta - x^\dagger\|$ (garantida na Observação 4.2.9). Então, para cada $\delta > 0$ suficientemente pequeno,

$$D_{\xi_{k+1}^\delta}(x^\dagger, x_{k+1}^\delta) \leq D_{\xi_k^\delta}(x^\dagger, x_k^\delta), \quad \text{para todo } k < k_\delta. \quad (79)$$

Demonstração. Seja $\tilde{\xi}_{k+1}^\delta = \nabla P(\tilde{x}_{k+1}^\delta)$. Assim, uma vez que \tilde{x}_{k+1}^δ satisfaz (70), e tendo em vista que $\delta < \frac{1}{\tau}\|Ax_k^\delta - b^\delta\|$ para $k < k_\delta$, segue que

$$\begin{aligned} \langle \tilde{\xi}_{k+1}^\delta - \xi_k^\delta, \tilde{x}_{k+1}^\delta - x^\dagger \rangle &= \frac{1}{\alpha_k^\delta} \langle A^*(b^\delta - A\tilde{x}_{k+1}^\delta), \tilde{x}_{k+1}^\delta - x^\dagger \rangle = \frac{1}{\alpha_k^\delta} \langle b^\delta - A\tilde{x}_{k+1}^\delta, A\tilde{x}_{k+1}^\delta - b \rangle \\ &= \frac{1}{\alpha_k^\delta} \left[\langle b^\delta - A\tilde{x}_{k+1}^\delta, A\tilde{x}_{k+1}^\delta - b^\delta \rangle + \langle b^\delta - A\tilde{x}_{k+1}^\delta, b^\delta - b \rangle \right] \\ &\leq -\frac{1}{\alpha_k^\delta} \|A\tilde{x}_{k+1}^\delta - b^\delta\| \left(\|A\tilde{x}_{k+1}^\delta - b^\delta\| - \delta \right) \\ &\leq -\frac{\eta_2}{\alpha_k^\delta} \|A\tilde{x}_{k+1}^\delta - b^\delta\| \left(\|Ax_k^\delta - b^\delta\| - \delta \right) \\ &\leq -\frac{\eta_2 \left(1 - \frac{1}{\tau}\right)}{\alpha_k^\delta} \|A\tilde{x}_{k+1}^\delta - b^\delta\| \|Ax_k^\delta - b^\delta\|. \end{aligned}$$

A partir da identidade dos três pontos (15), temos

$$\begin{aligned} D_{\tilde{\xi}_{k+1}^\delta}(x^\dagger, \tilde{x}_{k+1}^\delta) - D_{\xi_k^\delta}(x^\dagger, x_k^\delta) &= -D_{\xi_k^\delta}(\tilde{x}_{k+1}^\delta, x_k^\delta) + \langle \tilde{\xi}_{k+1}^\delta - \xi_k^\delta, \tilde{x}_{k+1}^\delta - x^\dagger \rangle \\ &\leq -\frac{\eta_2 \left(1 - \frac{1}{\tau}\right)}{\alpha_k^\delta} \|A\tilde{x}_{k+1}^\delta - b^\delta\| \|Ax_k^\delta - b^\delta\|. \quad (80) \end{aligned}$$

Novamente usando (15) e aplicando a técnica usada no Lema 4.2.1, segue que

$$\begin{aligned} D_{\xi_{k+1}^\delta}(x^\dagger, x_{k+1}^\delta) - D_{\tilde{\xi}_{k+1}^\delta}(x^\dagger, \tilde{x}_{k+1}^\delta) &= -D_{\tilde{\xi}_{k+1}^\delta}(x_{k+1}^\delta, \tilde{x}_{k+1}^\delta) + \langle \xi_{k+1}^\delta - \tilde{\xi}_{k+1}^\delta, x_{k+1}^\delta - x^\dagger \rangle \\ &\leq \|\xi_{k+1}^\delta - \tilde{\xi}_{k+1}^\delta\| \|x_{k+1}^\delta - x^\dagger\| \leq L \|x_{k+1}^\delta - \tilde{x}_{k+1}^\delta\| \|x_{k+1}^\delta - x^\dagger\| \\ &\leq \frac{L \|A\|}{\beta \alpha_k^\delta} f(\delta) \|Ax_k^\delta - b^\delta\| \|x_{k+1}^\delta - x^\dagger\| \leq \frac{c_0}{\alpha_k^\delta} f(\delta) \|Ax_k^\delta - b^\delta\|, \end{aligned}$$

onde $c_0 = \frac{L \|A\| c}{\beta} > 0$. De (80) temos,

$$\begin{aligned} D_{\xi_{k+1}^\delta}(x^\dagger, x_{k+1}^\delta) - D_{\xi_k^\delta}(x^\dagger, x_k^\delta) &= D_{\xi_{k+1}^\delta}(x^\dagger, x_{k+1}^\delta) - D_{\tilde{\xi}_{k+1}^\delta}(x^\dagger, \tilde{x}_{k+1}^\delta) \\ &\quad + D_{\tilde{\xi}_{k+1}^\delta}(x^\dagger, \tilde{x}_{k+1}^\delta) - D_{\xi_k^\delta}(x^\dagger, x_k^\delta) \\ &\leq \frac{c_0}{\alpha_k^\delta} \|Ax_k^\delta - b^\delta\| \left[f(\delta) - \frac{\eta_2 \left(1 - \frac{1}{\tau}\right)}{c_0} \|A\tilde{x}_{k+1}^\delta - b^\delta\| \right]. \end{aligned}$$

Agora, segue de (70) que

$$\|A\tilde{x}_{k+1}^\delta - b^\delta\| \geq (1 - \eta_2)\delta + \eta_2 \|Ax_k^\delta - b^\delta\| \geq (1 - \eta_2)\delta + \eta_2\delta = \delta,$$

e portanto, $D_{\xi_{k+1}^\delta}(x^\dagger, x_{k+1}^\delta) - D_{\xi_k^\delta}(x^\dagger, x_k^\delta) \leq \frac{c_0}{\alpha_k^\delta} \|Ax_k^\delta - b^\delta\| [f(\delta) - \widehat{C}\delta] \leq 0$. \square

Observação 4.2.11. A Proposição 4.2.10 vale não apenas para o x^\dagger mas para qualquer x^* solução de (1).

Observação 4.2.12. Em particular, quando $P(x) = \frac{1}{2} \|x\|^2$ temos $\|\nabla P(x) - \nabla P(y)\| = \|x - y\|$, e portanto P tem gradiente L -Lipschitz com $L = 1$.

Teorema 4.2.13 (Regularização). Assuma que P é Fréchet diferenciável com gradiente L -Lipschitz. Seja δ_j uma sequência de números positivos convergindo para zero e assumo que as sequências $x_k^{\delta_j}$ estão fixadas para cada $j \in \mathbb{N}$, com $0 \leq k \leq k_{\delta_j}$ onde $x_{k+1}^{\delta_j}$ é um sucessor de $x_k^{\delta_j}$, para $k = 0, \dots, k_{\delta_j} - 1$. Além disso, suponha que $f(\delta) \leq \widehat{C}\delta$ para $\delta > 0$ pequeno o suficiente, com $\widehat{C} > 0$ definido com em (78). Se x_0 não é solução do problema $Ax = b$, então

$$\lim_{j \rightarrow \infty} x_{k_{\delta_j}}^{\delta_j} = x^\dagger. \quad (81)$$

Demonstração. Primeiro vamos mostrar que a sequência k_{δ_j} não possui subsequência convergente. De fato, se para a subsequência δ_{j_m} de δ_j for verdade que $k_{\delta_{j_m}} \rightarrow n$ quando $m \rightarrow \infty$, então como $k_{\delta_{j_m}} \in \mathbb{N}$ para todo $m \in \mathbb{N}$, temos que $k_{\delta_{j_m}} = n$ para m grande o suficiente. Segue do Teorema 4.2.7, que a subsequência $x_n^{\delta_{j_m}}$ tem ela própria uma subsequência (denota também por $x_n^{\delta_{j_m}}$) convergindo para x_n . Dessa forma,

$$\lim_{m \rightarrow \infty} x_{k_{\delta_{j_m}}}^{\delta_{j_m}} = \lim_{m \rightarrow \infty} x_n^{\delta_{j_m}} = x_n.$$

Mas isso contradiz o fato de x_0 não ser solução de $Ax = b$, pois

$$\begin{aligned} 0 &\leq \eta^n \|Ax_0 - b\| = \lim_{m \rightarrow \infty} \eta^n \left(\|Ax_0 - b^{\delta_{jm}}\| - \delta_{jm} \right) \leq \lim_{m \rightarrow \infty} \left(\|Ax_n^{\delta_{jm}} - b^{\delta_{jm}}\| - \delta_{jm} \right) \\ &= \lim_{m \rightarrow \infty} \left(\|Ax_{k_{\delta_{jm}}}^{\delta_{jm}} - b^{\delta_{jm}}\| - \delta_{jm} \right) \leq \lim_{m \rightarrow \infty} (\tau - 1)\delta_{jm} = 0. \end{aligned}$$

Portanto devemos ter $k_{\delta_j} \rightarrow \infty$, quando $j \rightarrow \infty$.

Vamos agora provar que cada subsequência de $x_{k_{\delta_j}}^{\delta_j}$ tem ela própria uma subsequência convergindo para x^\dagger . Observe que, cada subsequência de δ_j é ela própria uma sequência positiva convergindo para zero. Por isso, é suficiente para provar que $x_{k_{\delta_j}}^{\delta_j}$ tem uma subsequência convergindo para x^\dagger . Com isso temos (81).

Considere a sequência sem ruídos x_k e a subsequência δ_{j_m} construídas no Corolário 4.2.8. Fixe $\varepsilon > 0$. Uma vez que x_k é uma sequência sem ruídos, temos que $D_{\xi_k}(x^\dagger, x_k) \rightarrow 0$ e $\|x_k - x^\dagger\| \rightarrow 0$, quando $k \rightarrow \infty$. Então, existe um número natural $N \geq \sqrt{\frac{3L}{\varepsilon}}$, tal que,

$$D_{\xi_N}(x^\dagger, x_N) < \frac{\varepsilon}{3} \quad \text{e} \quad \|x_N - x^\dagger\| < \sqrt{\frac{\varepsilon}{3L}}. \quad (82)$$

A partir do Corolário 4.2.8, temos

$$\|x_N^{\delta_{j_m}} - x_N\| \leq \frac{1}{N}, \quad \text{para todo } m \geq N.$$

Agora, como $k_{\delta_{j_m}} \rightarrow \infty$, quando $m \rightarrow \infty$, existe um número $M_1 \in \mathbb{N}$ tal que $k_{\delta_{j_m}} \geq N$ para todo $m \geq M_1$. Segue da Proposição 4.2.10 que existe $M_2 \in \mathbb{N}$ de modo que a monotonia (79) é válida para todo $m \geq M_2$. Isso junto com a identidade dos três pontos (15) implica que para $m \geq \max\{N, M_1, M_2\}$ temos

$$\begin{aligned} D_{\xi_{k_{\delta_{j_m}}}}^{\delta_{j_m}}(x^\dagger, x_{k_{\delta_{j_m}}}^{\delta_{j_m}}) &\leq D_{\xi_N}^{\delta_{j_m}}(x^\dagger, x_N^{\delta_{j_m}}) \\ &\leq D_{\xi_N}(x^\dagger, x_N) - D_{\xi_N}(x_N^{\delta_{j_m}}, x_N) + \|\xi_N^{\delta_{j_m}} - \xi_N\| \|x_N^{\delta_{j_m}} - x^\dagger\| \\ &\leq D_{\xi_N}(x^\dagger, x_N) + L \|x_N^{\delta_{j_m}} - x_N\| \left(\|x_N^{\delta_{j_m}} - x_N\| + \|x_N - x^\dagger\| \right) < \varepsilon. \end{aligned}$$

Isso implica que $D_{\xi_{k_{\delta_j}}}^{\delta_j}(x^\dagger, x_{k_{\delta_j}}^{\delta_j}) \rightarrow 0$, quando $j \rightarrow \infty$. Consequentemente, pela Observação 2.1.7 temos que $\|x_{k_{\delta_j}}^{\delta_j} - x^\dagger\| \rightarrow 0$, quando $j \rightarrow \infty$. Portanto (81) acontece. \square

4.3 OPÇÕES DE MÉTODOS PARA A ITERAÇÃO INTERNA

Nesta seção vamos apresentar alguns métodos iterativos de otimização, métodos esses que podem ser usados para aplicar as iterações internas do Algoritmo 2

ou dos métodos de um passo apresentados no Capítulo 3. Cada um desses métodos, quando empregados para aproximar o minimizador do funcional de Tikhonov (22) (resp. (50)), gera uma sequência satisfazendo a Hipótese 3.0.1 (resp. a Hipótese 4.1.1), e portanto, pode ser usado na implementação dos algoritmos apresentados no Capítulo 3 (resp. na iteração interna do Algoritmo 2). A ideia inicial para se fazer as iterações internas foi usar o HPE, introduzido em [30] e também discutido em [24]. Como mostrado no Teorema 2.2.9 a sequência gerada pelo HPE satisfaz a Hipótese 3.0.1 e a Hipótese 4.1.1 com

$$0 < \Lambda_L = \left(1 - \frac{1 - \sigma}{(1 + L\bar{c})^2(1 + \sigma)}\right)^{1/2} < 1,$$

onde $\bar{c} > 0$ e $\sigma \in (0, 1)$ são apresentados na Seção 2.2. O próximo resultado deixa claro que os operadores a serem usados no nosso método são L -Lipschitz contínuo, se encaixando assim na teoria feita por [30].

Teorema 4.3.1. *Seja $T_\alpha : X \rightarrow \bar{\mathbb{R}}$ o funcional de Tikhonov definido em (22), onde $P : X \rightarrow \bar{\mathbb{R}}$ é uma função fortemente convexa com constante $\beta > 0$. Então $(\partial T_\alpha)^{-1}$ é L -Lipschitz contínuo em torno da origem, com constante $L = \frac{1}{\beta\alpha}$.*

Demonstração. Considere $w \in (\partial T_\alpha)^{-1}(x)$ e $s \in (\partial T_\alpha)^{-1}(y)$. Então,

$$x \in \partial T_\alpha(w) = A^*(Aw - b^\delta) + \alpha\partial P(w) \quad \text{e} \quad y \in \partial T_\alpha(s) = A^*(As - b^\delta) + \alpha\partial P(s).$$

Assim, existem vetores $u \in \partial P(w)$ e $v \in \partial P(s)$ tais que,

$$x = A^*(Aw - b^\delta) + \alpha u \quad \text{e} \quad y = A^*(As - b^\delta) + \alpha v.$$

Com isso obtemos

$$\begin{aligned} \langle x - y, w - s \rangle &= \langle A^*A(w - s) + \alpha(u - v), w - s \rangle = \|A(w - s)\|^2 + \alpha\langle u - v, w - s \rangle \\ &\geq \alpha\langle u - v, w - s \rangle \geq \alpha\beta\|w - s\|^2, \end{aligned}$$

onde a última desigualdade acima segue da Observação 2.1.7. Agora, usando a desigualdade de Cauchy-Schwarz temos

$$\|w - s\| \leq \frac{1}{\beta\alpha}\|x - y\|.$$

Finalmente, uma vez que $w \in (\partial T_\alpha)^{-1}(x)$ e $s \in (\partial T_\alpha)^{-1}(y)$ são arbitrários, temos

$$\|(\partial T_\alpha)^{-1}(x) - (\partial T_\alpha)^{-1}(y)\| \leq L\|x - y\|,$$

onde $L = \frac{1}{\beta\alpha}$. □

Observação 4.3.2. *Com uma prova similar, podemos mostrar que $(\partial T_{\alpha_k^\delta}^\delta)^{-1}$ (onde $T_{\alpha_k^\delta}^\delta$ definido em (50)) é L -Lipschitz contínuo em torno da origem, com $L = \frac{1}{\beta\alpha_k^\delta}$.*

Pelo Teorema 4.3.1 (resp. Observação 4.3.2), concluímos que o HPE, quando aplicado para encontrar o zero do subdiferencial do funcional de Tikhonov T_α , definido em (22) (resp. $T_{\alpha_k^\delta}$, definido em (50)), gera uma sequência satisfazendo a Hipótese 3.0.1 (resp. a Hipótese 4.1.1).

Na literatura existe uma classe ampla de métodos iterativos que, ou se encaixam na estrutura HPE ou a sequência gerada satisfaz a hipótese necessária para o passo interno do nosso algoritmo. Em particular o método forward-backward splitting modificado e o método tipo-Newton Regularizado ambos podem ser encontrados em [30]. Outros métodos iterativos que satisfazem nossos propósitos são: O Método do Ponto Proximal, o Método Forward-Backward de Tseng, o Método Extragradiante de Korpelevich, o Método de Múltiplas Direções Alternadas (ADMM) e o Método Douglas-Rachford. [30, 24, 26, 31, 25, 11, 26, 8, 18]. O resultado também é válido para diferentes tipos de métodos de gradiente, o Método das Inversas Parciais de Spingarn [3, 4] e algumas variações como os métodos inexatos e métodos splitting, veja [25, 24, 4]. Sob certas restrições no funcional de penalização P , também podemos usar o Método de Newton extragradiante [24], entre outros.

5 MÉTODO DE TIKHONOV ITERADO KACZMARZ

Vamos apresentar neste capítulo um método tipo Kaczmarz iterado, inspirado em [54], onde cada passo é definido como no método de Tikhonov iterado [37, Sec. 1.2]. Aqui o multiplicador de Lagrange é escolhido de forma que a próxima iteração tenha resíduo não apenas igual a uma certa quantidade como em [7, 42], mas sim que esse resíduo esteja em um certo intervalo, assim como feito em [6] e no capítulo anterior, essa é uma das principais funcionalidades computacionais desta estratégia.

Depois de escrever as equações (11) de forma única, i.e, $Ax = b$, onde $A = (A_0, \dots, A_{N-1}) : X \rightarrow Y^N$ e $b^\delta = (b_0^\delta, \dots, b_{N-1}^\delta)$, o método clássico de Tikhonov Iterado (TI) para resolver o problema mal posto (11) sujeito a (10), é definido pela iteração

$$x_{k+1}^\delta = \arg \min_{x \in X} \{ \lambda_k \|Ax - b^\delta\|^2 + \|x - x_k^\delta\|^2 \} \quad (83)$$

ou, de forma equivalente, por

$$x_{k+1}^\delta = x_k^\delta - \lambda_k (I + \lambda_k A^* A)^{-1} A^* (Ax_k^\delta - b^\delta) = (\lambda_k^{-1} I + A^* A)^{-1} [\lambda_k^{-1} x_k^\delta + A^* b^\delta], \quad (84)$$

onde $A^* : Y^N \rightarrow X$ é o operador adjunto de A . O parâmetro $\lambda_k > 0$ representa o multiplicador de Lagrange para o problema de projetar x_k^δ sobre os conjunto de níveis de $\|Ax - b^\delta\|^2$. Se na iteração (84) a sequência $\{\lambda_k = \lambda\}$ é constante, o método é dito Tikhonov iterado estacionário [56, 46, 55], caso contrário é chamado de Tikhonov iterado não estacionário [44, 12, 37].

5.1 APRESENTAÇÃO DO MÉTODO

Nesta seção vamos apresentar o método Tikhonov Iterado Kaczmarz relaxado (TIKr). Este método que visa resolver o problema mal posto (10), (11) é definido em cada passo como o método TI em (84) e a escolha do multiplicador de Lagrange é adotada similar a [6]. A iteração do método é definida por

$$x_{k+1}^\delta = x_k^\delta + h_k, \quad (85)$$

onde

$$h_k = \begin{cases} \lambda_k (I + \lambda_k A_{[k]}^* A_{[k]})^{-1} A_{[k]}^* (b_{[k]}^\delta - A_{[k]} x_k^\delta), & \text{se } \|A_{[k]} x_k^\delta - b_{[k]}^\delta\| > \tau \delta_{[k]} \\ 0 & , \text{ caso contrário} \end{cases} \quad (86)$$

e

$$\lambda_k = \begin{cases} \text{escolhido como no Algoritmo 3,} & \text{se } \|A_{[k]} x_k^\delta - b_{[k]}^\delta\| > \tau \delta_{[k]} \\ 0 & , \text{ caso contrário} \end{cases} \quad (87)$$

Aqui $[k] = (k \bmod N) \in \{0, 1, \dots, N-1\}$, $x_0^\delta = x_0 \in X$ é o ponto inicial e $\tau > 1$ é uma constante fixada.

O termo $h_k \in X$ em (86) é inspirado no passo iterativo proposto por [6] para a $[k]$ -ésima equação $A_{[k]}x = b_{[k]}$ de (11), com $b_{[k]}^\delta$ dado como em (10). Note que, se $\|A_{[k]}x_k^\delta - b_{[k]}^\delta\| \leq \tau\delta_{[k]}$ então definimos $\lambda_k = 0$, $h_k = 0$ e $x_{k+1}^\delta = x_k^\delta$.

Em algoritmos tipo Kaczmarz, um grupo de N passos subsequentes (começando em algum múltiplo de N) é chamado de ciclo, dito de outra forma, é o grupo das N equações $A_jx = b_j$, em cada ciclo a equação $A_jx = b_j$ pode ser usada ou não. As iterações (85), (86) e (87) terminam quando pela primeira vez todos os x_k^δ são iguais dentro de um ciclo. Dessa forma, paramos a iteração em $k_* = k_*(\{\delta_j\}_j, \{b_j^\delta\}_j)$, onde

$$k_* := \min \{ lN : l \in \mathbb{N} \text{ e } x_{lN}^\delta = x_{lN+1}^\delta = \dots = x_{lN+N}^\delta \}. \quad (88)$$

Em outras palavras, $k_* \in \mathbb{N}$ é o menor múltiplo de N tal que $x_{k_*}^\delta = x_{k_*+1}^\delta = \dots = x_{k_*+N}^\delta$ ou, equivalentemente, $\lambda_{k_*} = \lambda_{k_*+1} = \dots = \lambda_{k_*+N} = 0$.

Para o restante deste capítulo, vamos chamar $i = [k]$ e supor que as seguintes afirmações sejam verdadeiras:

(A1) Existe $x^* \in X$ tal que $A_jx^* = b_j$, onde $b_j \in R(A_j)$, $i = 0, \dots, N-1$, corresponde ao dado exato.

(A2) Os operadores $A_j : X \rightarrow Y$ são lineares, limitados e mal postos, i.e., caso o operador $A_j^{-1} : R(A_j) \rightarrow X$ exista, é descontínuo.

A partir de (A2) segue a existência de $C > 0$ com $\max_j \|A_j\| \leq C$.

Conforme já discutido na introdução, a etapa iterativa do método TIKr é análogo a que propõe [6], mais precisamente, dado $k \in \mathbb{N}$, vamos definir para cada $\mu > 0$ os conjuntos de níveis $\Omega_\mu^i := \{x \in X; \|A_ix - b_i^\delta\| \leq \mu\}$ da função residual $x \mapsto \|A_ix - b_i^\delta\|$ da i -ésima equação do sistema (11). Se a iteração x_k^δ não pertence a $\Omega_{\delta_i}^i$, a próxima iteração x_{k+1}^δ é computada resolvendo o problema de projeção relaxado

$$\begin{cases} \min_x \|x - x_k^\delta\|^2 \\ \text{s.t. } \|A_ix - b_i^\delta\|^2 \leq \mu^2, \quad \bar{\Phi}(\|A_ix_k^\delta - b_i^\delta\|, \delta_i) \leq \mu \leq \bar{\bar{\Phi}}(\|A_ix_k^\delta - b_i^\delta\|, \delta_i) \end{cases} \quad (89)$$

para $(x, \mu) \in X \times \mathbb{R}$, onde $\bar{\Phi}(u, v) = \eta u + (1 - \eta)v$ e $\bar{\bar{\Phi}}(u, v) = \bar{\eta}u + (1 - \bar{\eta})v$, $\forall u, v \in \mathbb{R}$, com $0 < \eta < \bar{\eta} < 1$. O intervalo $[\bar{\Phi}(\|A_ix_k^\delta - b_i^\delta\|, \delta_i), \bar{\bar{\Phi}}(\|A_ix_k^\delta - b_i^\delta\|, \delta_i)]$ é não degenerado sempre que $\|A_ix_k^\delta - b_i^\delta\| > \delta_i$. Se (x', μ') é uma solução de (89), temos $x_{k+1}^\delta = x'$ e $\|A_ix_{k+1}^\delta - b_i^\delta\| = \mu'$ (veja o Lema 5.1.1). Assim como em [6], x_{k+1}^δ é gerado projetando x_k^δ em um dos conjuntos convexos $(\Omega_\mu^i)_{\bar{\Phi} \leq \mu \leq \bar{\bar{\Phi}}}$.

Uma vez que não temos unicidade da solução de (89), existe mais de uma maneira de escolher x_{k+1}^δ . O próximo lema, que contém uma prova simples, aborda essa questão.

Lema 5.1.1. Suponha que $\|A_ix_k^\delta - b_i^\delta\| > \delta_i$. As seguintes afirmações são equivalentes:

1. $x' = P_{\Omega_\mu}(x_k^\delta)$ e $\bar{\Phi}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i) \leq \mu' \leq \bar{\bar{\Phi}}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i)$;
2. $(x', \mu') \in X \times \mathbb{R}$ é solução do problema de projeção relaxado (89);
3. $x' = x_k^\delta - \lambda(I + \lambda A_i^* A_i)^{-1} A_i^*(A_i x_k^\delta - b_i^\delta)$, para algum $\lambda > 0$,

$$\bar{\Phi}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i) \leq \|A_i x' - b_i^\delta\| \leq \bar{\bar{\Phi}}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i),$$

$$\text{e } \mu' = \|A_i x' - b_i^\delta\|;$$

(aqui $P_{\Omega}(x)$ representa a projeção ortogonal de x sobre o conjunto convexo Ω).

Demonstração. Veja [6, Lema 2.3]. □

Segue direto do Lema 5.1.1 que resolvendo o problema de projeção relaxado em (89) temos $\bar{\Phi}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i) \leq \|A_i x' - b_i^\delta\| \leq \bar{\bar{\Phi}}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i)$, onde $x' = x_k^\delta - \lambda(I + \lambda A_i^* A_i)^{-1} A_i^*(A_i x_k^\delta - b_i^\delta)$ e $\mu' = \|A_i x' - b_i^\delta\|$.

A seguir vamos nos inspirar no Lema 5.1.1 para propor o método TIKr.

- [1] Escolha $x_0 \in X$ e $\lambda_{\max} > 0$; defina $k := 0$;
- [2] Escolha $\tau > 1$, e $0 < \eta < \bar{\eta} < 1$;
- [3] Faça
 - [3.1] $i = [k]$;
 - [3.2] Se $[\|A_i x_k^\delta - b_i^\delta\| > \tau \delta_i]$ então
 - compute $(\lambda_k, h_k) \in \mathbb{R}_+ \times X$ tal que

$$\begin{cases} h_k = -\lambda_k (I + \lambda_k A_i^* A_i)^{-1} A_i^*(A_i x_k^\delta - b_i^\delta) \\ \eta \|A_i x_k^\delta - b_i^\delta\| + (1 - \eta) \delta_i \leq \|A_i(x_k^\delta + h_k) - b_i^\delta\| \leq \bar{\eta} \|A_i x_k^\delta - b_i^\delta\| + (1 - \bar{\eta}) \delta_i \end{cases}$$
 - Se $[\lambda_k > \lambda_{\max}]$ então

$$\lambda_k = \lambda_{\max}; h_k = -\lambda_{\max} (I + \lambda_{\max} A_i^* A_i)^{-1} A_i^*(A_i x_k^\delta - b_i^\delta)$$
 - caso contrário

$$\lambda_k = 0; h_k = 0;$$
 - [3.3] $x_{k+1}^\delta = x_k^\delta + h_k$;
 - [3.4] $k = k + 1$;
 - até $[(k) = 0]$ e $(\lambda_{k-1} = \lambda_{k-2} = \dots = \lambda_{k-N} = 0)$];
- [4] $k_* = k - N$;

Algoritmo 3 – Método Tikhonov Iterado Kaczmarz relaxado (TIKr).

Observação 5.1.2 (Sobre a limitação dos multiplicadores de Lagrange no Algoritmo 3).

- No Algoritmo 3 os multiplicadores de Lagrange λ_k são acotados superiormente por $\lambda_{\max} > 0$. Isso é necessário para provar os resultados sobre convergência para dados exatos (veja Teoremas 5.2.2 e 5.2.5).

- É importante notar que no caso de dados com ruídos ($\delta > 0$, fixo), essa suposição não tem influência, nesse caso λ_{\max} pode ser escolhido arbitrariamente grande, pois o Algoritmo 3 sempre para após um número finito de passos (veja Corolário 5.1.8). Portanto o Algoritmo 3 pode ser implementado sem limitar os multiplicadores.
- Nos casos em que $\lambda_k < \lambda_{\max}$ e $\lambda_k = \lambda_{\max}$ a desigualdade $\eta \|A_i x_k^\delta - b_i^\delta\| + (1 - \eta)\delta_i \leq \|A_i(x_k^\delta + h_k) - b_i^\delta\|$ é sempre válida (adaptando [61, Lema 2.3 itens 2 e 4]), i.e, o resíduo é decrescente. Porém não podemos garantir que $\|A_i(x_k^\delta + h_k) - b_i^\delta\| \leq \bar{\eta} \|A_i x_k^\delta - b_i^\delta\| + (1 - \bar{\eta})\delta_i$ quando $\lambda_k = \lambda_{\max}$.
- No caso particular $N = 1$, o sistema (11) reduz a uma única equação e os resultados similares podem ser provados sem a limitação dos multiplicadores. Neste caso os resultados são um caso particular de [58] para equações de operadores não lineares.

Vamos agora dar início a uma sequência de resultados que serão importantes para nossa análise de convergência. O primeiro resultado é sobre a etapa iterativa do Algoritmo 3.

Para simplificar a notação vamos usar $d_k^\delta := b_i^\delta - A_i x_{k+1}^\delta = b_i^\delta - A_i h_k - A_i x_k^\delta$.

Lema 5.1.3. Assuma que (A1) e (A2) sejam verdade e sejam x_k^δ , h_k , λ_k definidos em (85), (86) e (87) respectivamente. No caso de dados com ruídos, as afirmações

1. $A_i x_{k+1}^\delta - b_i^\delta = (I + \lambda_k A_i A_i^*)^{-1} (A_i x_k^\delta - b_i^\delta)$;
2. $h_k = \lambda_k A_i^* (b_i^\delta - A_i x_{k+1}^\delta)$;
3. $\eta \|A_i x_k^\delta - b_i^\delta\| \leq \|A_i x_{k+1}^\delta - b_i^\delta\| \leq \|A_i x_k^\delta - b_i^\delta\|$;
valem para $0 \leq k < k_*$.

No caso de dados exatos 1. e 2. continuam válidos. Além disso,

4. $\eta \|A_i x_k - b_i\| \leq \|A_i x_{k+1} - b_i\| \leq \bar{\eta} \|A_i x_k - b_i\|$, sempre que $\lambda_k < \lambda_{\max}$;
5. $\eta \|A_i x_k - b_i\| \leq \|A_i x_{k+1} - b_i\| \leq \|A_i x_k - b_i\|$, sempre que $\lambda_k = \lambda_{\max}$.

Demonstração. A condição de otimalidade aplicado em (83) (com A_i) nos fornece $\lambda_k A_i^* (A_i x_{k+1}^\delta - b_i^\delta) + x_{k+1}^\delta - x_k^\delta = 0$, isso implica $\lambda_k A_i^* (A_i x_{k+1}^\delta - b_i^\delta) + A_i (x_{k+1}^\delta - x_k^\delta) = 0$. Somando $A_i x_{k+1}^\delta - b_i^\delta$ em ambos os lados dessa última igualdade temos $(I + \lambda_k A_i A_i^*) (A_i x_{k+1}^\delta - b_i^\delta) = A_i x_k^\delta - b_i^\delta$, o que prova 1., a afirmação 2. segue diretamente dessa condição de otimalidade. Para provar 3. note que no caso $\lambda_k < \lambda_{\max}$ a afirmação segue do passo [3.2] do Algoritmo 3. Se $\lambda_k = \lambda_{\max}$ a primeira desigualdade segue da Observação 5.1.2 e a segunda desigualdade segue de (83). O item 4. segue do passo [3.2] do Algoritmo 3 com $\delta_i = 0$. Quanto ao item 5. a primeira desigualdade segue da Observação 5.1.2 e a segunda desigualdade segue de (83).

□

Proposição 5.1.4. Assuma que (A1) e (A2) valem e sejam x_k^δ , h_k , λ_k definidos por (85), (86) e (87) respectivamente. Então

$$\|x_{k+1}^\delta - x^*\|^2 - \|x_k^\delta - x^*\|^2 \leq 2\eta \lambda_k \|d_k^\delta\|(\delta_i - \|A_i x_k^\delta - b_i^\delta\|) - \|x_{k+1}^\delta - x_k^\delta\|^2, \quad (90)$$

para $k = 0, \dots, k_* - 1$. Em particular no caso de dados exatos ($b_i^\delta = b_i$) temos

$$\|x_{k+1} - x^*\|^2 - \|x_k - x^*\|^2 \leq -2\eta^2 \lambda_k \|A_i x_k - b_i\|^2 - \|x_{k+1} - x_k\|^2, \quad (91)$$

para $k = 0, \dots$

Demonstração. Primeiro vamos analisar o caso $\delta_i > 0$. Se $\|A_i x_k^\delta - b_i^\delta\| \leq \tau \delta_i$, então $\lambda_k = 0$ e $x_{k+1}^\delta = x_k^\delta$. Assim, (90) é trivial. Caso contrário, duas situações podem ocorrer. Podemos ter $\lambda_k < \lambda_{\max}$, nesse caso segue a partir do Lema 5.1.3 item 2. que

$$\begin{aligned} & \|x_{k+1}^\delta - x^*\|^2 - \|x_k^\delta - x^*\|^2 \\ &= 2 \langle x_{k+1}^\delta - x_k^\delta, x_{k+1}^\delta - x^* \rangle - \|x_{k+1}^\delta - x_k^\delta\|^2 \\ &= 2\lambda_k \langle b_i^\delta - A_i x_{k+1}^\delta, A_i(x_{k+1}^\delta - x^*) \rangle - \|x_{k+1}^\delta - x_k^\delta\|^2 \\ &= 2\lambda_k \langle b_i^\delta - A_i x_{k+1}^\delta, A_i x_{k+1}^\delta - b_i^\delta + b_i^\delta - A_i x^* \rangle - \|x_{k+1}^\delta - x_k^\delta\|^2 \\ &\leq 2\lambda_k [-\|d_k^\delta\|^2 + \|d_k^\delta\| \delta_i] - \|x_{k+1}^\delta - x_k^\delta\|^2. \end{aligned} \quad (92)$$

Segue do Passo [3.2] do Algoritmo 3 que $\bar{\Phi}(\|A_i x_k^\delta - b_i^\delta\|, \delta_i) = \eta \|A_i x_k^\delta - b_i^\delta\| + (1 - \eta)\delta_i \leq \|d_k^\delta\|$. Consequentemente $-\|d_k^\delta\|^2 + \|d_k^\delta\| \delta_i = \|d_k^\delta\|(\delta_i - \|d_k^\delta\|) \leq \eta \|d_k^\delta\|(\delta_i - \|A_i x_k^\delta - b_i^\delta\|)$, assim (90) é verificado. O segundo caso é quando $\lambda_k = \lambda_{\max}$, nesse caso segue da Observação 5.1.2 que também vale $\eta \|A_i x_k^\delta - b_i^\delta\| + (1 - \eta)\delta_i \leq \|d_k^\delta\|$, implicando que vale (90).

No caso de dados exatos, segue a partir do Lema 5.1.3 itens 4. e 5. (quando $\lambda_k < \lambda_{\max}$ e $\lambda_k = \lambda_{\max}$) que $-\|d_k\|^2 \leq -\eta^2 \|A_i x_k - b_i\|^2$, assim (92) (com $\delta_i = 0$) implica em (91). \square

O próximo corolário nos diz que a cada passo do Algoritmo 3, a sequência gerada se aproxima de uma solução do problema (11).

Corolário 5.1.5. Assuma que (A1) e (A2) valem e sejam x_k^δ , h_k , λ_k definidos em (85), (86) e (87) respectivamente. Então

$$\|x_{k+1}^\delta - x^*\|^2 \leq \|x_k^\delta - x^*\|^2, \quad k = 0, \dots, k_* - 1. \quad (93)$$

Adicionalmente, no caso de dados exatos temos $\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2$, para $k = 0, 1, \dots$

Demonstração. Os dois casos seguem diretamente de (90) e (91) respectivamente. \square

Corolário 5.1.6. Assuma que (A1) e (A2) valem e sejam x_k^δ , h_k , λ_k definidos em (85), (86) e (87) respectivamente. Então

$$\lambda_k \geq \min \left\{ \frac{(\|A_i x_k^\delta - b_i^\delta\| - \|d_k^\delta\|) \|A_i x_k^\delta - b_i^\delta\|}{\|A_i^*(A_i x_k^\delta - b_i^\delta)\|^2}, \lambda_{\max} \right\}, \quad k = 0, \dots, k_* - 1. \quad (94)$$

Além disso, se $\|A_i x_k^\delta - b_i^\delta\| > \tau \delta_i$ (para algum $0 \leq k \leq k_* - 1$) então

$$\lambda_k > \min \left\{ C^{-2}(1 - \bar{\eta})(1 - 1/\tau), \lambda_{\max} \right\} =: \theta, \text{ com } C > 0 \text{ definido anteriormente.}$$

Adicionalmente, no caso de dados exatos temos $\lambda_k \geq \min\{\lambda_{\max}, C^{-2}(1 - \bar{\eta})\}$ para $k \geq 0$.

Demonstração. Seja $0 \leq k < k_*$. Se $\lambda_k = \lambda_{\max}$, então não há o que provar. Caso contrário, se $\lambda_k < \lambda_{\max}$ duas situações podem ocorrer. Primeiro, se $\|A_i x_k^\delta - b_i^\delta\| \leq \tau \delta_i$ então $\lambda_k = 0$ e $x_{k+1}^\delta = x_k^\delta$. Assim, (94) é trivial. A outra situação é $\|A_i x_k^\delta - b_i^\delta\| > \tau \delta_i$, neste caso a prova de (94) é conforme em [6, Corolário 2.5].

Para provar a segunda afirmação, ou $\lambda_k = \lambda_{\max}$ (nada há fazer) ou $\lambda_k < \lambda_{\max}$ e vale que $\|d_k^\delta\| \leq \bar{\eta} \|A_i x_k^\delta - b_i^\delta\| + (1 - \bar{\eta})\delta$. Consequentemente, $\|A_i x_k^\delta - b_i^\delta\| - \|d_k^\delta\| \geq (1 - \bar{\eta})(\|A_i x_k^\delta - b_i^\delta\| - \delta_i)$. Assim, segue a partir de (94) que

$$\lambda_k \geq C^{-2} \frac{\|A_i x_k^\delta - b_i^\delta\| - \|d_k^\delta\|}{\|A_i x_k^\delta - b_i^\delta\|} \geq C^{-2}(1 - \bar{\eta}) \left(1 - \frac{\delta_i}{\|A_i x_k^\delta - b_i^\delta\|}\right) > C^{-2}(1 - \bar{\eta})(1 - 1/\tau).$$

No caso de dados exatos, dado $k \in \mathbb{N}$, temos $\lambda_k < \lambda_{\max}$ e $\|d_k\| \leq \bar{\eta} \|A_i x_k - b_i\|$ (veja Lema 5.1.3 2.) ou $\lambda_k = \lambda_{\max}$ (nesse caso não há o que provar). No primeiro caso, segue de (94) que

$$\lambda_k \geq C^{-2} \left(1 - \frac{\|d_k\|}{\|A_i x_k - b_i\|}\right) \geq C^{-2}(1 - \bar{\eta}).$$

Consequentemente, $\lambda_k \geq \min\{\lambda_{\max}, C^{-2}(1 - \bar{\eta})\}$. □

Corolário 5.1.7. Assuma que (A1) e (A2) valem e sejam x_k , h_k , λ_k definidos em (85), (86) e (87) no caso de dados exatos (i.e. $b_i^\delta = b_i$, $i = 0, \dots, N - 1$). Então as séries

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2, \quad \sum_{k=0}^{\infty} \lambda_k \|A_{[k]} x_k - b_{[k]}\|^2, \quad \sum_{k=0}^{\infty} \lambda_k \|d_k\|^2 \text{ e } \sum_{k=0}^{\infty} \|A_{[k]} x_k - b_{[k]}\|^2$$

convergem.

Demonstração. A convergência da primeira e da segunda série segue de (91), usando o argumento de soma telescópica. A terceira soma segue da convergência da segunda série e do Lema 5.1.3 itens 4. e 5. . A última segue da segunda soma e do Corolário 5.1.6. □

Corolário 5.1.8. *Assuma que (A1) e (A2) valem e sejam x_k^δ , h_k , λ_k definidos em (85), (86) e (87) respectivamente. Então o índice de parada k_* definido em (88) é finito e*

$$k_* \leq N \left(1 + \|x_0 - x^*\|^2 \left[2\eta^2 \theta \tau \delta_{\min}^2 (\tau - 1) \right]^{-1} \right), \quad (95)$$

onde $\delta_{\min} = \min\{\delta_i\}$, com $i \in \{0, \dots, N-1\}$.

Demonstração. Assuma por contradição que k_* não seja finito, i.e., em cada ciclo $\{IN, \dots, IN+N-1\}$, $l \in \mathbb{N}$, do Algoritmo 3, existe pelo menos um índice $j(l) \in \{0, \dots, N-1\}$ tal que $\|A_{j(l)} x_{IN+j(l)} - b_{j(l)}^\delta\| \geq \tau \delta_{j(l)}$. De (90) e do fato de $\|A_{[k]} x_k^\delta - b_{[k]}^\delta\| \geq \tau \delta_{[k]}$ ou $\lambda_k = 0$, obtemos para todo $l \in \mathbb{N}$

$$\begin{aligned} \|x_0 - x^*\|^2 &\geq 2\eta \sum_{k=0}^{IN} \lambda_k \|d_k^\delta\| (\|A_i x_k^\delta - b_i^\delta\| - \delta_i) \\ &\geq 2\eta^2 \sum_{k=0}^{IN} \lambda_k \|A_i x_k^\delta - b_i^\delta\| (\|A_i x_k^\delta - b_i^\delta\| - \delta_i) \\ &\geq 2\eta^2 \sum_{s=0}^l \lambda_{sN+j(s)} \|A_{j(s)} x_{sN+j(s)}^\delta - b_{j(s)}^\delta\| \left(\|A_{j(s)} x_{sN+j(s)}^\delta - b_{j(s)}^\delta\| - \delta_{j(s)} \right) \\ &\geq 2\eta^2 \sum_{s=0}^l \lambda_{sN+j(s)} \tau \delta_{j(s)}^2 (\tau - 1) \geq l 2\eta^2 \theta \tau \delta_{\min}^2 (\tau - 1), \end{aligned} \quad (96)$$

(a última desigualdade segue do Corolário 5.1.6). Uma vez que o lado direito de (96) se torna ilimitado quando $l \rightarrow \infty$ isso gera uma contradição, e portanto k_* é finito. A estimativa (95) segue substituindo $l = (k_* - N)/N$ em (96). \square

5.2 ANÁLISE DE CONVERGÊNCIA

Nesta seção vamos apresentar os principais resultado de convergência do método TIKr, começando com o teorema para dados sem ruídos, $\delta_i = 0$, $i = 0, \dots, N-1$. Note que no caso de dados exatos, $\lambda_k = 0$ e $h_k = x_{k+1} - x_k = 0$ se, e somente se, $\|A_i x_k - b_i\| = 0$ para $i = 0, \dots, N-1$ (veja o Passo [3.2] do Algoritmo 3).

Definição 5.2.1. *Vamos chamar de solução x_0 -mínima do problema (11), o elemento x^\dagger tal que $\|x^\dagger - x_0\| = \inf\{\|x - x_0\|; Ax = b\}$. A existência de x^\dagger é assegurada pelo Lema 2.3.3. Além disso, x^\dagger é a única solução de (11) com essa propriedade.*

Teorema 5.2.2 (Convergência para dados exatos). *Assuma que (A1) e (A2) valem e sejam x_k , h_k , λ_k definidos em (85), (86) e (87) respectivamente em caso de dados exatos (i.e. $b_i^\delta = b_i$, $i = 0, \dots, N-1$). Então x_k termina após um número finito de passos com uma solução de (11) ou $x_k \rightarrow x^\dagger$, quando $k \rightarrow \infty$, onde x^\dagger é a solução x_0 -mínima de (11).*

Demonstração. Se após um número k de iterações, os resíduos de todas as equações são iguais a zero, então o algoritmo termina e x_k é uma solução de (11). Caso contrário, defina $e_k := x^* - x_k$. Segue do Corolário 5.1.5 que $\|e_k\|$ é monótona não crescente. Então, $\|e_k\|$ converge para algum $\epsilon \geq 0$. Vamos mostrar que e_k é uma sequência de Cauchy.

Para mostrar que e_k é uma sequência de Cauchy é suficiente provar que $|\langle e_n - e_k, e_n \rangle| \rightarrow 0$, $|\langle e_n - e_l, e_n \rangle| \rightarrow 0$ quando $k, l \rightarrow \infty$ com $k \leq l$ para algum $k \leq n \leq l$ [50, Teorema 2.3]. Sejam $k \leq l$ arbitrários e escreva $k = k_0N + k_1$, $l = l_0N + l_1$, com $k_1, l_1 \in \{0, \dots, N-1\}$. Seja $n_0 \in \{k_0, \dots, l_0\}$ um índice tal que para todo $j_0 \in \{k_0, \dots, l_0\}$

$$\sum_{s=0}^{N-1} \lambda_{n_0N+s} \|A_s x_{n_0N+s} - b_s\| \leq \sum_{s=0}^{N-1} \lambda_{j_0N+s} \|A_s x_{j_0N+s} - b_s\|, \quad (97)$$

e seja $n = n_0N + N - 1$. Usando a notação $j = j_0N + j_1$ temos

$$\begin{aligned} |\langle e_n - e_k, e_n \rangle| &= \left| \sum_{j=k}^{n-1} \langle (x_{j+1} - x_j), (x_n - x^*) \rangle \right| = \left| \sum_{j=k}^{n-1} \lambda_j \langle A_{[j]} x_{j+1} - b_{[j]}, A_{[j]} x_n - A_{[j]} x^* \rangle \right| \\ &\leq \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \lambda_j \|A_{j_1} x_{j+1} - b_{j_1}\| \|A_{j_1} x_n - b_{j_1}\| \\ &\leq \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \lambda_j \|d_{j_1}\| \|A_{j_1} x_n - b_{j_1}\|. \end{aligned} \quad (98)$$

O último termo do lado direito de (98) pode ser estimado por

$$\begin{aligned} \|A_{j_1} x_n - b_{j_1}\| &= \|A_{j_1} x_{n_0N+N-1} - b_{j_1}\| \\ &\leq \|A_{j_1} x_{n_0N+j_1+1} - b_{j_1}\| + \sum_{s=j_1+1}^{N-2} \|A_{j_1} x_{n_0N+s+1} - A_{j_1} x_{n_0N+s}\| \\ &\leq \|A_{j_1} x_{n_0N+j_1+1} - b_{j_1}\| + \sum_{s=j_1+1}^{N-2} C \|x_{n_0N+s+1} - x_{n_0N+s}\| \\ &\leq \|A_{j_1} x_{n_0N+j_1+1} - b_{j_1}\| + \sum_{s=j_1+1}^{N-2} C \lambda_{n_0N+s} \|A_s^* (b_s - A_s x_{n_0N+s})\| \\ &\leq \|A_{j_1} x_{n_0N+j_1+1} - b_{j_1}\| + \sum_{s=0}^{N-1} C^2 \lambda_{n_0N+s} \|A_s x_{n_0N+s} - b_s\| \\ &\leq \sum_{s=0}^{N-1} (1 + C^2 \lambda_{n_0N+s}) \|A_s x_{n_0N+s} - b_s\| \\ &\leq (\lambda_{\min}^{-1} + C^2) \sum_{s=0}^{N-1} \lambda_{n_0N+s} \|A_s x_{n_0N+s} - b_s\|. \end{aligned}$$

(com $\lambda_{\min} = \min\{\lambda_{\max}, C^{-2}(1 - \bar{\eta})\}$, veja Corolário 5.1.6). Dessa forma, segue de (97) que $\|A_{j_l} x_n - b_{j_l}\| \leq (\lambda_{\min}^{-1} + C^2) \sum_{s=0}^{N-1} \lambda_{j_0 N+s} \|A_s x_{j_0 N+s} - b_s\|$, para todo $j_0 \in \{k_0, \dots, l_0\}$.

Dessa última desigualdade junto com (98) obtemos

$$\begin{aligned} |\langle e_n - e_k, e_n \rangle| &\leq (\lambda_{\min}^{-1} + C^2) \sum_{j_0=k_0}^{n_0} \left[\sum_{j_1=0}^{N-1} \lambda_j \|d_j\| \right] \left[\sum_{s=0}^{N-1} \lambda_{j_0 N+s} \|A_s x_{j_0 N+s} - b_s\| \right] \\ &= (\lambda_{\min}^{-1} + C^2) \sum_{j_0=k_0}^{n_0} \left[\sum_{j_1=0}^{N-1} \lambda_j \|d_j\| \right] \left[\sum_{j_1=0}^{N-1} \lambda_j \|A_{j_1} x_j - b_{j_1}\| \right] \\ &\leq (\lambda_{\min}^{-1} + C^2) \frac{N\lambda_{\max}}{2} \sum_{l_0=k_0}^{n_0} \left[\sum_{l_1=0}^{N-1} \lambda_{l_1} \|d_{l_1}\|^2 + \sum_{l_1=0}^{N-1} \lambda_{l_1} \|A_{l_1} x_{l_1} - b_{l_1}\|^2 \right]. \end{aligned} \quad (99)$$

Segue do Corolário 5.1.7 que o lado direito de (99) tende a zero quando $k, l \rightarrow \infty$. Analogamente mostra-se que $|\langle e_n - e_l, e_n \rangle| \rightarrow 0$ quando $k, l \rightarrow \infty$.

Portando, e_k é uma seqüência de Cauchy e x_k converge para algum $x^+ \in X$. Uma vez que $\|A_{[k]} x_k - b_{[k]}\| \rightarrow 0$ quando $k \rightarrow \infty$ (veja Corolário 5.1.7), x^+ é uma solução de (11).

Segue do Lema 5.1.3 2. que $x_{k+1} - x_k \in \mathcal{R}(A_i^*) \subset \mathcal{N}(A_i)^\perp \subset \mathcal{N}(A)^\perp$. Consequentemente $x^+ \in x_0 + \mathcal{N}(A)^\perp$, pela Definição 5.2.1 x^+ é a única solução x_0 -mínima de (11), com isso o Teorema está provado. \square

Definição 5.2.3. Para $k < k_*$, um vetor $z \in X$ é chamado sucessor de x_k^δ se existe um par $(\lambda_k \geq 0, h_k \in X)$ definido no Passo [3.2] do Algoritmo 3, tal que $z = x_k^\delta + h_k$.

Definição 5.2.4. Uma seqüência sem ruídos é uma seqüência $x_k \subset X$ gerada pelo Algoritmo 3 com $\delta_i = 0, i = 0, \dots, N-1$. Note que

1. x_{k+1} é um sucessor de x_k para todo $k \in \mathbb{N}$;
2. $x_{k+1} = \arg \min_x T_{k, \lambda_k}(x)$, onde $T_{k, \lambda_k}(x) := \lambda_k \|A_i x - b_i\|^2 + \|x - x_k\|^2$ e $\lambda_k \geq 0$ é definido no Passo [3.2] do Algoritmo 3;
3. $h_k = x_{k+1} - x_k = 0$ se, e somente se $A_i x_k = b_i$ (veja o Passo [3.2] do Algoritmo 3); nesse caso, o único sucessor de x_k é $z = x_k$, ele mesmo.

Teorema 5.2.5 (Estabilidade). Assuma (A1) e (A2). Sejam $\delta^j := (\delta_0^j, \dots, \delta_{N-1}^j)$ uma seqüência de números reais positivos convergindo a zero e $b^{\delta^j} = (b_0^{\delta^j}, \dots, b_{N-1}^{\delta^j})$ a seqüência de dados com ruídos satisfazendo (10). Para cada $j \in \mathbb{N}$, seja $x_{k+1}^{\delta^j}$ um sucessor de $x_k^{\delta^j}$ para $0 \leq k \leq k_*(\delta^j, b^{\delta^j})$. Então, existe uma seqüência sem ruídos x_k tal que, para cada $k \in \mathbb{N}$ fixado, existe uma subsequência δ^{j_m} (dependendo de k) de δ^j , satisfazendo

$$x_l^{\delta^{j_m}} \rightarrow x_l, \text{ quando } m \rightarrow \infty, \text{ para } l = 0, \dots, k.$$

Demonstração. Vamos usar indução. Uma vez que $x_0^\delta = x_0$ para todo $\delta \geq 0$, a afirmação é válida para $k = 0$. O argumento consiste em escolher repetidamente uma subsequência da subsequência corrente. Por simplicidade, vamos chamar uma subsequência de δ^j também de δ^j . Suponha por indução que a afirmação valha para algum $k \in \mathbb{N}$, i.e., que existe uma subsequência δ^j e x_l satisfazendo

$$x_l^{\delta^j} \rightarrow x_l, \text{ quando } j \rightarrow \infty, \text{ para } l = 0, \dots, k, \quad (100)$$

com x_{l+1} um sucessor de x_l para $l = 0, \dots, k-1$.¹

Desde que $x_{k+1}^{\delta^j}$ é um sucessor de $x_k^{\delta^j}$ (para cada δ^j) existe um número positivo $\lambda_k^{\delta^j}$ tal que $x_{k+1}^{\delta^j} = x_k^{\delta^j} - \lambda_k^{\delta^j} (I + \lambda_k^{\delta^j} A_i^* A_i)^{-1} A_i^* (A_i x_k^{\delta^j} - b_i^{\delta^j})$. Depois disso vamos provar a existência de um sucessor x_{k+1} de x_k e de uma subsequência δ^j da atual subsequência tal que $x_{k+1}^{\delta^j} \rightarrow x_{k+1}$ quando $j \rightarrow \infty$, completando assim o argumento de indução. Vamos dividir essa prova em três passos:

Passo 1. Vamos definir o elemento $x_{k+1} \in X$.

Passo 2. Vamos provar que

$$x_{k+1}^{\delta^j} \rightarrow x_{k+1}, \text{ quando } j \rightarrow \infty. \quad (101)$$

Passo 3. Vamos mostrar que x_{k+1} é um sucessor de x_k .

Prova do Passo 1. Segue a partir do Algoritmo 3 que $\lambda_k^{\delta^j} \leq \lambda_{\max}$ para cada $k \in \mathbb{N}$ fixado (veja Observação 5.1.2). Consequentemente, para alguma subsequência δ^j existe um $\lambda_k > 0$ satisfazendo

$$\lambda_k := \lim_{j \rightarrow \infty} \lambda_k^{\delta^j} < \infty. \quad (102)$$

Agora definimos $x_{k+1} := \arg \min_{x \in X} T_{k, \lambda_k}(x)$, com $x_{k+1} = x_k + h_k$ e T_{k, λ_k} como na Definição 5.2.4 item 2.

Prova do Passo 2. Da definição de $x_{k+1}^{\delta^j}$ e de x_{k+1} segue que

$$\lambda_k^{\delta^j} A_i^* (A_i x_{k+1}^{\delta^j} - b_i^{\delta^j}) + x_{k+1}^{\delta^j} - x_k^{\delta^j} = 0 = \lambda_k A_i^* (A_i x_{k+1} - b_i) + x_{k+1} - x_k,$$

com isso obtemos

$$0 = \lambda_k^{\delta^j} A_i^* A_i (x_{k+1}^{\delta^j} - x_{k+1}) + x_{k+1}^{\delta^j} - x_{k+1} + (\lambda_k^{\delta^j} - \lambda_k) A_i^* A_i x_{k+1} + \lambda_k b_i - \lambda_k^{\delta^j} b_i^{\delta^j} - x_k^{\delta^j} + x_k.$$

¹ Note que $k_*(\delta^j, b^{\delta^j}) \geq k$ para j grande o suficiente.

Multiplicando essa última expressão por $x_{k+1}^{\delta_j} - x_{k+1}$ temos

$$\begin{aligned} \|x_{k+1}^{\delta_j} - x_{k+1}\|^2 &\leq \lambda_k^{\delta_j} \|A_i(x_{k+1}^{\delta_j} - x_{k+1})\|^2 + \|x_{k+1}^{\delta_j} - x_{k+1}\|^2 \\ &= |\lambda_k - \lambda_k^{\delta_j}| \langle A_i^* A_i x_{k+1}, (x_{k+1}^{\delta_j} - x_{k+1}) \rangle + \langle x_k - x_k^{\delta_j}, x_{k+1}^{\delta_j} - x_{k+1} \rangle \\ &\quad + \langle \lambda_k^{\delta_j} b_i^{\delta_j} - \lambda_k b_i, x_{k+1}^{\delta_j} - x_{k+1} \rangle \\ &\leq C^2 |\lambda_k - \lambda_k^{\delta_j}| \|x_{k+1}\| \|x_{k+1}^{\delta_j} - x_{k+1}\| + \|x_k - x_k^{\delta_j}\| \|x_{k+1}^{\delta_j} - x_{k+1}\| \\ &\quad + \|\lambda_k b_i - \lambda_k^{\delta_j} b_i^{\delta_j}\| \|x_{k+1}^{\delta_j} - x_{k+1}\|. \end{aligned}$$

Consequentemente $\|x_{k+1}^{\delta_j} - x_{k+1}\| \leq C^2 |\lambda_k - \lambda_k^{\delta_j}| \|x_{k+1}\| + \|x_k - x_k^{\delta_j}\| + \|\lambda_k b_i - \lambda_k^{\delta_j} b_i^{\delta_j}\|$. Portanto (101) segue de (10), (100) e (102).

Prova do Passo 3. Vamos dividir em dois casos: Primeiro caso, $\|A_i x_k - b_i\| = 0$: nesse caso devemos concluir que $x_{k+1} = x_k$. Se $h_k = 0$, temos que $x_{k+1} = x_k$, caso contrário segue da definição de x_{k+1} que $0 = \lambda_k A_i^* (A_i(x_{k+1} - x_k) + A_i x_k - b_i) + x_{k+1} - x_k = (\lambda_k A_i^* A_i + I)(x_{k+1} - x_k)$, isso implica $x_{k+1} = x_k$.

Segundo caso, $\|A_i x_k - b_i\| > 0$: Se $\lambda_k^{\delta_j} < \lambda_{\max}$ segue de (10) e (100) que

$$\lim_{j \rightarrow \infty} \|A_i x_k^{\delta_j} - b_i^{\delta_j}\| = \|A_i x_k - b_i\| > 0. \quad (103)$$

Do passo [3.2] do Algoritmo 3 temos que

$\|A_i x_{k+1}^{\delta_j} - b_i^{\delta_j}\| \in [\bar{\Phi}(\|A_i x_k^{\delta_j} - b_i^{\delta_j}\|, \delta_j^j), \bar{\Phi}(\|A_i x_k^{\delta_j} - b_i^{\delta_j}\|, \delta_j^j)]$. Fazendo $j \rightarrow \infty$ e considerando (100) e (102) segue que $\|A_i x_{k+1} - b_i\| \in [\bar{\Phi}(\|A_i x_k - b_i\|, 0), \bar{\Phi}(\|A_i x_k - b_i\|, 0)]$, portanto x_{k+1} é um sucessor de x_k . Agora se $\lambda_k^{\delta_j} = \lambda_{\max}$, segue de (103) que existe $J \in \mathbb{N}$ tal que $\|A_i x_k^{\delta_j} - b_i^{\delta_j}\| > \tau \delta_j^j$ para $j > J$; $\lambda_k^{\delta_j} < \lambda_{\max}$. Se $\#\mathcal{J} < \infty$, a menos de uma subsequência temos $\lambda_k^{\delta_j} = \lambda_{\max}$. Então $\lambda_k = \lim_{j \rightarrow \infty} \lambda_k^{\delta_j} = \lambda_{\max}$.

Além do mais $\bar{\Phi}(\|A_i x_k^{\delta_j} - b_i^{\delta_j}\|, \delta_j^j) \leq \|A_i x_{k+1}^{\delta_j} - b_i^{\delta_j}\|$, fazendo $j \rightarrow \infty$ concluímos que $\bar{\Phi}(\|A_i x_k - b_i\|, 0) \leq \|A_i x_{k+1} - b_i\|$, portanto x_{k+1} é um sucessor de x_k . □

No próximo resultado vamos tirar proveito dos Teoremas 5.2.2 e 5.2.5 para mostrar que $x_{k_*(\delta)}^{\delta}$ converge para uma solução de (11) quando $\delta \rightarrow 0$, onde $k_*(\delta) = k_*$ é definido em (88).

Teorema 5.2.6 (Regularização). Assuma (A1) e (A2). Seja $\delta^j := (\delta_0^j, \dots, \delta_{N-1}^j)_{j \in \mathbb{N}}$ uma seqüência de números reais positivos convergindo para zero e $b^{\delta^j} = (b_0^{\delta^j}, \dots, b_{N-1}^{\delta^j})_{j \in \mathbb{N}}$ a seqüência de dados com ruídos satisfazendo (10). Para cada $j \in \mathbb{N}$, seja $x_{k+1}^{\delta^j}$ um sucessor de $x_k^{\delta^j}$ para $0 \leq k \leq k_*^j = k_*(\delta^j, b^{\delta^j})$. Então, toda subsequência de $x_{k_*^j}^{\delta^j}$ tem ela própria uma subsequência convergindo para uma solução de (11). Além disso, se $k_*^j \rightarrow \infty$ quando $j \rightarrow \infty$, então $x_{k_*^j}^{\delta^j} \rightarrow x^\dagger$.

Demonstração. Fixe uma subsequência qualquer de δ^j e denote novamente por δ^j . Devemos considerar dois casos: Primeiro, assuma que a seqüência k_*^j é limitada. Então

esta tem um ponto de acumulação finito. Nesse caso, devemos mostrar que existe uma subsequência δ_j tal que $x_{k_*^j}^{\delta_j}$ converge para uma solução do problema inverso. Agora, uma vez que $k_*^j \in \mathbb{N}$, para todo j , podemos extrair um subsequência δ^{jm} de δ^j tal que $k_*^{jm} = n$ para algum $n \in \mathbb{N}$, e para todo m . Segue do Teorema 5.2.5, que a menos de uma subsequência

$$\lim_{m \rightarrow \infty} x_{k_*^{jm}}^{\delta^{jm}} = \lim_{m \rightarrow \infty} x_n^{\delta^{jm}} = x_n.$$

Note que x_n é uma solução de (11). De fato, para $i \in \{0, \dots, N-1\}$ temos

$$\begin{aligned} \|A_i x_n - b_i\| &= \lim_{m \rightarrow \infty} \|A_i x_{k_*^{jm}}^{\delta^{jm}} - b_i\| \leq \lim_{m \rightarrow \infty} \left(\|A_i x_{k_*^{jm}}^{\delta^{jm}} - b_i^{\delta^{jm}}\| + \|b_i^{\delta^{jm}} - b_i\| \right) \\ &\leq \lim_{m \rightarrow \infty} (\tau + 1) \delta^{jm} = 0. \end{aligned}$$

No segundo caso assuma que k_*^j não é limitada. Assim existe uma subsequência crescente de k_*^j , denotada também por k_*^j , tal que $k_*^j \rightarrow \infty$ quando $j \rightarrow \infty$. Vamos provar agora que $x_{k_*^j}^{\delta^j}$ tem uma subsequência convergindo para x^\dagger . Isso provará que toda subsequência de $x_{k_*^j}^{\delta^j}$ tem ela própria uma subsequência convergente para x^\dagger , o que provará o resultado.

Fixe $\varepsilon > 0$ e seja x_k a sequência sem ruídos construída no Teorema 5.2.5. Segue do Teorema 5.2.2 que $x_k \rightarrow x^\dagger$, onde x^\dagger é uma solução x_0 -mínima de (11). Então existe um índice $L = L(\varepsilon) \in \mathbb{N}$ tal que,

$$\|x_k - x^\dagger\| < \frac{\varepsilon}{2}, \quad k \geq L. \quad (104)$$

Além disso, como $k_*^j \rightarrow \infty$ quando $j \rightarrow \infty$, existe $J \in \mathbb{N}$ tal que $k_*^j \geq L$, para todo $j \geq J$. Com isso, segue do Corolário 5.1.5 que

$$j \geq J \implies \|x_{k_*^j}^{\delta^j} - x^\dagger\| \leq \|x_L^{\delta^j} - x^\dagger\|,$$

e a partir do Teorema 5.2.5 segue a existência de uma subsequência δ^{jm} (dependendo de $L(\varepsilon)$ e de $M \in \mathbb{N}$) tal que

$$m \geq M \implies \|x_L^{\delta^{jm}} - x_L\| < \frac{\varepsilon}{2}.$$

Dessa forma, para $m \geq \max\{J, M\}$ (o que implica $j_m \geq \max\{J, M\}$) temos

$$\|x_{k_*^{j_m}}^{\delta^{j_m}} - x^\dagger\| \leq \|x_L^{\delta^{j_m}} - x^\dagger\| \leq \|x_L^{\delta^{j_m}} - x_L\| + \|x_L - x^\dagger\| < \varepsilon. \quad (105)$$

Note que a subsequência δ^{j_m} depende de ε . Vamos construir uma subsequência independente: escolha $\varepsilon = 1$, então podemos encontrar uma subsequência δ^j e um número $j_1 \in \mathbb{N}$ tal que

$$\|x_{k_*^{j_1}}^{\delta^{j_1}} - x^\dagger\| < 1. \quad (106)$$

Como a subsequência atual δ^j é uma sequência de números reais positivos convergindo a zero, a ideia anterior pode ser aplicada para $\varepsilon = \frac{1}{2}$ usando uma subsequência da sequência atual. Podemos escolher $j_2 \geq j_1$ tal que

$$\|x_{k_*^{j_2}}^{\delta^{j_2}} - x^\dagger\| < \frac{1}{2}. \quad (107)$$

Usando indução, é possível construir uma subsequência δ^{j_m} com a propriedade

$$\|x_{k_*^{j_m}}^{\delta^{j_m}} - x^\dagger\| < \frac{1}{m}, \quad \text{para todo } m \in \mathbb{N}, \quad (108)$$

isso implica que $\|x_{k_*^{j_m}}^{\delta^{j_m}} - x^\dagger\| \rightarrow 0$, quando $m \rightarrow \infty$. □

6 CONCLUSÕES E TRABALHOS FUTUROS

Na primeira parte deste trabalho nós analisamos versões inexatas de métodos tipo Tikhonov, usados como técnicas de regularização para resolver problemas lineares mal postos. Os minimizadores dos funcionais de Tikhonov são computados de forma inexata usando uma iteração interna. O método a ser usado na iteração interna deve satisfazer a Hipótese 3.0.1 (no caso do método de Tikhonov de um passo) ou a Hipótese 4.1.1 (no caso de métodos de Tikhonov iterado). A iteração interna termina conforme a regra (39) (no caso de métodos de Tikhonov de um passo) ou (55) (no caso de métodos de Tikhonov iterado). Muitos métodos conhecidos de otimização convexa, satisfazem as hipóteses necessárias e portanto podem ser usados na iteração interna para aproximar os minimizadores, veja a Seção 4.3. No Capítulo 3 analisamos uma versão do método de Tikhonov de um passo penalizado por uma função uniformemente convexa. Neste capítulo consideramos ainda as regras a priori (4) e a posteriori (7) para computar os parâmetros de regularização. Para essa versão do método de Tikhonov de um passo, obtemos os seguintes resultados:

- *Boa definição do método, Lema 3.2.1;*
- *Propriedades de regularização, Teoremas 3.1.1 e 3.2.2;*
- *Taxas de convergência ótima (sob a hipótese de condição de fonte), Teoremas 3.1.2 e 3.2.3.*

No Capítulo 4 analisamos uma versão inexata do método Tikhonov Iterado penalizado com a distância de Bregman induzida por um funcional fortemente convexo. A sequência dos parâmetros de regularização é escolhida a posteriori, conforme (56). Para esse caso obtemos os seguintes resultados:

- *Boa definição do método, o índice de parada do método é finito e a monotonia da iteração, Lema 4.2.1, Proposição 4.2.3 e Corolário 4.2.4 e Proposição 4.2.10;*
- *Estabilidade do método, Teorema 4.2.7;*
- *Regularização do método, Teorema 4.2.13.*

Na segunda parte deste trabalho, investigamos uma versão de Kaczmarz do método de Tikhonov iterado não estacionário usado para computar soluções aproximadas estáveis de sistemas de equações lineares mal postos, veja Capítulo 5. A maior vantagem deste algoritmo está na estratégia usada para escolher os multiplicadores de Lagrange, a qual nos permite provar convergência do método para dados exatos de uma maneira diferente como foi feito em [6], veja Teorema 5.2.2. Além disso, provamos os seguintes resultados:

- *O índice de parada do método é finito, Corolário 5.1.8;*
- *Convergência para dados exatos ($\delta = 0$), Teorema 5.2.2;*
- *Estabilidade do método, Teorema 5.2.5;*
- *Regularização do método, Teorema 5.2.6.*

Algumas possíveis extensões deste trabalho são:

- *Estudar uma versão não linear do método Tikhonov iterado com minimização inexata;*
- *Estudar e analisar uma versão não linear do método TIKr;*
- *Estudar e analisar uma versão randômica do método TIKr;*
- *Estudar e analisar uma versão randômica não linear do método TIKr.*

Apêndices

APÊNDICE A – APLICAÇÕES

Vamos apresentar dois modelos de aplicações para o Algoritmo 3 (TIKr) analisado no Capítulo 5, a saber: O problema de desfoque de imagem (deblurring), [36] e o Problema do Potencial Inverso, [51]. Em [2] é feita uma análise qualitativa e quantitativa da performance do algoritmo 3 aplicado a esses dois modelos. Neste Apêndice apresentaremos apenas a análise quantitativa, para mais detalhes sobre tais aplicações bem como figuras mostrando o número de iterações por ciclo e o erro relativo $\|x^* - x_k^\delta\|/\|x^*\|$, veja [2]. O cálculo de nossos experimentos numéricos foi conduzido usando o MATLAB 2012a.

A performance do método TIKr é comparada com outros três métodos tipo Kaczmarz, a saber: O Tikhonov iterado geométrico Kaczmarz (giTK) com $\lambda_k = 2^k$, o método Tikhonov iterado estacionário Kaczmarz (siTK) com $\lambda_k = 2$ e o Landweber Kaczmarz (LWK). A iteração do (LWK) usando a notação $i = \lfloor k \rfloor$ é dada por

$$x_{k+1}^\delta = x_k^\delta + A_i^*(b_i^\delta - A_i x_k^\delta), \text{ para cada } k \in \mathbb{N}.$$

A.1 DESFOQUE DE IMAGEM

Esses são problemas de dimensão finita modelados em geral, por sistemas lineares $Ax = b$ de alta dimensão. Aqui o vetor $x \in X = \mathbb{R}^n$ representa os valores de pixel de uma imagem verdadeira desconhecida, enquanto o vetor $b \in Y = \mathbb{R}^n$ representa os valores de pixel da imagem observada (desfocada). O operador $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ descreve o fenômeno de desfoque (discretizado) [35, 36].

Na configuração contínua, o processo de desfoque é representado por um operador integral do tipo convolução. Assim, o modelo matemático corresponde a uma equação integral de primeiro tipo, veja [9]. Na configuração discreta, depois de incorporar as condições de contorno adequadas no modelo, a convolução discreta deve ainda ser avaliada de um algoritmo apropriado.

Em nossos experimentos de remoção de borrão, veja [2], consideramos a seguinte configuração:

- A imagem verdadeira $x \in \mathbb{R}^n$, $n = 256^2$ (Cameraman 256×256);
- A função de propagação de pontos (point spread function) é rotacionalmente simétrico ao filtro gaussiano de tamanho $[256 \times 256]$ e desvio padrão $\sigma = 4$;
- Os dados exatos $b = Ax \in \mathbb{R}^n$ (imagem desfocada).

Os operadores lineares $A_j : \mathbb{R}^{16} \rightarrow \mathbb{R}^{16}$ em (11) correspondem aos blocos de linhas do operador desfoque A (i.e., $N = 16$ blocos com 32 linhas cada), enquanto os dados b_j são definidos adequadamente de modo que $A_j x = b_j$. Observe que cada ciclo

consiste em 16 etapas iterativas. Na implementação do método TIKr para o problema em questão usamos os valores $\eta = 0.1$ e $\bar{\eta} = 0.5$.

Nos experimentos comparamos o método TIKr com os métodos LWK, siTK com $\lambda_k = 2$ e giTK com $\lambda_k = 2^k$ usando o nível de ruído $\delta_j = 1\%$. As tabelas a seguir representam o erro relativo $\|x^* - x_k^\delta\|/\|x^*\|$ (calculado ao final do ciclo apresentado na primeira coluna da tabela) e número de equações resolvidas (calculadas ao final do ciclo apresentado na primeira coluna da tabela) respectivamente:

Ciclo	TIKr	LWK	siTK	giTK
10	0.05	0.22	0.055	0.05
30	–	0.20	0.055	0.045
50	–	0.17	0.052	0.045

Tabela 1 – Erro relativo para o nível de ruídos $\delta_j = 1\%$.

Ciclo	TIKr	LWK	siTK	giTK
10	7	16	14	14
30	–	16	14	10
50	–	16	14	7

Tabela 2 – O número de equações resolvidas por ciclos.

O método TIKr atinge o critério de parada no ciclo 24. Ele mostrou bom desempenho em relação aos métodos LWK, siTK e giTK quando comparado o erro relativo e número de equações resolvidas por ciclo.

No próximo experimento usamos ruído relativo $\|b_i - b_i^\delta\|/\|b_i\|$ de três níveis distintos, a saber 0.1%, 1% e 10%. Esses números representam o quão desfocada está a imagem a ser melhorada, dito de outra forma, representam o quão "distante" está a imagem desfocada da imagem verdadeira. A tabela 3 abaixo compara o método TIKr com o giTK, onde o número de ciclos computados para cada método é mostrado, bem como o número de iterações realizadas. Vale ressaltar que os métodos LWK e siTK não atingiram os critérios de parada para os níveis de ruídos analisados (esses métodos foram interrompidos após 10^4 ciclos e os resultados obtidos foram omitidos).

δ	TIKr	giTK
10%	15 (124)	111 (850)
1%	24 (212)	116 (939)
0.1%	29 (262)	119 (1023)

Tabela 3 – Imagem desfocada: Primeiro o número de ciclos computados; entre parênteses o número de iterações computadas.

Reiteramos que nos experimentos o parâmetro λ_{\max} não desempenha papel tendo em vista a Observação 5.1.2. No método TIKr os multiplicadores de Lagrange foram calculados de forma similar a estratégia apresentada em [6].

A.2 PROBLEMA DO POTENCIAL INVERSO

O problema do Potencial Inverso (IPP) é um problema de identificação de parâmetros para EDP's elípticas [45, 39, 51, 62]. Generalizações deste problema inverso aparecem em muitas aplicações relevantes incluindo Gravimetria Inversa [53, 62], eletroencefalografia (EEG) [43] e eletromiográfico (EMG) [63].

O problema direto consiste em resolver num domínio Lipschitz $\Omega \subset \mathbb{R}^d$, para uma determinada função de fonte $x \in L_2(\Omega)$, o problema de contorno

$$-\Delta u = x, \text{ em } \Omega, \quad u = 0 \text{ em } \partial\Omega. \quad (109)$$

O problema inverso correspondente é chamado de problema do Potencial Inverso (IPP), que consiste em recuperar uma função x em L_2 , a partir da medida dos dados de Neumann do seu potencial correspondente $u \in H^1$ na fronteira de Ω , i.e., $b := u_\nu|_{\partial\Omega} \in L_2(\partial\Omega)$.

O IPP é modelado pelo operador linear $A : L_2(\Omega) \rightarrow L_2(\partial\Omega)$ definido por $Ax := u_\nu|_{\partial\Omega}$ onde $u \in H_0^1(\Omega)$ é a única solução de (109), veja [51]. Usando essa notação, o IPP pode ser escrito na forma abreviada $Ax = b$

Nossos experimentos são inspirados em [6]. Na configuração experimental escolhemos $\Omega = (0, 1) \times (0, 1)$ e assumindo que o parâmetro procurado x^* é uma função H^1 . A função x^* é mostrada na figura abaixo, bem como o iterado x_{72}^δ e o erro de iteração $\|x^* - x_{72}^\delta\|$ para o nível de ruído $\delta = 0.1\%$, obtidos utilizando o algoritmo TIKr com as configurações descritas abaixo.

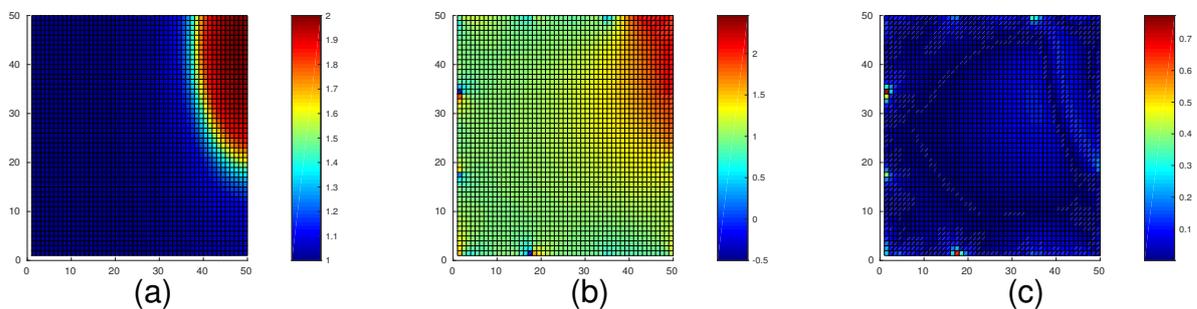


Figura 1 – IPP: Cenário com ruídos $\delta = 0.1\%$. (a) Solução exata x^* ; (b) Solução aproximada x_{72}^δ (método TIKr); (c) Erro de iteração $\|x^* - x_{72}^\delta\|$ (erro absoluto em termos de pixels).

Na configuração discreta, a solução dos problemas de valor de contorno elípticos envolvidos é calculada usando diferenças finitas em uma malha uniforme com 50^2 nós. Assim o operador A é aproximado por uma matriz $A_D : \mathbb{R}^{2500} \rightarrow \mathbb{R}^{192}$, a fronteira

de Ω é dividida em $N = 12$ segmentos, i.e. $\partial\Omega = \cup_{i=0}^{11} \Gamma_i$, e cada Γ_i é discretizado usando fronteiras com 16 nós.

Dessa forma, os operadores lineares em (11) são definidos por $A_i := \gamma_i \circ A_d$, onde $\gamma_i : \mathbb{R}^{192} \rightarrow \mathbb{R}^{16}$ é a discretização correspondente do operador de projeção em $\partial\Omega$ para Γ_i . O dado de fronteira (discretizado) $\gamma_i(b) = b_j$ são definidos de forma ainda que $A_i x = b_j$.

Os dados perturbados b_j^δ são gerados adicionando aos dados de Neumann exatos b_j um ruído normalmente distribuído com média zero e variância adequada para alcançar um nível de ruído relativo prescrito. Nas implementações numéricas que definimos $\eta = 0.1$, $\bar{\eta} = 0.5$, $\tau = 2$ e o ponto inicial $x_0 \equiv 1.5$ (a função constante em Ω).

Assim como na Seção A.1, três cenários distintos são considerados para os métodos TIKr, giTK, siTK e LWK onde o nível de ruído relativo $\|b - b^\delta\|/\|b\|$ corresponde a 1%, 0.1% e 0.025% respectivamente. Os resultados são resumidos na Tabela 4 abaixo.

δ	TIKr	giTK	siTK	LWK
1%	2 (10)	3 (21)	6 (32)	10 (56)
0.1%	6 (43)	5 (55)	20 (165)	35 (298)
0.025%	7 (64)	7 (73)	44 (358)	88 (669)

Tabela 4 – IPP: Primeiro o número de ciclos computados; entre parênteses o número de iterações computadas.

Note que o TIKr é um método com o cálculo dos multiplicadores dado a posteriori enquanto que giTK, siTK e LWK são todos a priori. Portanto, possivelmente precisamos minimizar mais de um funcional por iteração no TIKr, já os demais exigem apenas uma minimização por iteração. Mesmo tendo que minimizar, em geral, mais de um funcional por iteração o TIKr computa (veja Tabela 4) uma quantidade de ciclo/iteração compatível ao giTK, que é um método de escolha a priori dos multiplicadores. Levando em consideração o esforço computacional por iteração e número de ciclos/iterações computadas o TIKr apresentou uma performance satisfatória quando comparado com os métodos giTK, siTK e LWK.

No próximo experimento comparamos o método TIKr com os métodos LWK, siTK $\lambda_k = 2$ e giTK $\lambda_k = 2^k$ no cenário com ruídos $\delta = 0.1\%$. As tabelas a seguir apresentam o erro relativo $\|x^* - x_k^\delta\|/\|x^*\|$ (calculado ao final do ciclo apresentado na primeira coluna da tabela) e número de equações resolvidas (calculadas ao final do ciclo apresentado na primeira coluna da tabela) respectivamente:

Ciclo	TIKr	LWK	siTK	giTK
4	0.1	0.22	0.16	0.1
12	–	0.18	0.11	–
20	–	0.11	0.1	–

Tabela 5 – Erro de iteração para o nível de ruídos $\delta_j = 0.1\%$.

Ciclo	TIKr	LWK	siTK	giTK
4	3	12	12	6
12	–	12	8	–
20	–	–	–	–

Tabela 6 – O número de equações resolvidas por ciclos.

Nesse modelo, o método TIKr atinge o critério de parada no ciclo 6 e o giTK no ciclo 5. O problema direto e o problema inverso foram resolvidos usando diferentes níveis de discretização. Para o problema de potencial inverso, em todos os cenários de ruídos, ambos TIKr e giTK requerem número semelhante de ciclos para atingir os mesmos critérios de parada. Por outro lado, para o problema de imagem desfocada o TIKr usa menos ciclos do que o giTK. É importante notar que para ambos os problemas inversos o TIKr computa menos passos do que o giTK (compare os números entre parênteses nas Tabelas 3 e 4), desse ponto de vista o TIKr demonstra maior desempenho do que o giTK.

Nas implementações numéricas do Algoritmo 3 para ambos os problemas inversos acima, é possível verificar que as desigualdades no passo [3.2] são satisfeitas, veja [2], assim como podemos contar quantas iterações foram efetivamente computadas (ou seja, $h_k \neq 0$) em cada um dos ciclos do método TIKr.

BIBLIOGRAFIA

- 1 *F. Margotti and J. Rabelo. Tikhonov-like methods whit inexact minimização for solving linear ill-posed problems. Inverse Problems. 36(12)- 125013, 2020.*

- 2 *Filippozi. Rafaela, J. Rabelo, Boiger. Romana, A. Leitão. A range-relaxed criteria for choosing the Lagrange multipliers in the iterated Tikhonov Kaczmarz method for solving systems of linear ill-posed equations. Accepted for publication. Inverse Problems. 37 (04)-045005, 2020.*

- 3 *Maryam A. Alghamdi, Abdullah Alotaibi, Patrick L. Combettes, and Naseer Shahzad. A primal-dual method of partial inverses for composite inclusions. Optim. Lett., 8(8):2271–2284, 2014.*

- 4 *Maicon Marques Alves and Samara Costa Lima. An inexact Spingarns partial inverse method with applications to operator splitting and composite optimization. J Optim Theory Appl, 175:818–847, 2017.*

- 5 *Heinz H. Bauschke and Patrick L. Combettes. Convex analysis and monotone operator theory in Hilbert spaces. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011. With a foreword by Hédý Attouch.*

- 6 *R. Boiger, A. Leitão, and B. F. Svaiter. Range-relaxed criteria for choosing the Lagrange multipliers in nonstationary iterated Tikhonov method. IMA J. Numer. Anal., 40(1):606–627, 2020.*

- 7 *Marco Donatelli and Martin Hanke. Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring. Inverse Problems, 29(9):095008, aug 2013.*

- 8 *Jonathan Eckstein and Dimitri P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Programming, 55(3, Ser. A):293– 318, 1992.*

- 9 *Heinz W. Engl, Martin Hanke, and Andreas Neubauer. Regularization of inverse problems, volume 375 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1996.*

- 10 Heinz W. Engl, Karl Kunisch, and Andreas Neubauer. Convergence rates for Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, 5(4):523–540, 1989.
- 11 Max L. N. Gonçalves, Jefferson G. Melo, and Renato D. C. Monteiro. Improved pointwise iteration- complexity of a regularized ADMM and of a regularized non-Euclidean HPE framework. *SIAM J. Optim.*, 27(1):379–407, 2017.
- 12 M. Hanke and C. W. Groetsch. Nonstationary iterated Tikhonov regularization. *J. Optim. Theory Appl.*, 98(1):37–53, 1998.
- 13 Qinian Jin and Linda Stals. Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces. *Inverse Problems*, 28(10):104011, 15, 2012.
- 14 Qinian Jin and Hongqi Yang. Levenberg-Marquardt method in Banach spaces with general convex regularization terms. *Numerische Mathematik*, 133(4):655–684, Aug 2016.
- 15 Qinian Jin and Min Zhong. Nonstationary iterated Tikhonov regularization in Banach spaces with uniformly convex penalty terms. *Numer. Math.*, 127(3):485–513, 2014.
- 16 Barbara Kaltenbacher, Andreas Neubauer, and Otmar Scherzer. *Iterative regularization methods for nonlinear ill-posed problems, volume 6 of Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH Co. KG, Berlin, 2008.
- 17 Andreas Kirsch. *An introduction to the mathematical theory of inverse problems, volume 120 of Applied Mathematical Sciences*. Springer, New York, second edition, 2011.
- 18 P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.
- 19 M. P. Machado, F. Margotti, and A. Leitão. On the choice of Lagrange multipliers in the iterated Tikhonov method for linear ill-posed equations in Banach spaces. *Inverse Probl. Sci. Eng.*, 28(6):796– 826, 2020.

- 20 F. Margotti, A. Leitão and B.F. Svaiter. Range-relaxed criteria for choosing the Lagrange multipliers in Levenberg-Marquardt method. *IMA Journal of Numerical Analysis*. DOI 10.1093/imanum/draa050, 2020.
- 21 M. P. Machado, F. Margotti, and Antonio Leitão. On nonstationary iterated Tikhonov methods for ill-posed equations in Banach spaces. In *New trends in parameter identification for mathematical models*, Trends Math., pages 175–193. Birkhäuser/Springer, Cham, 2018.
- 22 Fábio Margotti. *On inexact Newton methods in Banach spaces*. PhD. Thesis, Karlsruhe, 2015.
- 23 Fábio Margotti and Andreas Rieder. An inexact Newton regularization in Banach spaces based on the nonstationary iterated Tikhonov method. *J. Inverse Ill-Posed Probl.*, 23(4):373–392, 2015.
- 24 Renato D. C. Monteiro and B. F. Svaiter. On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean. *SIAM J. Optim.*, 20(6):2755–2787, 2010.
- 25 Renato D. C. Monteiro and B. F. Svaiter. Complexity of variants of Tseng's modified F-B splitting and Korpelevich's methods for hemivariational inequalities with applications to saddle-point and convex optimization problems. *SIAM J. Optim.*, 21(4):1688–1720, 2011.
- 26 Renato D. C. Monteiro and Benar F. Svaiter. Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers. *SIAM J. Optim.*, 23(1):475–507, 2013.
- 27 Andreas Rieder. *Keine Probleme mit inversen Problemen*. Friedr. Vieweg Sohn, Braunschweig, 2003. *Eine Einführung in ihre stabile Lösung*. [An introduction to their stable solution].
- 28 Otmar Scherzer. Convergence rates of iterated Tikhonov regularized solutions of nonlinear ill-posed problems. *Numer. Math.*, 66:259–279, 1993.

- 29 Barbara Schuster, Thomas and, Bernd Hofmann, and Kamil S. Kazimierski. *Regularization methods in Banach spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH Co. KG, Berlin, 2012.
- 30 M. V. Solodov and B. F. Svaiter. *A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator*. *Set-Valued Anal.*, 7(4):323–345, 1999.
- 31 Paul Tseng. *A modified forward-backward splitting method for maximal monotone mappings*. *SIAM J. Control Optim.*, 38(2):431–446, 2000.
- 32 R.S. Burachik, A.N. Iusem, and B.F. Svaiter. *Enlargement of mono- tone operators with applications to variational inequalities*. *Set-Valued Analysis*, 5:159–180, 1997.
- 33 A.B. Bakushinsky and M.Y. Kokurin, *Iterative Methods for Approximate Solution of In- verse Problems, Mathematics and Its Applications*, vol. 577, Springer, Dordrecht, 2004.
- 34 J. Baumeister, *Stable Solution of Inverse Problems, Advanced Lectures in Mathematics*, Friedr. Vieweg Sohn, Braunschweig, 1987. MR 889048.
- 35 M. Bertero, *Image deblurring with Poisson data: from cells to galaxies*, *Inverse Problems* 25 (2009), no. 12, 123006.
- 36 M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging, Advanced Lectures in Mathematics*, IOP Publishing, Bristol, 1998.
- 37 M. Brill and E. Schock, *Iterative solution of ill-posed problems: A survey*, ch. in *Model Optimization in Exploration Geophysics*, Ed. A. Vogel, pp. 13–38, Vieweg, Braunschweig, 1987.
- 38 M. Burger and B. Kaltenbacher, *Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems*, *SIAM J. Numer. Anal.* 44 (2006), 153–182.
- 39 A. De Cezaro, A. Leitão, and X.-C. Tai, *On multiple level-set regularization methods for inverse problems*, *Inverse Problems* 25 (2009), 035004.

40 A. De Cezaro, J. Baumeister, and A. Leitão, *Modified iterated Tikhonov methods for solving systems of nonlinear ill-posed equations*, *Inverse Probl. Imaging* 5 (2011), no. 1, 1–17.

41 A. De Cezaro, M. Haltmeier, A. Leitão, and O. Scherzer, *On steepest-descent-Kaczmarz methods for regularizing systems of nonlinear ill-posed equations*, *Appl. Math. Comput.* 202 (2008), no. 2, 596–607.

42 Marco Donatelli, *On nondecreasing sequences of regularization parameters for nonstationary iterated Tikhonov*, *Numer. Algorithms* 60 (2012), no. 4, 651–668. MR 2946487.

43 A. El Badia and M. Farah, *Identification of dipole sources in an elliptic equation from boundary measurements: application to the inverse EEG problem*, *J. Inverse Ill-Posed Probl.* 14 (2006), no. 4, 331–353.

44 A.G. Fakeev, *A class of iterative processes for solving degenerate systems of linear algebraic equations*, *U.S.S.R. Comput. Math. Math. Phys.* 21 (1981), no. 3, 15–22.

45 F. Frühaufer, O. Scherzer, and A. Leitão, *Analysis of Regularization Methods for the Solution of Ill-Posed Problems Involving Discontinuous Operators*, *SIAM J. Numer. Anal.* 43 (2005), 767–786.

46 C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, *Research Notes in Mathematics*, vol. 105, Pitman (Advanced Publishing Program), Boston, MA, 1984.

47 M. Haltmeier, R. Kowar, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. II. Applications*, *Inverse Probl. Imaging* 1 (2007), no. 3, 507–523.

48 M. Haltmeier, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. I. convergence analysis*, *Inverse Probl. Imaging* 1 (2007), no. 2, 289–298.

- 49 M. Haltmeier, A. Leitão, and E. Resmerita, *On regularization methods of EM-Kaczmarz type*, *Inverse Problems* 25 (2009), 075008.
- 50 M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of Landweber iteration for nonlinear ill-posed problems*, *Numer. Math.* 72 (1995), 21–37.
- 51 F. Hettlich and W. Rundell, *Iterative methods for the reconstruction of an inverse potential problem*, *Inverse Problems* 12 (1996), 251–266.
- 52 B. Hofmann, *Regularization for Applied Inverse and Ill-Posed Problems*, *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*, vol. 85, BSB B. G. Teubner Verlags- gesellschaft, Leipzig, 1986, *A numerical approach*, With German, French and Russian summaries.
- 53 Victor Isakov, *Inverse Problems for Partial Differential Equations*, second ed., *Applied Mathematical Sciences*, vol. 127, Springer, New York, 2006.
- 54 S. Kaczmarz, *Approximate solution of systems of linear equations*, *Internat. J. Control* 57 (1993), no. 6, 1269–1271.
- 55 L. J. Lardy, *A series representation for the generalized inverse of a closed linear operator*, *Atti della Accademia Nazionale dei Lincei, Rendiconti della Classe di Scienze Fisiche, Matematiche, e Naturali, Serie VIII* 58 (1975), 152–157.
- 56 A. Louis, *Inverse und schlecht gestellte Probleme*, B.G. Teubner, Stuttgart, 1989.
- 57 V.A. Morozov, *Regularization Methods for Ill-Posed Problems*, CRC Press, Boca Raton, 1993.
- 58 A. Leitão, F. Margotti, and B. F. Svaiter, *Range-relaxed criteria for choosing the Lagrange multipliers in the Levenberg-Marquardt method*, *IMA J. of Numerical Analysis* (2020), to appear.
- 59 F. Natterer, *The mathematics of computerized tomography*, SIAM, Philadelphia, 2001.

-
- 60 F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM, Philadelphia, 2001.
- 61 A. Leitão, F. Margotti, and B. F. Svaiter, *Range-relaxed criteria for choosing the Lagrange multipliers in the Levenberg-Marquardt method*, *IMA J. of Numerical Analysis* (2020), to appear.
- 62 K. van den Doel, U. M. Ascher, and A. Leitão, *Multiple Level Sets for Piecewise Constant Surface Reconstruction in Highly Ill-Posed Problems*, *Journal of Scientific Computing* 43 (2010), no. 1, 44–66.
- 63 Kees van den Doel, Uri M. Ascher, and Dinesh K. Pai, *Computed myography: three-dimensional reconstruction of motor functions from surface EMG data*, *Inverse Problems* 24 (2008), no. 6, 065010, 17.
- 64 J. Hadamard. *Lectures on the Cauchy Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, 1923.