

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE MATEMÁTICA**

Samara Vendramin Pieta

CÁLCULO PROXIMAL EM OTIMIZAÇÃO CONVEXA

Florianópolis

2013

Samara Vendramin Pieta

CÁLCULO PROXIMAL EM OTIMIZAÇÃO CONVEXA

Dissertação submetida ao Programa
de Pós-Graduação em Matemática Pura
e Aplicada para a obtenção do Grau
de Mestre em Matemática.
Orientador: Prof. Dr. Maicon Mar-
ques Alves (UFSC)

Florianópolis

2013

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

Pieta, Samara Vendramin

Cálculo Proximal em Otimização Convexa [dissertação] /
Samara Vendramin Pieta ; orientador, Maicon Marques Alves
- Florianópolis, SC, 2013.

85 p. ; 21cm

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro de Ciências Físicas e Matemáticas.
Programa de Pós-Graduação em Matemática Pura e
Aplicada.

Inclui referências

1. Matemática Pura e Aplicada. 2. Otimização convexa. 3.
Operador não-expansivo ponderado. 4. Operador proximal. I.
Marques Alves, Maicon . II. Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em Matemática Pura e
Aplicada. III. Título.

Samara Vendramin Pieta

CÁLCULO PROXIMAL EM OTIMIZAÇÃO CONVEXA

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Matemática”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Matemática Pura e Aplicada.

Florianópolis, 20 de fevereiro 2013.

Prof. Dr. Daniel Gonçalves
Coordenador do Curso

Banca Examinadora:

Prof. Dr. Maicon Marques Alves (UFSC)
Orientador

Prof. Dr. Antonio Carlos Gardel Leitão (UFSC)

Prof. Dr. Jefferson Gonçalves de Melo (UFG)

Profa. Dra. Melissa Weber Mendonça (UFSC)

AGRADECIMENTOS

À minha família, em especial aos meus pais, Stelamar e Hélio, por todo o apoio necessário para a conclusão de mais esta etapa.

Ao meu orientador, Maicon Marques Alves, pela atenção, e principalmente, pela dedicação e presteza na orientação deste trabalho.

Aos colegas, pelas trocas de experiências e conhecimentos, que muito me ajudaram no decorrer deste curso. Em especial, à Asteroide Santana, que me estendeu a mão inúmeras vezes, agradeço por toda ajuda e paciência.

À todos os professores, que tive o prazer de encontrar durante a minha vida acadêmica, pelos ensinamentos e contribuições.

Aos professores Antonio Carlos Gardel Leitão, Jefferson Gonçalves de Melo e Melissa Weber Mendonça, por gentilmente terem aceitado participar da banca examinadora.

Aos professores e funcionários do Departamento de Matemática da Universidade Federal de Santa Catarina. Em particular, à Elisa Barbosa Amaral.

À CAPES, pelo suporte financeiro nesses dois anos.

À todos, que de alguma forma contribuíram para esta realização, muito obrigada!

RESUMO

O operador proximal, introduzido por Moreau em 1962, é uma ferramenta importante na análise e solução numérica de problemas de otimização convexa. Neste trabalho, apresentamos a teoria baseada na noção de operadores proximais, utilizada para estudar o problema de minimizar a soma de duas funções convexas com certas propriedades de regularidade, em espaços de Hilbert. Analisamos a convergência de um algoritmo forward-backward e uma aplicação em problemas de recuperação de sinais.

Palavras-chave: Espaços de Hilbert, otimização convexa, operador não-expansivo ponderado, operador proximal, decomposição de Moreau, algoritmo forward-backward, recuperação de sinais.

ABSTRACT

The proximity operator, introduced by Moreau in 1962, is an important tool in the analysis and numerical solution of convex optimization problems. In this work, we present the theory based on the notion of proximity operators, used to study the problem of minimizing the sum of two convex functions with certain regularity properties, in Hilbert spaces. We also present a convergence analysis of a forward-backward algorithm and an application in signal recovery problems.

Keywords: Hilbert spaces, convex optimization, averaged nonexpansive operator, proximity operator, Moreau's decomposition, forward-backward algorithm, signal recovery

SUMÁRIO

INTRODUÇÃO	15
1 CONCEITOS BÁSICOS DE ANÁLISE CONVEXA ..	19
1.1 FUNÇÕES CONVEXAS	19
1.2 OTIMIZAÇÃO CONVEXA	22
1.3 CONVOLUÇÃO INFIMAL	24
1.4 CONJUGADA DE FENCHEL	26
1.5 SUBDIFERENCIAL	28
2 APROXIMAÇÃO DE PONTO FIXO DE OPERA- DORES NÃO-EXPANSIVOS	33
2.1 OPERADORES NÃO-EXPANSIVOS	33
2.2 OPERADORES PONDERADOS	35
2.3 SEQUÊNCIAS FEJÉR-MONÓTONAS	39
2.4 ITERAÇÕES DE PONTO FIXO	42
3 OPERADOR PROXIMAL	47
3.1 REGULARIZAÇÃO DE MOREAU	47
3.2 CÁLCULO PROXIMAL	53
3.3 DECOMPOSIÇÃO DE MOREAU	58
3.4 EXEMPLOS DE OPERADORES PROXIMAIS	62
4 MINIMIZANDO A SOMA DE DUAS FUNÇÕES CON- VEXAS	71
4.1 ALGORITMO FORWARD-BACKWARD	71
4.2 APLICAÇÃO EM PROBLEMAS DE RECUPERAÇÃO DE SINAIS	74
CONSIDERAÇÕES FINAIS	81
REFERÊNCIAS	83

INTRODUÇÃO

Desde o princípio, procuramos formas instintivas para obter o melhor resultado possível em qualquer de nossas atividades, como por exemplo, escolher entre várias possibilidades um trajeto mais rápido ou mais curto entre a casa e o trabalho. Para problemas mais complexos, a intuição deu lugar ao uso de artifícios técnicos desenvolvidos para otimizar essas atividades, dando origem à otimização matemática.

A otimização é o ramo da matemática que estuda problemas cujo objetivo é determinar o menor ou o maior valor que uma função pode assumir dentro de um conjunto viável. Apesar da sua importância nos dias atuais, historicamente os primeiros resultados para o desenvolvimento desta área, propostos por George Bernard Dantzig (1914-2005), só apareceram na década de 40, após a Segunda Guerra Mundial; até então era restrita às operações militares.

Uma área especial da otimização matemática é a otimização convexa, que trata de problemas onde a função e o conjunto viável são ambos convexos. Devido à teoria bem desenvolvida e à grande frequência com que ocorrem na prática, já existem vários métodos para resolver problemas desse tipo. As aplicações para essa classe estão em diversas áreas, por exemplo, na engenharia: em sistemas de controle, análise e processamento de sinais, análise de circuitos, projeto de sistemas de energia elétrica, etc.

Uma aplicação da otimização convexa que vem crescendo consideravelmente nas últimas décadas é na solução de problemas que buscam determinar causas desconhecidas a partir de efeitos medidos ou observados, conhecidos como problemas inversos. Podemos citar, problemas em recuperação de sinais, que abrangem problemas em que um sinal multidimensional deve ser deduzido a partir da observação de dados, consistindo de sinais fisicamente ou matematicamente relacionados a ele [24]. Um exemplo prático é a reconstrução tomográfica [14]. Matematicamente, problemas em recuperação de sinais podem ser formulados como o problema de minimizar a soma de duas funções em um espaço de Hilbert com certas propriedades de regularidade, com o objetivo de incorporar informações a priori e impor algum grau de consistência nos dados observados. Uma dificuldade que surge na resolução dos mesmos decorre do fato que, normalmente, uma das funções não é diferenciável, o que exclui as técnicas convencionais de otimização.

O objetivo deste trabalho é apresentar com detalhes os concei-

tos teóricos utilizados para provar propriedades como existência, unicidade e caracterização de soluções, além da convergência de um método numérico, para problemas que aparecem no seguinte formato:

$$(P) \quad \underset{x \in \mathcal{H}}{\text{minimize}} \quad f_1(x) + f_2(x),$$

onde $f_1 : \mathcal{H} \rightarrow (-\infty, +\infty]$ e $f_2 : \mathcal{H} \rightarrow \mathbb{R}$ são funções convexas, semi-contínuas inferiormente e próprias, tal que f_2 é diferenciável em \mathcal{H} com gradiente $(1/\beta)$ -Lipschitz contínuo, para algum $\beta \in \mathbb{R}_{++}$. Apesar de sua simplicidade, o Problema (P) cobre uma ampla gama de formulações aparentemente não relacionadas com recuperação de sinais, incluindo problemas de mínimos quadrados restrito [23], de regularização esparsa [12], de regularização de Fourier [15], de decomposição de imagem [2], de viabilidade [9], de projeção alternada [22], de distância quadrada mínima [7], entre outros.

Nosso estudo baseia-se na análise convexa, em particular na noção de operador proximal, que foi introduzida por Moreau em [18] e posteriormente investigada em [19, 20], como uma generalização da noção de um operador de projeção em conjuntos convexos. Utilizamos como principais referências o artigo de Combettes e Wajs [11] e o livro de Bauschke e Combettes [3]. A organização deste trabalho encontra-se da seguinte maneira:

No primeiro capítulo, relembramos conceitos básicos da análise convexa, definimos convolução infimal, conjugada de Fenchel e subdiferencial, e fornecemos um teorema de existência e unicidade de minimizadores para certas funções convexas. No capítulo 2, estudamos operadores não-expansivos e sequências Fejér-monótonas a fim de obter um resultado de convergência para interações de ponto fixo envolvendo operadores não-expansivos ponderados. No capítulo 3, introduzimos operadores proximais, onde discutimos suas principais propriedades e descrevemos fórmulas fechadas para calcular tais operadores juntamente com exemplos específicos. Também apresentamos o princípio de decomposição de Moreau que diz respeito a decomposição de um ponto em termos do operador proximal. No último capítulo, estudamos as propriedades do Problema (P) e analisamos a convergência de um algoritmo forward-backward para resolvê-lo. Tal algoritmo tem como princípio usar as funções f_1 e f_2 separadamente, mais especificamente, o núcleo de uma interação consiste em um passo de gradiente forward (explícito) em f_2 , seguido de um passo backward (implícito) em f_1 . Por fim, aplicamos os resultados obtidos em problemas de recuperação de sinais.

Notações:

$\mathcal{H}, \mathcal{K}, \mathcal{Z}$	espaços de Hilbert
$B(x, \rho)$	bola fechada de centro x e raio ρ
\overline{C}	fecho de um conjunto C
\rightharpoonup	convergência fraca
\rightarrow	convergência forte
\mathbb{R}_+	intervalo $[0, +\infty)$
\mathbb{R}_{++}	intervalo $(0, +\infty)$
$f(\mathcal{H})$	conjunto imagem da função $f : \mathcal{H} \rightarrow (-\infty, +\infty]$
I	operador identidade
L^{-1}	inverso de L
L^*	adjunto de L
$2^{\mathcal{H}}$	conjunto das partes de \mathcal{H}
$C + D$	soma de Minkowski dos conjuntos C e D
αC	scaling do conjunto C por um número real α

1 CONCEITOS BÁSICOS DE ANÁLISE CONVEXA

Neste capítulo, descrevemos alguns conceitos básicos de análise convexa envolvendo funções convexas definidas em espaços de Hilbert, que serão utilizados ao longo deste trabalho. Detalhes e demonstrações podem ser encontrados em [3]. Assumimos como conhecidos resultados de análise funcional (ver [16]).

Seja \mathcal{H} um espaço de Hilbert real com produto escalar $\langle \cdot | \cdot \rangle$. Denotamos por $\| \cdot \|$ a norma e por d a distância associadas, respectivamente.

Seja C um subconjunto de \mathcal{H} . Se C é não-vazio, a distância de um ponto $x \in \mathcal{H}$ a C é definida por

$$d_C(x) = \inf_{y \in C} \|x - y\|. \quad (1.1)$$

Se C é fechado e convexo, então para todo $x \in \mathcal{H}$ existe um único ponto $P_C x \in C$ tal que

$$\|x - P_C x\| = d_C(x). \quad (1.2)$$

O ponto $P_C x$ é a projeção de x em C e é caracterizado pelas relações

$$P_C x \in C \quad \text{e} \quad (\forall y \in C) \langle y - P_C x | x - P_C x \rangle \leq 0. \quad (1.3)$$

Sejam $x, y \in \mathcal{H}$ e $\alpha \in \mathbb{R}$. A identidade

$$\|\alpha x + (1 - \alpha)y\|^2 + \alpha(1 - \alpha)\|x - y\|^2 = \alpha\|x\|^2 + (1 - \alpha)\|y\|^2 \quad (1.4)$$

é bastante utilizada quando trabalhamos com resultados que envolvem a norma ao quadrado.

1.1 FUNÇÕES CONVEXAS

Uma função $f : \mathcal{H} \rightarrow [-\infty, +\infty]$ é *convexa* se para todo $x, y \in \mathcal{H}$ e $\lambda \in (0, 1)$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad (1.5)$$

desde que o lado direito de (1.5) esteja bem definido. Isto é equivalente a dizer que o *epigrafo* de f , definido por

$$\text{epi} f = \{(x, \eta) \in \mathcal{H} \times \mathbb{R} \mid f(x) \leq \eta\}, \quad (1.6)$$

é um subconjunto convexo de $\mathcal{H} \times \mathbb{R}$. Se f é convexa, então seu *domínio efetivo* ou, simplesmente, *domínio*, definido por

$$\text{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}, \quad (1.7)$$

é convexo.

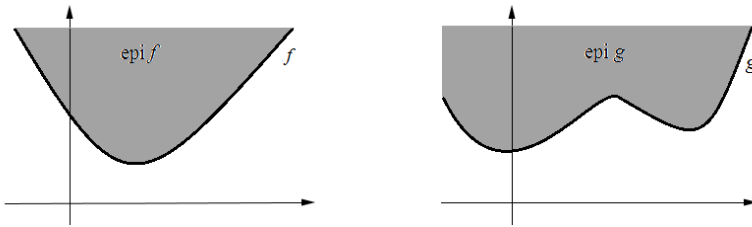


Figura 1: Representação de uma função convexa f e de uma função não convexa g , com seus respectivos epigrafos.

Se a desigualdade (1.5) é estrita para x diferente de y , então a função f é dita *estritamente convexa*.

Exemplo 1.1.1. A função $f : \mathcal{H} \rightarrow \mathbb{R} : x \mapsto \|x\|^2$ é *estritamente convexa*, uma vez que se considerarmos x diferente de y e utilizarmos (1.4),

$$\begin{aligned} \|\lambda x + (1 - \lambda)y\|^2 &< \|\lambda x + (1 - \lambda)y\|^2 + \lambda(1 - \lambda)\|x - y\|^2 \\ &= \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 \end{aligned}$$

para todo $x, y \in \mathcal{H}$ e $\lambda \in (0, 1)$.

Uma função f é *própria* se $f > -\infty$ e $\text{dom} f \neq \emptyset$, em outras palavras, f é própria se $f > -\infty$ e não é identicamente $+\infty$. Além disso, a função f é *semicontínua inferiormente* (s.c.i.) em $x_0 \in \mathcal{H}$ se ela é majorada pelo seu limite inferior¹, isto é,

$$f(x_0) \leq \liminf_{x \rightarrow x_0} f(x). \quad (1.8)$$

¹O limite inferior de f em x_0 é definido por $\liminf_{x \rightarrow x_0} f(x) = \sup_{\delta \in \delta_0} \inf_{x \in \delta} f(x)$, onde δ_0 é o conjunto de vizinhanças de x_0 .

Dizemos que f é s.c.i. em \mathcal{H} se f é s.c.i. em todos os pontos de \mathcal{H} , o que é equivalente a dizer que o epigrafo de f é um subconjunto fechado de $\mathcal{H} \times \mathbb{R}$.

Uma função f convexa e própria é s.c.i. em \mathcal{H} se, e somente se, é s.c.i. com respeito à topologia fraca em \mathcal{H} , pois um conjunto convexo é fechado se, e somente se, é fechado fraco. Em particular, o epigrafo da f é fechado se, e somente se, é fechado fraco.

Com respeito à continuidade das funções convexas convém mencionar o seguinte.

Proposição 1.1.1. [3, Teorema 8.29] *Sejam $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ própria e convexa e $x_0 \in \text{dom}f$. Então, são equivalentes:*

(i) f é contínua em x_0 .

(ii) f é limitada superiormente em uma vizinhança de x_0 .

A classe de todas as funções de \mathcal{H} em $[-\infty, +\infty]$ convexas, s.c.i. e próprias é denotada por $\Gamma_0(\mathcal{H})$. Esta classe de funções desempenha um papel central neste trabalho, tendo em vista que é sobre $\Gamma_0(\mathcal{H})$ que todos os resultados são desenvolvidos. Convém observar que, se f e g são funções em $\Gamma_0(\mathcal{H})$ tal que a interseção de seus domínios é não-vazia, então a função $f + g$ está em $\Gamma_0(\mathcal{H})$.

Uma propriedade chave das funções em $\Gamma_0(\mathcal{H})$ por vezes citada ao longo deste trabalho é a seguinte.

Lema 1.1.1. [4, Proposição 1.10] *Se $f \in \Gamma_0(\mathcal{H})$, então f é limitada inferiormente por uma função afim.*

Como um importante exemplo de função convexa temos a função indicadora de um conjunto convexo. Dado um subconjunto A (não-vazio) de \mathcal{H} , a função definida por

$$\iota_A : \mathcal{H} \rightarrow [-\infty, +\infty] : x \mapsto \begin{cases} 0, & \text{se } x \in A \\ +\infty, & \text{se } x \notin A \end{cases} \quad (1.9)$$

é chamada de *função indicadora* de A . Observemos que a função indicadora de A é própria e seu domínio é o conjunto A ; é convexa se, e somente se, A é convexo; é s.c.i. se, e somente se, A é fechado, pois seu epigrafo é o conjunto $A \times \mathbb{R}$, que é fechado se A é fechado. Portanto, a função ι_A pertence à classe $\Gamma_0(\mathcal{H})$ se, e somente se, A é um subconjunto não-vazio, fechado e convexo de \mathcal{H} .

1.2 OTIMIZAÇÃO CONVEXA

A otimização convexa é uma das principais áreas de aplicação da análise convexa. Nesta seção, trazemos uma breve introdução às funções coercivas, a fim de apresentar um teorema de existência e unicidade para problemas de minimização.

Lembremos que, dados uma função própria $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ e x em \mathcal{H} , o ponto x é um minimizador de f se

$$f(x) = \inf f(\mathcal{H}),$$

isto é, $f(x) = \min f(\mathcal{H}) \in \mathbb{R}$. O conjunto dos minimizadores de f é denotado por $\arg \min f$.

Seja $f : \mathcal{H} \rightarrow (-\infty, +\infty]$. Então, f é *coerciva* se

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty,$$

e *supercoerciva* se

$$\lim_{\|x\| \rightarrow +\infty} \frac{f(x)}{\|x\|} = +\infty.$$

Note que, uma função supercoerciva é coerciva, visto que se f é supercoerciva e x está \mathcal{H} , temos

$$\lim_{\|x\| \rightarrow +\infty} f(x) = \lim_{\|x\| \rightarrow +\infty} \frac{f(x)}{\|x\|} \cdot \|x\| = +\infty.$$

Uma função coerciva pode não ser supercoerciva, como é o caso da função $f = \|\cdot\|$.

Exemplo 1.2.1. A função $f : \mathcal{H} \rightarrow \mathbb{R} : x \mapsto \|x\|^2$ é supercoerciva, uma vez que

$$\lim_{\|x\| \rightarrow +\infty} \frac{\|x\|^2}{\|x\|} = \lim_{\|x\| \rightarrow +\infty} \|x\| = +\infty.$$

A soma de duas funções coercivas é coerciva, pela propriedade do limite da soma. Se uma das funções não for coerciva, obtemos pela seguinte proposição uma condição para que a soma seja.

Proposição 1.2.1. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $g : \mathcal{H} \rightarrow (-\infty, +\infty]$ supercoerciva. Então, $f + g$ é supercoerciva.*

Demonstração. Como f está em $\Gamma_0(\mathcal{H})$, pelo Lema 1.1.1 ela é limitada inferiormente por uma função afim, digamos $\langle \cdot | u \rangle + \eta$, onde $u \in \mathcal{H}$ e

$\eta \in \mathbb{R}$. Assim, usando este fato e a desigualdade de Cauchy-Schwarz,

$$(\forall x \in \mathcal{H}) \quad f(x) + g(x) \geq \langle x|u \rangle + \eta + g(x) \geq -\|x\|\|u\| + \eta + g(x).$$

Logo,

$$\frac{f(x) + g(x)}{\|x\|} \geq -\|u\| + \frac{\eta + g(x)}{\|x\|} \rightarrow +\infty$$

quando $\|x\| \rightarrow +\infty$, uma vez que g é supercoerciva. Portanto, $f + g$ é supercoerciva. ■

O próximo teorema nos dá condições que garantem a existência e unicidade de minimizadores para funções em $\Gamma_0(\mathcal{H})$.

Teorema 1.2.1. *Seja $f \in \Gamma_0(\mathcal{H})$ coerciva. Então, f admite um minimizador em \mathcal{H} . Além disso, se f é estritamente convexa, então o minimizador é único.*

Demonstração. Defina $a = \inf_{x \in \mathcal{H}} f(x)$. Seja (x_n) uma seqüência em \mathcal{H} tal que $f(x_n) \rightarrow a$. A seqüência (x_n) é limitada, pois caso contrário, existiria uma subsequência (x_{n_j}) tal que $0 \neq \|x_{n_j}\| \rightarrow +\infty$. Como f é coerciva, por hipótese, teríamos $f(x_{n_j}) \rightarrow +\infty$, o que é absurdo, pois $f(x_n) \rightarrow a$. Assim, como \mathcal{H} é um espaço de Hilbert e (x_n) é uma seqüência limitada, existe uma subsequência (x_{n_k}) que converge fracamente para \bar{x} em \mathcal{H} . E, como f é s.c.i., temos que

$$(\forall x \in \mathcal{H}) \quad f(\bar{x}) \leq \liminf f(x_{n_k}) = a \leq f(x).$$

Logo, f admite um minimizador \bar{x} em \mathcal{H} .

Suponha que f é estritamente convexa e que \hat{x} é um minimizador da f em \mathcal{H} . Se \hat{x} é diferente de \bar{x} , então

$$\begin{aligned} (\forall \lambda \in (0, 1)) \quad f(\hat{x}) &\leq f(\lambda \hat{x} + (1 - \lambda)\bar{x}) \\ &< \lambda f(\hat{x}) + (1 - \lambda)f(\bar{x}) \\ &\leq \lambda f(\hat{x}) + (1 - \lambda)f(\hat{x}) \\ &= f(\hat{x}), \end{aligned}$$

ou seja, $f(\hat{x}) < f(\hat{x})$, que é absurdo. Portanto, \hat{x} é igual a \bar{x} e o minimizador da f é único. ■

Corolário 1.2.1. *Sejam $f, g \in \Gamma_0(\mathcal{H})$ tal que $\text{dom}f \cap \text{dom}g \neq \emptyset$ e f é supercoerciva. Então, $f + g$ é coerciva e tem um minimizador em \mathcal{H} . Se f ou g é estritamente convexa, então $f + g$ tem exatamente um minimizador em \mathcal{H} .*

Demonstração. Como já observamos a função $f + g$ está em $\Gamma_0(\mathcal{H})$, e é coerciva, pela Proposição 1.2.1. Portanto, o resultado segue do Teorema 1.2.1. ■

1.3 CONVOLUÇÃO INFIMAL

Sejam f e g funções em $\Gamma_0(\mathcal{H})$. A *convolução infimal* de f e g é a função definida por

$$f \square g : \mathcal{H} \rightarrow [-\infty, +\infty] : x \mapsto \inf_{y \in \mathcal{H}} f(y) + g(x - y). \quad (1.10)$$

Esta função é exata em um ponto $x \in \mathcal{H}$ se existe $y \in \mathcal{H}$ tal que $(f \square g)(x) = f(y) + g(x - y)$. Dizemos que $f \square g$ é exata se é exata em todos os pontos de seu domínio. Observemos que (1.10) pode ser reescrita como

$$f \square g : \mathcal{H} \rightarrow [-\infty, +\infty] : x \mapsto \inf_{\substack{u, v \in \mathcal{H} \\ u+v=x}} f(u) + g(v).$$

Assim, podemos ver que² $\text{dom}(f \square g) = \text{dom} f + \text{dom} g$.

Se as funções f e g são convexas, então $f \square g$ é convexa. No entanto, a convolução infimal de duas funções em $\Gamma_0(\mathcal{H})$ não é necessariamente exata ou s.c.i.. Por exemplo, se

$$f : \mathbb{R} \rightarrow (-\infty, +\infty] : x \mapsto \begin{cases} 1/x, & \text{se } x > 0 \\ +\infty, & \text{se } x \leq 0 \end{cases}$$

e $g : x \mapsto f(-x)$, então

- (i) $f \in \Gamma_0(\mathbb{R})$ e $g \in \Gamma_0(\mathbb{R})$.
- (ii) $f \square g = 0$ e $f \square g$ não é exata.
- (iii) As funções ι_C e ι_D , onde $C = \text{epi} f$ e $D = \text{epi} g$, estão em $\Gamma_0(\mathbb{R}^2)$, pois segue de (i) que C e D são subconjuntos não-vazios, fechados e convexas de \mathbb{R}^2 . Porém, o conjunto $C + D$ é o semi-plano superior aberto de \mathbb{R}^2 e portanto $\iota_C \square \iota_D = \iota_{C+D}$ não é s.c.i..

²Sejam C e D subconjuntos de \mathcal{H} . A soma de Minkowski dos conjuntos C e D é definida por $C + D = \{c + d \mid c \in C, d \in D\}$.

Apresentamos abaixo uma condição para que a convolução infimal de duas funções em $\Gamma_0(\mathcal{H})$ seja exata e esteja em $\Gamma_0(\mathcal{H})$.

Proposição 1.3.1. *Sejam $f, g \in \Gamma_0(\mathcal{H})$ e f supercoerciva. Então, $f \square g$ é exata e está em $\Gamma_0(\mathcal{H})$.*

Demonstração. Primeiramente, vamos mostrar que $f \square g$ é exata. Observe que

$$\text{dom} f \square g = \text{dom} f + \text{dom} g \neq \emptyset, \quad (1.11)$$

pois o domínio da f e o domínio da g são não-vazios por hipótese. Então, tome $x \in \text{dom} f \square g$. Temos que $x = y + z$, onde $y \in \text{dom} f$ e $z \in \text{dom} g$. Segue que,

$$g(x - y) = g(x - (x - z)) = g(z) < +\infty,$$

ou seja, $y \in \text{dom} g(x - \cdot)$. Logo, $y \in \text{dom} f \cap \text{dom} g(x - \cdot)$. Portanto, a função $f + g(x - \cdot)$ está em $\Gamma_0(\mathcal{H})$, pois é soma de duas funções convexas, s.c.i. e próprias tal que $\text{dom} f \cap \text{dom} g(x - \cdot) \neq \emptyset$.

Temos pela Proposição 1.2.1 que $f + g(x - \cdot)$ é coerciva, já que f é supercoerciva. Assim, pelo Teorema 1.2.1, a função $f + g(x - \cdot)$ tem um minimizador em \mathcal{H} . Logo, para todo $x \in \text{dom} f \square g$,

$$(f \square g)(x) = \min_{y \in \mathcal{H}} f(y) + g(x - y) \in \mathbb{R}. \quad (1.12)$$

Isto implica que $f \square g$ é exata em todos os pontos de seu domínio.

Agora, vamos verificar que $f \square g$ está em $\Gamma_0(\mathcal{H})$. A convolução infimal de duas funções convexas é convexa, conforme já mencionado, e segue de (1.11) e (1.12) que $f \square g$ é própria.

Para concluir, falta mostrar que $f \square g$ é s.c.i., mas isto é equivalente a mostrar que $f \square g$ é sequencialmente s.c.i. Para isto, sejam $x \in \mathcal{H}$ e (x_n) uma sequência em \mathcal{H} tal que $x_n \rightarrow x$. Devemos verificar que

$$(f \square g)(x) \leq \liminf_{n \rightarrow +\infty} (f \square g)(x_n).$$

Suponha sem perda de generalidade que $(f \square g)(x_n) \rightarrow \mu \in \mathbb{R}$. Seja (y_n) uma sequência em \mathcal{H} tal que

$$(\forall n \in \mathbb{N}) \quad (f \square g)(x_n) = f(y_n) + g(x_n - y_n).$$

Note que (y_n) existe, pois $f \square g$ é exata. A sequência (y_n) é limitada. Para verificar esta afirmação, suponha que (y_n) não é limitada. Então,

existe uma subsequência (y_{n_k}) tal que $0 \neq \|y_{n_k}\| \rightarrow +\infty$. Pelo Lema 1.1.1, como $g \in \Gamma_0(\mathcal{H})$, temos que g é limitada inferiormente por uma função afim, digamos $\langle \cdot | u \rangle + \eta$, onde $u \in \mathcal{H}$ e $\eta \in \mathbb{R}$. Usando a desigualdade de Cauchy-Schwarz, a supercoercividade de f e o fato que $\langle x_{n_k} | u \rangle \rightarrow \langle x | u \rangle$, segue que

$$\begin{aligned}
 (f \square g)(x_{n_k}) &= f(y_{n_k}) + g(x_{n_k} - y_{n_k}) \\
 &\geq f(y_{n_k}) + \langle x_{n_k} - y_{n_k} | u \rangle + \eta \\
 &= f(y_{n_k}) + \langle x_{n_k} | u \rangle - \langle y_{n_k} | u \rangle + \eta \\
 &\geq f(y_{n_k}) - \|y_{n_k}\| \|u\| + \langle x_{n_k} | u \rangle + \eta \\
 &= \|y_{n_k}\| \left(\frac{f(y_{n_k})}{\|y_{n_k}\|} - \|u\| \right) + \langle x_{n_k} | u \rangle + \eta \\
 &\rightarrow +\infty
 \end{aligned}$$

quando $\|y_{n_k}\| \rightarrow +\infty$, contradizendo a hipótese que $(f \square g)(x_n) \rightarrow \mu \in \mathbb{R}$. Logo, a sequência (y_n) é limitada. Então, existe uma subsequência (y_{n_j}) tal que $y_{n_j} \rightharpoonup y \in \mathcal{H}$. Assim, temos que $x_{n_j} - y_{n_j} \rightharpoonup x - y$. Com isto e o fato de f e g serem s.c.i, quando $j \rightarrow +\infty$ obtemos que

$$\begin{aligned}
 \mu &= \lim (f \square g)(x_{n_j}) \\
 &= \lim f(y_{n_j}) + g(x_{n_j} - y_{n_j}) \\
 &\geq \liminf f(y_{n_j}) + \liminf g(x_{n_j} - y_{n_j}) \\
 &\geq f(y) + g(x - y) \\
 &\geq (f \square g)(x),
 \end{aligned}$$

ou seja, $f \square g$ é s.c.i. na topologia fraca, mas isto é equivalente a ser s.c.i. na topologia forte, pois $f \square g$ é convexa.

Portanto, $f \square g$ é exata e está em $\Gamma_0(\mathcal{H})$. ■

Considerando C um subconjunto não-vazio, fechado e convexo de \mathcal{H} , segue diretamente da proposição acima que a função distância ao conjunto C definida por $d_C : \mathcal{H} \rightarrow [0, +\infty) : x \mapsto d_C(x)$ pertence a classe $\Gamma_0(\mathcal{H})$ e é exata, sendo que é a raiz quadrada da função $d_C^2 = \iota_C \square \|\cdot\|^2$.

1.4 CONJUGADA DE FENCHEL

Na análise clássica, as transformações funcionais tornam possível investigar problemas a partir de uma perspectiva diferente e, por ve-

zes, simplificar a sua análise. Em análise convexa, a transformação que melhor desempenha este papel é a conjugada de Fenchel, que leva uma função à sua conjugada de Fenchel.

A *conjugada de Fenchel* de uma função $f \in \Gamma_0(\mathcal{H})$ é a função $f^* \in \Gamma_0(\mathcal{H})$ definida por

$$f^* : \mathcal{H} \rightarrow [-\infty, +\infty] : u \mapsto \sup_{x \in \mathcal{H}} \langle x|u \rangle - f(x). \quad (1.13)$$

Observe que f^* está bem definida, uma vez que f é própria. A *biconjugada* de f é definida por $f^{**} = (f^*)^*$. Como \mathcal{H} é um espaço de Hilbert, temos que $f^{**} = f$ (ver [4, Teorema 1.11]).

Por exemplo, a conjugada de Fenchel da função indicadora de um subconjunto C não-vazio, fechado e convexo de \mathcal{H} é a *função suporte* de C , isto é,

$$\iota_C^* = \sigma_C : u \mapsto \sup_{x \in C} \langle x|u \rangle. \quad (1.14)$$

Consequentemente, a conjugada de Fenchel da função suporte é a função indicadora,

$$\sigma_C^* = \iota_C^{**} = \iota_C. \quad (1.15)$$

Exemplo 1.4.1. Se $f = \|\cdot\|^2/2$, então $f^* = f$. De fato,

$$\begin{aligned} f^*(u) &= \sup_{x \in \mathcal{H}} \langle x|u \rangle - \frac{\|x\|^2}{2} \\ &= \sup_{r>0} \left(\sup_{\|x\|=r} \langle x|u \rangle \right) - \frac{r^2}{2} \\ &= \sup_{r>0} \left(\sup_{\|\frac{x}{r}\|=1} \langle r^{-1}x|ru \rangle \right) - \frac{r^2}{2} \\ &= \sup_{r>0} r\|u\| - \frac{r^2}{2}. \end{aligned} \quad (1.16)$$

Calculando a primeira e segunda derivada da função $r \mapsto r\|u\| - r^2/2$, verificamos que ela atinge o valor máximo em $r = \|u\|$. Substituindo em (1.16), concluímos que $f^*(u) = \|u\|^2/2$ para todo $u \in \mathcal{H}$, isto é, $f^* = \|\cdot\|^2/2$.

Uma consequência imediata de (1.13) é a desigualdade de Fenchel-Young:

$$(\forall x, u \in \mathcal{H}) \quad f(x) + f^*(u) \geq \langle x|u \rangle, \quad (1.17)$$

utilizada para fornecer a caracterização (1.19), na Seção 1.5, para o subdiferencial de uma função.

Considere agora o problema de minimizar a soma de duas funções f e g em $\Gamma_0(\mathcal{H})$. Temos que o *problema primal* associado é

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(x)$$

e seu *problema dual* é

$$\underset{u \in \mathcal{H}}{\text{minimize}} \quad f^*(u) + g^*(-u).$$

O principal resultado que envolve o problema primal e seu problema dual é o seguinte.

Teorema 1.4.1 (Dualidade de Fenchel-Rockafellar). [4, Teorema 1.12] *Sejam $f, g \in \Gamma_0(\mathcal{H})$. Se existe $x_0 \in \text{dom} f \cap \text{dom} g$ tal que g é contínua em x_0 , então*

$$\inf_{x \in \mathcal{H}} f(x) + g(x) = \max_{u \in \mathcal{H}} -f^*(u) - g^*(-u).$$

1.5 SUBDIFERENCIAL

Uma ferramenta fundamental na análise de funções convexas não-diferenciáveis é o subdiferencial, sendo que no contexto de funções convexas é uma generalização da derivada.

Dada uma função $f \in \Gamma_0(\mathcal{H})$, o *subdiferencial* da f é o operador ponto conjunto $\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ definido por

$$\partial f(x) = \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x|u \rangle + f(x) \leq f(y)\} \quad (1.18)$$

ou, equivalentemente, usando a definição (1.13) e a desigualdade (1.17),

$$\partial f(x) = \{u \in \mathcal{H} \mid f(x) + f^*(u) = \langle x|u \rangle\}. \quad (1.19)$$

Seja $x \in \mathcal{H}$. Então, f é *subdiferenciável* em x se $\partial f(x) \neq \emptyset$. Os elementos de $\partial f(x)$ são os *subgradientes* da f em x . Por definição, se $x \notin \text{dom} f$, temos que $\partial f(x) = \emptyset$. Além disso, se f é uma função diferenciável em x com gradiente $\nabla f(x)$, então o único subgradiente de f em x é $\nabla f(x)$. Esta relação é descrita no seguinte lema.

Lema 1.5.1. [13, Proposição 5.3] *Seja $f : \mathcal{H} \rightarrow [-\infty, +\infty]$ uma função convexa. Se f é diferenciável em $x \in \mathcal{H}$, então f é subdiferenciável em x e*

$$\partial f(x) = \{\nabla f(x)\}.$$

Um exemplo relevante para este trabalho é o caso da função $f : y \mapsto \|x - y\|^2 / (2\gamma)$, diferenciável em \mathcal{H} com $\nabla f(y) = (y - x) / \gamma$. Pelo Lema 1.5.1, temos que f é subdiferenciável em \mathcal{H} e

$$\partial \left(\frac{1}{2\gamma} \|x - \cdot\|^2 \right) (y) = \left\{ \frac{y - x}{\gamma} \right\}.$$

Utilizando a noção de subdiferencial podemos caracterizar os minimizadores da função f conforme segue

$$\arg \min f = \{x \in \mathcal{H} \mid 0 \in \partial f(x)\}. \quad (1.20)$$

Por exemplo, a função $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$ é subdiferenciável em todo $x \in \mathbb{R}$ e seu subdiferencial é dado por

$$\partial f(x) = \begin{cases} 1, & \text{se } x > 0 \\ [-1, 1], & \text{se } x = 0 \\ -1, & \text{se } x < 0. \end{cases}$$

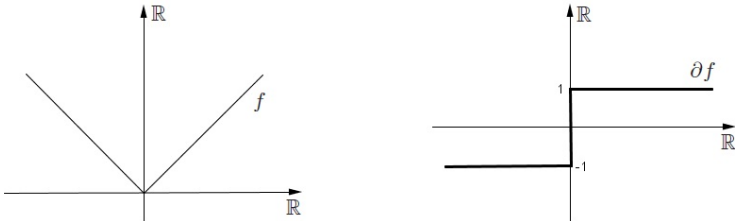


Figura 2: Gráfico da função $f : x \mapsto |x|$ e do seu subdiferencial.

Em geral, a função $f : \mathcal{H} \rightarrow \mathbb{R} : x \mapsto \|x\|$ é subdiferenciável em todo

$x \in \mathcal{H}$ e

$$\partial f(x) = \begin{cases} x/\|x\|, & \text{se } x \neq 0 \\ B(0,1), & \text{se } x = 0. \end{cases}$$

Em ambos os casos, o minimizador da função f é o ponto $x = 0$.

Apresentamos a seguir três regras de cálculo para subdiferenciais de funções em $\Gamma_0(\mathcal{H})$. A primeira proposição, refere-se ao cálculo de subdiferenciais no espaço produto.

Proposição 1.5.1. [3, Proposição 16.8] *Seja $I = \{1, \dots, m\}$ um conjunto de índices. Seja $(\mathcal{H}_i)_{i \in I}$ uma família de espaços de Hilbert reais. Considere $\mathcal{H}^m = \mathcal{H}_1 \times \dots \times \mathcal{H}_m$. Para todo $i \in I$, sejam $f_i \in \Gamma_0(\mathcal{H}_i)$ e $x_i \in \mathcal{H}_i$. Defina $x = (x_i)_{i \in I}$ e $f : \mathcal{H}^m \rightarrow (-\infty, +\infty]$ por $f(x) = \sum_{i \in I} f_i(x_i)$. Então,*

$$\partial f(x) = \partial f_1(x_1) \times \dots \times \partial f_m(x_m).$$

A proposição seguinte estabelece uma fórmula para calcular o subdiferencial de transformações de funções convexas em termos do subdiferencial dessas funções.

Proposição 1.5.2. [3, Teorema 16.37] *Considere \mathcal{K} um espaço de Hilbert real. Sejam $f = g \circ L$, onde $g \in \Gamma_0(\mathcal{K})$ e $L : \mathcal{H} \rightarrow \mathcal{K}$ é um operador linear, limitado e bijetivo tal que $L^{-1} = L^*$. Então, $\partial f = L^* \circ (\partial g) \circ L$.*

Por fim, pela próxima proposição obtemos condições e a fórmula para o cálculo do subdiferencial da soma de duas funções.

Proposição 1.5.3. [3, Corolário 16.38] *Sejam $f, g \in \Gamma_0(\mathcal{H})$. Se existe $x_0 \in \text{dom} f \cap \text{dom} g$ tal que f é contínua em x_0 , então $\partial(f+g) = \partial f + \partial g$.*

Considere C um subconjunto não-vazio, fechado e convexo de \mathcal{H} , introduzimos dois importantes exemplos de subdiferenciais. Um é o caso em que f é a função indicadora do conjunto C . Por definição, $u \in \partial \iota_C$ se, e somente se, para todo $y \in \mathcal{H}$,

$$\langle y - x | u \rangle + \iota_C(x) \leq \iota_C(y).$$

Se $x \in C$, esta condição diz que, para todo $y \in \mathcal{H}$,

$$\langle y - x | u \rangle \leq 0,$$

isto é, que u é normal a C em x . Logo, o subdiferencial da função indicadora de C em x é o cone normal de C em x , se $x \in C$, e é vazio, se $x \notin C$.

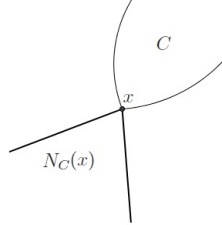


Figura 3: Cone normal de C em x .

Portanto, o operador cone normal de C é definido por

$$N_C = \partial \iota_C : x \mapsto \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \langle y - x, u \rangle \leq 0\}, & \text{se } x \in C \\ \emptyset, & \text{se } x \notin C. \end{cases} \quad (1.21)$$

O outro exemplo é o caso onde f é a função distância a C , o subdiferencial da função distância (ver [3, Exemplo 16.49]) é dado por

$$(\forall x \in \mathcal{H}) \quad \partial d_C = \begin{cases} N_C(x) \cap B(0, 1), & \text{se } x \in C \\ \left\{ \frac{x - P_C x}{d_C(x)} \right\}, & \text{se } x \notin C. \end{cases} \quad (1.22)$$

2 APROXIMAÇÃO DE PONTO FIXO DE OPERADORES NÃO-EXPANSIVOS

Neste capítulo, introduzimos a teoria básica sobre sequências Fejér-monótonas e a aplicamos para obter de uma forma sistemática resultados de convergência para iterações de ponto fixo envolvendo operadores não-expansivos. Denotamos por $\text{Fix } T$ o conjunto dos pontos fixos de um operador $T : \mathcal{H} \rightarrow \mathcal{H}$, isto é,

$$\text{Fix } T = \{x \in \mathcal{H} \mid Tx = x\}.$$

2.1 OPERADORES NÃO-EXPANSIVOS

Operadores não-expansivos são operadores Lipschitz contínuos com constante de Lipschitz um. Nesta seção, apresentamos algumas caracterizações para estes operadores.

Definição 2.1.1. *Um operador $T : \mathcal{H} \rightarrow \mathcal{H}$ é firmemente não-expansivo se*

$$(\forall x, y \in \mathcal{H}) \quad \|Tx - Ty\|^2 + \|(I - T)x - (I - T)y\|^2 \leq \|x - y\|^2 \quad (2.1)$$

e não-expansivo se é Lipschitz contínuo com constante um, isto é,

$$(\forall x, y \in \mathcal{H}) \quad \|Tx - Ty\| \leq \|x - y\|. \quad (2.2)$$

Segue diretamente da definição acima que um operador firmemente não-expansivo é não-expansivo. Porém, a recíproca não é verdadeira. Como exemplo, considere o operador $T = -I$, onde I denota o operador identidade em \mathcal{H} . Temos que T é não-expansivo, mas não é firmemente não-expansivo.

Proposição 2.1.1. *Seja $T : \mathcal{H} \rightarrow \mathcal{H}$. Então, são equivalentes:*

- (i) T é firmemente não-expansivo.
- (ii) $I - T$ é firmemente não-expansivo.
- (iii) $2T - I$ é não-expansivo.
- (iv) $(\forall x, y \in \mathcal{H}) \quad \|Tx - Ty\|^2 \leq \langle x - y | Tx - Ty \rangle$.

Demonstração.

(i) \Leftrightarrow (ii) Segue da simetria entre T e $I - T$ em (2.1).

(i) \Leftrightarrow (iii) Sejam x e y em \mathcal{H} . Defina

$$\mu = \|Tx - Ty\|^2 + \|(I - T)x - (I - T)y\|^2 - \|x - y\|^2$$

e

$$\nu = \|(2T - I)x - (2T - I)y\|^2 - \|x - y\|^2.$$

Por (1.4),

$$\begin{aligned} \|(2T - I)x - (2T - I)y\|^2 &= \\ &= \|2(Tx - Ty) + (1 - 2)(x - y)\|^2 \\ &= 2\|Tx - Ty\|^2 - \|x - y\|^2 + 2\|(I - T)x - (I - T)y\|^2. \end{aligned}$$

Disso, vemos que $\nu = 2\mu$. Assim, $2T - I$ é não-expansivo $\Leftrightarrow \nu \leq 0 \Leftrightarrow \mu \leq 0 \Leftrightarrow T$ é firmemente não-expansivo.

(i) \Leftrightarrow (iv) Basta escrever

$$\|(I - T)x - (I - T)y\|^2 = \|x - y\|^2 + \|Tx - Ty\|^2 - 2\langle x - y | Tx - Ty \rangle$$

em (2.1). ■

Definição 2.1.2. *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ e $\beta \in \mathbb{R}_{++}$. Dizemos que T é β -cocoercivo se βT é firmemente não-expansivo, isto é,*

$$(\forall x, y \in \mathcal{H}) \quad \langle x - y | Tx - Ty \rangle \geq \beta \|Tx - Ty\|^2.$$

O lema a seguir é utilizado para demonstrar a Proposição 2.1.2.

Lema 2.1.1. [21, Lema 1.2.3] *Seja $f : \mathcal{H} \rightarrow \mathbb{R}$ convexa e diferenciável em \mathcal{H} tal que ∇f é $1/\beta$ -Lipschitz contínuo, para algum $\beta \in \mathbb{R}_{++}$. Então,*

$$(\forall x, y \in \mathcal{H}) \quad f(y) \leq f(x) + \langle \nabla f(x) | y - x \rangle + \frac{1}{2\beta} \|x - y\|^2.$$

Vimos que um operador não-expansivo pode não ser firmemente não-expansivo. Observe que o próximo resultado, com $\beta = 1$, afirma que isto é verdade para o gradiente de uma função convexa $f : \mathcal{H} \rightarrow \mathbb{R}$.

Proposição 2.1.2. *Seja $f : \mathcal{H} \rightarrow \mathbb{R}$ convexa e diferenciável em \mathcal{H} tal que ∇f é $1/\beta$ -Lipschitz contínuo, para algum $\beta \in \mathbb{R}_{++}$. Então, ∇f é β -cocoercivo.*

Demonstração. Seja $x \in \mathcal{H}$. Defina a função

$$\phi : \mathcal{H} \rightarrow \mathbb{R} : y \mapsto f(y) - \langle \nabla f(x) | y \rangle.$$

Notemos que, ϕ está bem definida e é convexa, pois é soma de funções convexas. Ainda, seu gradiente $\nabla \phi(y) = \nabla f(y) - \nabla f(x)$ é $1/\beta$ -Lipschitz contínuo. Como ϕ é convexa e $\nabla \phi(x) = 0$, temos que x é minimizador de ϕ . Então, usando o Lema 2.1.1 para a função ϕ ,

$$\begin{aligned} (\forall y \in \mathcal{H}) \quad \phi(x) &\leq \phi(y - \beta \nabla \phi(y)) \\ &\leq \phi(y) + \langle \nabla \phi(y) | -\beta \nabla \phi(y) \rangle + \frac{1}{2\beta} \|\beta \nabla \phi(y)\|^2 \\ &= \phi(y) - \frac{\beta}{2} \|\nabla \phi(y)\|^2, \end{aligned}$$

isto é,

$$f(x) - \langle \nabla f(x) | x \rangle \leq f(y) - \langle \nabla f(x) | y \rangle - \frac{\beta}{2} \|\nabla f(y) - \nabla f(x)\|^2.$$

Segue que, para todo $x, y \in \mathcal{H}$,

$$f(x) + \langle \nabla f(x) | y - x \rangle + \frac{\beta}{2} \|\nabla f(y) - \nabla f(x)\|^2 \leq f(y). \quad (2.3)$$

Trocando x por y e y por x na equação (2.3), obtemos

$$f(y) + \langle \nabla f(y) | x - y \rangle + \frac{\beta}{2} \|\nabla f(x) - \nabla f(y)\|^2 \leq f(x). \quad (2.4)$$

Somando (2.3) e (2.4), concluímos que, para todo $x, y \in \mathcal{H}$,

$$\langle \nabla f(x) - \nabla f(y) | x - y \rangle \geq \beta \|\nabla f(x) - \nabla f(y)\|^2,$$

ou seja, que ∇f é β -cocoercivo. ■

2.2 OPERADORES PONDERADOS

Uma classe especial de operadores não-expansivos são os operadores ponderados, cujas propriedades são exploradas nesta seção.

Definição 2.2.1. *Seja $\alpha \in (0, 1)$. Um operador $T : \mathcal{H} \rightarrow \mathcal{H}$ é ponderado com constante α , ou α -ponderado, se existe um operador não-expansivo $R : \mathcal{H} \rightarrow \mathcal{H}$ tal que $T = (1 - \alpha)I + \alpha R$.*

Proposição 2.2.1. *Seja $T : \mathcal{H} \rightarrow \mathcal{H}$.*

- (i) *Se T é α -ponderado, então T é não-expansivo.*
- (ii) *Se T é não-expansivo, T não é necessariamente α -ponderado.*
- (iii) *T é firmemente não-expansivo se, e somente se, T é $1/2$ -ponderado.*
- (iv) *Seja $\beta \in \mathbb{R}_{++}$. T é β -cocoercivo se, e somente se, βT é $1/2$ -ponderado.*

Demonstração.

(i) Sejam x e y em \mathcal{H} , da Definição 2.2.1 segue que

$$\begin{aligned} \|Tx - Ty\| &= \|((1 - \alpha)I + \alpha R)x - ((1 - \alpha)I + \alpha R)y\| \\ &= \|(1 - \alpha)(x - y) + \alpha(Rx - Ry)\| \\ &\leq (1 - \alpha)\|x - y\| + \alpha\|Rx - Ry\| \\ &\leq (1 - \alpha)\|x - y\| + \alpha\|x - y\| \\ &= \|x - y\|. \end{aligned}$$

(ii) Considere $T = -I$, que é não-expansivo. Sejam $\alpha \in (0, 1)$ e $x \in \mathcal{H}$, suponha que T é α -ponderado. Então, existe um operador não-expansivo $R : \mathcal{H} \rightarrow \mathcal{H}$ tal que

$$(1 - \alpha)x + \alpha Rx = -x, \text{ ou seja, } Rx = \frac{\alpha - 2}{\alpha}x.$$

Assim,

$$\|Rx - Ry\| = \frac{|\alpha - 2|}{\alpha}\|x - y\| \leq \|x - y\|.$$

Disto, temos que $|\alpha - 2|/\alpha \leq 1$, contradizendo o fato de $\alpha \in (0, 1)$. Logo, não existe R não-expansivo tal que $T = (1 - \alpha)I + \alpha R$.

(iii) Se T é firmemente não-expansivo, pela Proposição 2.1.1, $2T - I$ é não-expansivo. Defina $R = 2T - I$. Logo, existe um operador não-expansivo R tal que $T = (1 - 1/2)I + (1/2)R$. Portanto, T é $1/2$ -ponderado. Agora, se T é $1/2$ -ponderado, temos que $T = (I + R)/2$ com R não-expansivo. Sejam x e y em \mathcal{H} . Segue que,

$$\|Tx - Ty\|^2 + \|(I - T)x - (I - T)y\|^2 =$$

$$\begin{aligned}
&= \left\| \frac{x + Rx}{2} - \frac{y + Ry}{2} \right\|^2 + \left\| \frac{x - Rx}{2} - \frac{y - Ry}{2} \right\|^2 \\
&= \frac{1}{4} (\|x - y + Rx - Ry\|^2 + \|x - y - (Rx - Ry)\|^2) \\
&= \frac{1}{2} (\|x - y\|^2 + \|Rx - Ry\|^2) \\
&\leq \frac{1}{2} (\|x - y\|^2 + \|x - y\|^2) \\
&= \|x - y\|^2.
\end{aligned}$$

Logo, T é firmemente não-expansivo.

(iv) Segue da Definição 2.1.2 e do item (iii). ■

Proposição 2.2.2. *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ e $\alpha \in (0, 1)$. Então, são equivalentes:*

(i) T é α -ponderado.

(ii) $(1 - 1/\alpha)I + (1/\alpha)T$ é não-expansivo.

(iii) $(\forall x, y \in \mathcal{H}) \|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1 - \alpha}{\alpha} \|(I - T)x - (I - T)y\|^2$.

(iv) $(\forall x, y \in \mathcal{H}) \|Tx - Ty\|^2 + (1 - 2\alpha)\|x - y\|^2 \leq 2(1 - \alpha)\langle x - y | Tx - Ty \rangle$.

Demonstração. Fixe x e y em \mathcal{H} , defina $R = (1 - \lambda)I + \lambda T$, onde $\lambda = 1/\alpha$, e note que $T = (1 - \alpha)I + \alpha R$.

(i) \Leftrightarrow (ii) Por definição T é α -ponderado se, e somente se, R é não-expansivo.

(i) \Leftrightarrow (iii) Usando (1.4), temos que

$$\begin{aligned}
\|Rx - Ry\|^2 &= \|(1 - \lambda)x + \lambda Tx - (1 - \lambda)y - \lambda Ty\|^2 \\
&= \|(1 - \lambda)(x - y) + \lambda(Tx - Ty)\|^2 \\
&= (1 - \lambda)\|x - y\|^2 + \lambda\|Tx - Ty\|^2 \\
&\quad - \lambda(1 - \lambda)\|x - y - (Tx - Ty)\|^2,
\end{aligned}$$

ou seja,

$$\begin{aligned}
\alpha(\|x - y\|^2 - \|Rx - Ry\|^2) &= \\
&= \|x - y\|^2 - \|Tx - Ty\|^2 - \frac{1 - \alpha}{\alpha} \|(I - T)x - (I - T)y\|^2. \quad (2.5)
\end{aligned}$$

Assim, R é não-expansivo se, e somente se, o lado esquerdo de (2.5) é não-negativo, o que é equivalente a

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1 - \alpha}{\alpha} \|(I - T)x - (I - T)y\|^2.$$

(iii) \Leftrightarrow (iv) Basta escrever

$$\|(I - T)x - (I - T)y\|^2 = \|x - y\|^2 + \|Tx - Ty\|^2 - 2\langle x - y | Tx - Ty \rangle$$

em (iii). ■

Observação 2.2.1. *Segue da Proposição 2.2.2 que se $T : \mathcal{H} \rightarrow \mathcal{H}$ é α -ponderado com $\alpha \in (0, 1/2]$, então T é firmemente não-expansivo. Basta notar que se $\alpha \in (0, 1/2]$, então $(\alpha - 1)/\alpha \leq -1$.*

A próxima proposição afirma que operadores ponderados são fechados em relação à operação de composição. Descrevemos este resultado considerando a composição entre dois operadores; o caso geral pode ser visto em [3, Proposição 4.32].

Proposição 2.2.3. *Sejam $T_1, T_2 : \mathcal{H} \rightarrow \mathcal{H}$ e $\alpha_1, \alpha_2 \in (0, 1)$ tais que T_1 e T_2 são ponderados com constantes α_1 e α_2 , respectivamente. Defina*

$$T = T_1 \circ T_2 \quad e \quad \alpha = \frac{2}{1 + \frac{1}{\max\{\alpha_1, \alpha_2\}}}. \quad (2.6)$$

Então, T é α -ponderado.

Demonstração. Defina

$$\kappa_1 = \frac{\alpha_1}{1 - \alpha_1}, \quad \kappa_2 = \frac{\alpha_2}{1 - \alpha_2} \quad e \quad \kappa = \max\{\kappa_1, \kappa_2\}.$$

Sejam x e y em \mathcal{H} . Segue da convexidade de $x \mapsto \|x\|^2$ e da equivalência (i) \Leftrightarrow (iii) na Proposição 2.2.2 que

$$\begin{aligned} & \|(I - T)x - (I - T)y\|^2 / 2 = \\ & = \|(x - y) - (T_2x - T_2y) + (T_2x - T_2y) - (T_1T_2x - T_1T_2y)\|^2 / 2 \\ & = \|(I - T_2)x - (I - T_2)y + (I - T_1)T_2x - (I - T_1)T_2y\|^2 / 2 \\ & \leq \|(I - T_2)x - (I - T_2)y\|^2 + \|(I - T_1)T_2x - (I - T_1)T_2y\|^2 \\ & \leq \kappa_2(\|x - y\|^2 - \|T_2x - T_2y\|^2) + \kappa_1(\|T_2x - T_2y\|^2 \\ & \quad - \|T_1T_2x - T_1T_2y\|^2) \end{aligned}$$

$$\begin{aligned} &\leq \kappa(\|x - y\|^2 - \|T_2x - T_2y\|^2 + \|T_2x - T_2y\|^2 - \|T_1T_2x - T_1T_2y\|^2) \\ &= \kappa(\|x - y\|^2 - \|Tx - Ty\|^2). \end{aligned}$$

Assim,

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1}{2\kappa} \|(I - T)x - (I - T)y\|^2.$$

Observe que,

$$\frac{1}{2\kappa} = \frac{1 - \alpha}{\alpha}, \quad \text{pois} \quad \alpha = \frac{2}{2 + 1/\kappa}.$$

Pela Proposição 2.2.2 novamente, concluímos que T é α -ponderado. ■

Outra operação que preserva a ponderabilidade é a seguinte.

Proposição 2.2.4. *Sejam $\beta \in \mathbb{R}_{++}$, $T : \mathcal{H} \rightarrow \mathcal{H}$ um operador β -cocoercivo e $\gamma \in (0, 2\beta)$. Então, $I - \gamma T$ é $\gamma/2\beta$ -ponderado.*

Demonstração. Como T é β -cocoercivo, pela Proposição 2.2.1(iv), temos que βT é $1/2$ -ponderado, ou seja, existe um operador $R : \mathcal{H} \rightarrow \mathcal{H}$ não-expansivo tal que $\beta T = (I + R)/2$. Por sua vez,

$$I - \gamma T = I - \gamma \frac{I + R}{2\beta} = \left(1 - \frac{\gamma}{2\beta}\right)I + \frac{\gamma}{2\beta}(-R).$$

Além disso, $\gamma/2\beta \in (0, 1)$ e $-R$ é não-expansivo. Portanto, $I - \gamma T$ é $\gamma/2\beta$ -ponderado. ■

2.3 SEQUÊNCIAS FEJÉR-MONÓTONAS

A noção de sequências Fejér-monótonas é central no estudo de pontos fixos de operadores não-expansivos. Apresentamos nesta seção algumas propriedades destas sequências.

Definição 2.3.1. *Seja C um subconjunto não-vazio de \mathcal{H} . Uma sequência (x_n) em \mathcal{H} é Fejér-monótona com respeito a C se*

$$(\forall x \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|.$$

Exemplo 2.3.1. *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ não-expansivo tal que $\text{Fix } T \neq \emptyset$ e $x_0 \in \mathcal{H}$. Defina $x_{n+1} = Tx_n$, para todo $n \in \mathbb{N}$.*

A sequência (x_n) é Fejér-monótona com respeito a $\text{Fix } T$, uma vez que, $x \in \text{Fix } T$ e a não-expansividade de T implicam

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| = \|Tx_n - Tx\| \leq \|x_n - x\|.$$

As sequências Fejér-monótonas possuem propriedades que simplificam a análise do seu comportamento assintótico. Observemos que:

- Para todo $x \in C$, a sequência $(\|x_n - x\|)$ converge, visto que é uma sequência monótona e limitada em \mathbb{R} .
- (x_n) é uma sequência limitada, pois tomando $x \in C$ temos que

$$(\forall n \in \mathbb{N}) \quad \|x_n\| \leq \|x_n - x\| + \|x\| < M + \|x\|,$$

onde M é tal que $\|x_n - x\| < M$.

O seguinte resultado diz respeito à convergência fraca de sequências Fejér-monótonas.

Teorema 2.3.1 (Browder). *Sejam C um subconjunto não-vazio de \mathcal{H} e (x_n) uma sequência Fejér-monótona com respeito a C . A sequência (x_n) converge fracamente para um ponto em C se, e somente se, todo ponto aderente fraco de (x_n) pertence a C .*

Demonstração.

(\Rightarrow) Trivial.

(\Leftarrow) Como (x_n) é limitada, existe pelo menos um ponto aderente fraco. Sejam \bar{x} e \hat{x} pontos aderentes e $(x_{n_k}), (x_{n_j})$ subseqüências de (x_n) tais que $x_{n_k} \rightharpoonup \bar{x}$ e $x_{n_j} \rightharpoonup \hat{x}$. Por hipótese, temos que $\bar{x} \in C$ e $\hat{x} \in C$. Como $(\|x_n - \bar{x}\|)$ e $(\|x_n - \hat{x}\|)$ são convergentes, existem $\bar{\ell}$ e $\hat{\ell}$ tais que

$$\|x_n - \bar{x}\| \rightarrow \bar{\ell} \quad \text{e} \quad \|x_n - \hat{x}\| \rightarrow \hat{\ell}.$$

Em particular,

$$\|x_{n_k} - \bar{x}\| \rightarrow \bar{\ell} \quad \text{e} \quad \|x_{n_j} - \bar{x}\| \rightarrow \bar{\ell},$$

bem como,

$$\|x_{n_k} - \hat{x}\| \rightarrow \hat{\ell} \quad \text{e} \quad \|x_{n_j} - \hat{x}\| \rightarrow \hat{\ell}.$$

Disso, segue que

$$\begin{aligned}
 \|x_{n_k} - \bar{x}\|^2 &= \|x_{n_k} - \hat{x}\|^2 + \|\hat{x} - \bar{x}\|^2 + 2\langle x_{n_k} - \hat{x} | \hat{x} - \bar{x} \rangle \\
 (\bar{\ell})^2 &= (\hat{\ell})^2 + \|\hat{x} - \bar{x}\|^2 + 2\langle \bar{x} - \hat{x} | \hat{x} - \bar{x} \rangle \\
 (\bar{\ell})^2 &= (\hat{\ell})^2 - \|\hat{x} - \bar{x}\|^2,
 \end{aligned} \tag{2.7}$$

por outro lado,

$$\begin{aligned}
 \|x_{n_j} - \hat{x}\|^2 &= \|x_{n_j} - \bar{x}\|^2 + \|\bar{x} - \hat{x}\|^2 + 2\langle x_{n_j} - \bar{x} | \bar{x} - \hat{x} \rangle \\
 (\hat{\ell})^2 &= (\bar{\ell})^2 + \|\bar{x} - \hat{x}\|^2 + 2\langle \hat{x} - \bar{x} | \bar{x} - \hat{x} \rangle \\
 (\hat{\ell})^2 &= (\bar{\ell})^2 - \|\bar{x} - \hat{x}\|^2,
 \end{aligned} \tag{2.8}$$

Somando (2.7) e (2.8), concluímos que $\|\bar{x} - \hat{x}\|^2 = 0$, isto é, $\bar{x} = \hat{x}$. Logo, a sequência (x_n) tem um único ponto aderente $x \in C$, ou seja, (x_n) converge fracamente para $x \in C$. ■

No caso de dimensão finita, para que a sequência Fejér monótona com respeito a C convirja fracamente para um ponto em C , basta que apenas um ponto aderente fraco esteja em C . Mas, isto não é suficiente se a dimensão for infinita.

Consideremos $\mathcal{H} = \ell^2(\mathbb{R})$, $e_i = (0, \dots, 0, 1, 0, \dots)$, com um na i -ésima posição, e (x_n) uma sequência em \mathcal{H} definida por

$$x_n = \begin{cases} a_n e_n, & \text{se } n \text{ é ímpar} \\ a_n e_1, & \text{se } n \text{ é par,} \end{cases}$$

onde $a_n = 1 + 1/n$, para todo $n \in \mathbb{N}$. Temos que

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1}\| = |a_{n+1}| < |a_n| = \|x_n\|,$$

isto é, (x_n) é uma sequência Fejér-monótona com respeito a $\{0\}$. Além disso, zero é um ponto aderente fraco de (x_n) , pois

$$\begin{aligned}
 (\forall x \in \mathcal{H}) \quad \langle x_{2k+1} | x \rangle &= |a_{2k+1}| \langle e_{2k+1} | x \rangle \\
 &= |a_{2k+1}| x^{2k+1} \rightarrow 0,
 \end{aligned}$$

ou seja, a subsequência (x_{2k+1}) converge fraco para zero. Por outro lado, a subsequência (x_{2k}) converge fraco para um, visto que $\langle x_{2k} | e_1 \rangle = |a_{2k}| \rightarrow 1$. Logo, (x_n) não converge fraco para zero. Portanto, a sequência (x_n) é Fejér-monótona com respeito a $\{0\}$, tem zero como um ponto aderente fraco, mas não converge fracamente para zero.

2.4 ITERAÇÕES DE PONTO FIXO

Nesta seção, descrevemos condições que garantem a convergência de iterações de ponto fixo envolvendo operadores não-expansivos. O principal resultado é o Krasnosel'skiĭ-Mann para operadores ponderados, Teorema 2.4.4.

Teorema 2.4.1. *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ não-expansivo, (x_n) uma sequência em \mathcal{H} e $x \in \mathcal{H}$. Se $x_n \rightarrow x$ e $x_n - Tx_n \rightarrow 0$, então $x \in \text{Fix } T$.*

Demonstração. Obtemos, através de manipulações algébricas, que

$$\begin{aligned} \|x - Tx\|^2 &= \|x - x_n + x_n - Tx\|^2 \\ &= \|x - x_n\|^2 + \|x_n - Tx\|^2 + 2\langle x - x_n | x_n - x + x - Tx \rangle \\ &= \|x - x_n\|^2 + \|x_n - Tx_n + Tx_n - Tx\|^2 - 2\|x - x_n\|^2 \\ &\quad + 2\langle x - x_n | x - Tx \rangle \\ &= \|x_n - Tx_n\|^2 + \|Tx_n - Tx\|^2 + 2\langle x_n - Tx_n | Tx_n - Tx \rangle \\ &\quad - \|x - x_n\|^2 + 2\langle x - x_n | x - Tx \rangle. \end{aligned}$$

Como T é não-expansivo, usando (2.2), segue que

$$\begin{aligned} \|x - Tx\|^2 &\leq \|x_n - Tx_n\|^2 + \|x_n - x\|^2 + 2\langle x_n - Tx_n | Tx_n - Tx \rangle \\ &\quad - \|x - x_n\|^2 + 2\langle x - x_n | x - Tx \rangle \\ &= \|x_n - Tx_n\|^2 + 2\langle x_n - Tx_n | Tx_n - Tx \rangle \\ &\quad + 2\langle x - x_n | x - Tx \rangle. \end{aligned}$$

Observemos que:

- (a) $\|x_n - Tx_n\|^2 \rightarrow 0$, pois $x_n - Tx_n \rightarrow 0$;
- (b) $2\langle x - x_n | x - Tx \rangle \rightarrow 0$, pois $x_n \rightarrow x$;
- (c) $2\langle x_n - Tx_n | Tx_n - Tx \rangle \rightarrow 0$, pois $x_n - Tx_n \rightarrow 0$ e

$$Tx_n - Tx = (Tx_n - x_n) + (x_n - Tx) \rightarrow 0 + x - Tx = x - Tx,$$

em particular $(Tx_n - Tx)$ é limitada.

Assim, (a), (b) e (c) implicam que $\lim_{n \rightarrow +\infty} \|x - Tx\|^2 \leq 0$, ou seja, $\|x - Tx\|^2 = 0$. Logo, $x \in \text{Fix } T$. ■

Dado um operador não-expansivo T , a sequência gerada pela iteração $x_{n+1} = Tx_n$ pode não convergir para um ponto fixo de T . Um exemplo simples desta situação é $T = -I$ e $x_0 \neq 0$. Neste caso, a propriedade de regularidade assintótica $x_n - Tx_n \rightarrow 0$, não é satisfeita. Como veremos a seguir, esta propriedade é fundamental.

Teorema 2.4.2. *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ não-expansivo tal que $\text{Fix } T \neq \emptyset$ e $x_0 \in \mathcal{H}$. Defina*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

Se $x_n - Tx_n \rightarrow 0$, então a sequência (x_n) converge fracamente para um ponto em $\text{Fix } T$.

Demonstração. Vimos no Exemplo 2.3.1 que (x_n) é uma sequência Fejér-monótona com respeito a $\text{Fix } T$. Seja $x \in \mathcal{H}$ um ponto aderente fraco de (x_n) . Então, existe uma subsequência (x_{n_k}) tal que $x_{n_k} \rightharpoonup x$. Como $x_{n_k} - Tx_{n_k} \rightarrow 0$, aplicando o Teorema 2.4.1 para (x_{n_k}) , temos que $x \in \text{Fix } T$. Portanto, pelo Teorema 2.3.1, a sequência (x_n) converge fracamente para um ponto em $\text{Fix } T$. ■

O próximo resultado diz respeito a um método iterativo, conhecido como algoritmo Krasnosel'skiĭ-Mann.

Teorema 2.4.3 (Krasnosel'skiĭ-Mann). *Sejam $T : \mathcal{H} \rightarrow \mathcal{H}$ não-expansivo tal que $\text{Fix } T \neq \emptyset$, (λ_n) uma sequência tal que $0 < \underline{\lambda} \leq \lambda_n \leq \bar{\lambda} < 1$, para todo $n \in \mathbb{N}$, e $x_0 \in \mathcal{H}$. Defina*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Então, a sequência (x_n) converge fracamente para um ponto em $\text{Fix } T$.

Demonstração. A sequência (x_n) é Fejér-monótona com respeito a $\text{Fix } T$. De fato, seja $y \in \text{Fix } T$. Usando (1.4) e (2.2) com o fato que $y = Ty$ e T é não-expansivo, segue que, para todo $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - y\|^2 &= \\ &= \|x_n + \lambda_n(Tx_n - x_n) - y\|^2 \\ &= \|\lambda_n(Tx_n - y) + (1 - \lambda_n)(x_n - y)\|^2 \\ &= \lambda_n \|Tx_n - y\|^2 + (1 - \lambda_n) \|x_n - y\|^2 - \lambda_n(1 - \lambda_n) \|Tx_n - x_n\|^2 \\ &\leq \lambda_n \|x_n - y\|^2 + (1 - \lambda_n) \|x_n - y\|^2 - \lambda_n(1 - \lambda_n) \|Tx_n - x_n\|^2 \\ &= \|x_n - y\|^2 - \lambda_n(1 - \lambda_n) \|Tx_n - x_n\|^2. \end{aligned} \tag{2.9}$$

Em particular, $\|x_{n+1} - y\| \leq \|x_n - y\|$ para todo $y \in \text{Fix } T$. Por (2.9), temos que

$$\lambda_n(1 - \lambda_n) \|Tx_n - x_n\|^2 \leq \|x_n - y\|^2 - \|x_{n+1} - y\|^2.$$

Então, como $0 < \underline{\lambda} \leq \lambda_n \leq \bar{\lambda} < 1$ para todo $n \in \mathbb{N}$,

$$\underline{\lambda}(1 - \bar{\lambda})\|Tx_n - x_n\|^2 \leq \|x_n - y\|^2 - \|x_{n+1} - y\|^2. \quad (2.10)$$

Tomando o somatório em ambos os lados da desigualdade (2.10), obtemos

$$\begin{aligned} \underline{\lambda}(1 - \bar{\lambda}) \sum_{n=0}^k \|Tx_n - x_n\|^2 &\leq \sum_{n=0}^k (\|x_n - y\|^2 - \|x_{n+1} - y\|^2) \\ &= \|x_0 - y\|^2 - \|x_{k+1} - y\|^2 \\ &\leq \|x_0 - y\|^2. \end{aligned}$$

Logo,

$$\sum_{n \in \mathbb{N}} \|Tx_n - x_n\|^2 \leq \frac{1}{\underline{\lambda}(1 - \bar{\lambda})} \|x_0 - y\|^2 < +\infty.$$

Portanto, $\|x_n - Tx_n\|^2 \rightarrow 0$, isto é, $x_n - Tx_n \rightarrow 0$. Seja $x \in \mathcal{H}$ um ponto aderente fraco de (x_n) , digamos $x_{n_k} \rightharpoonup x$. Como $x_{n_k} - Tx_{n_k} \rightarrow 0$, pelo Teorema 2.4.1 temos que $x \in \text{Fix } T$. Em vista do Teorema 2.3.1, o resultado está provado. ■

Um caso especial do Teorema 2.4.3 é o caso em que T é um operador α -ponderado.

Teorema 2.4.4 (Krasnosel'skiĭ-Mann). *Sejam $\alpha \in (0, 1)$, $T : \mathcal{H} \rightarrow \mathcal{H}$ um operador α -ponderado tal que $\text{Fix } T \neq \emptyset$, (λ_n) uma seqüência tal que $0 < \underline{\lambda}/\alpha \leq \lambda_n \leq \bar{\lambda}/\alpha < 1$, para todo $n \in \mathbb{N}$, e $x_0 \in \mathcal{H}$. Defina*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Então, a seqüência (x_n) converge fracamente para um ponto em $\text{Fix } T$.

Demonstração. Notemos que

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad x_{n+1} &= x_n + \lambda_n(Tx_n - x_n) \\ &= x_n + \alpha\lambda_n \left(\frac{Tx_n - x_n}{\alpha} \right) \\ &= x_n + \alpha\lambda_n \left[\frac{1}{\alpha}Tx_n + \left(1 - \frac{1}{\alpha}\right)x_n - x_n \right] \\ &= x_n + \delta_n(Rx_n - x_n), \end{aligned}$$

onde $\delta_n = \alpha\lambda_n$ e $R = (1/\alpha)T + (1 - 1/\alpha)I$. Por hipótese, temos

$0 < \underline{\lambda} \leq \alpha\lambda_n \leq \bar{\lambda} < 1$; então $0 < \underline{\lambda} \leq \delta_n \leq \bar{\lambda} < 1$ para todo $n \in \mathbb{N}$. Como T é α -ponderado, pela Proposição 2.2.2, segue que R é um operador não-expansivo. Aplicando o Teorema 2.4.3 para R e (δ_n) , concluímos que (x_n) converge fracamente para um ponto x em $\text{Fix } R$. Mas,

$$\begin{aligned} x \in \text{Fix } R &\Leftrightarrow x = Rx \\ &\Leftrightarrow x = \frac{1}{\alpha}Tx + \left(1 - \frac{1}{\alpha}\right)x \\ &\Leftrightarrow x = Tx \\ &\Leftrightarrow x \in \text{Fix } T. \end{aligned}$$

Portanto, a sequência (x_n) converge fracamente para um ponto em $\text{Fix } T$. ■

3 OPERADOR PROXIMAL

O operador proximal de uma função convexa é uma extensão natural da noção de um operador de projeção sobre um conjunto convexo. Foi proposto por Moreau ao substituir ι_C , função indicadora de um conjunto fechado e convexo C pertencente a $\Gamma_0(\mathcal{H})$, por uma função qualquer $f \in \Gamma_0(\mathcal{H})$ no problema

$$\underset{y \in \mathcal{H}}{\text{minimize}} \quad \iota_C(y) + \frac{1}{2} \|x - y\|^2,$$

cujas soluções são as projeções de x em C .

Neste capítulo, estudamos com detalhes as principais propriedades deste operador.

3.1 REGULARIZAÇÃO DE MOREAU

Seja $f \in \Gamma_0(\mathcal{H})$. Quando f é diferenciável podemos encontrar seus minimizadores utilizando técnicas conhecidas, como o método do gradiente. Mas, como muitas funções convexas não são diferenciáveis, o problema de minimizar f é mais difícil. Neste caso, uma abordagem é aproximar a função f a uma função diferenciável e converter o problema em um de minimizar uma função diferenciável. A regularização de Moreau definida abaixo é um exemplo disso.

Definição 3.1.1. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $\gamma \in \mathbb{R}_{++}$. A regularização de Moreau de índice γ da função f é a função*

$$f_\gamma = f \square \left(\frac{1}{2\gamma} \|\cdot\|^2 \right) : \mathcal{H} \rightarrow \mathbb{R}$$

$$x \mapsto \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|x - y\|^2. \quad (3.1)$$

Veremos na Proposição 3.1.1 que f_γ está bem definida.

Exemplo 3.1.1. *Sejam C um subconjunto não-vazio de \mathcal{H} e $\gamma \in \mathbb{R}_{++}$. Se $f = \iota_C$, então*

$$f_\gamma = d_C^2 / (2\gamma).$$

Note que, se y não está em C então $f(y) = +\infty$. Assim, podemos tomar o ínfimo em (3.1) apenas sobre y em C ; neste caso $f(y) = 0$, e

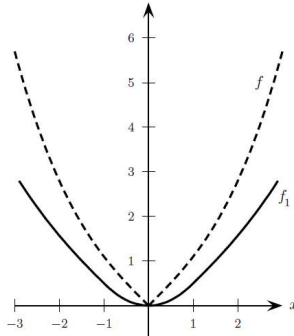


Figura 4: Gráfico de $f : x \mapsto |x| + 0.1|x|^3$ e sua regularização de Moreau f_1 .

temos que

$$(\forall x \in \mathcal{H}) \quad f_\gamma(x) = \inf_{y \in C} \frac{1}{2\gamma} \|x - y\|^2 = \frac{1}{2\gamma} \inf_{y \in C} \|x - y\|^2 = \frac{1}{2\gamma} d_C^2(x).$$

A regularização de Moreau é considerada uma função muito “boa”, pela regularidade que apresenta. Na seguinte proposição, examinamos suas propriedades.

Proposição 3.1.1. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $\gamma \in \mathbb{R}_{++}$. Então, f_γ é convexa, exata, contínua e $\text{dom} f_\gamma = \mathcal{H}$. Além disso, para todo $x \in \mathcal{H}$, o ínfimo em (3.1) é unicamente atingido.*

Demonstração. Por hipótese f está em $\Gamma_0(\mathcal{H})$. Defina $g = \|\cdot\|^2/(2\gamma)$. Temos que: g é s.c.i, pois é contínua; g é própria, pois $g > -\infty$ e $\text{dom} g = \mathcal{H}$; g é estritamente convexa, pelo Exemplo 1.1.1. Logo, g está em $\Gamma_0(\mathcal{H})$. Ainda, pelo Exemplo 1.2.1, temos que g é supercoerciva. Assim, segue da Proposição 1.3.1 que $f_\gamma = f \square g$ é exata e está em $\Gamma_0(\mathcal{H})$.

Vamos mostrar agora que f_γ é contínua em \mathcal{H} . Observe que

$$\text{dom} f_\gamma = \text{dom} f + \text{dom} g = \text{dom} f + \mathcal{H} = \mathcal{H},$$

pois $\text{dom} f \subseteq \mathcal{H}$. Então, pela Proposição 1.1.1, basta mostrar que f_γ é limitada superiormente em uma vizinhança de x_0 , para todo $x_0 \in \mathcal{H}$. Para isto, tome $\rho \in \mathbb{R}_{++}$ e $x_0 \in \mathcal{H}$ e fixe $z \in \text{dom} f$. Segue que, para

todo $y \in B(x_0, \rho)$,

$$\begin{aligned}
 f_\gamma(y) &\leq f(z) + \|y - z\|^2/(2\gamma) \\
 &\leq f(z) + (\|y - x_0\| + \|x_0 - z\|)^2/(2\gamma) \\
 &= f(z) + (\|y - x_0\|^2 + \|x_0 - z\|^2 + 2\|y - x_0\|\|x_0 - z\|)/(2\gamma) \\
 &\leq f(z) + (\rho^2 + \|x_0 - z\|^2 + 2\rho\|x_0 - z\|)/(2\gamma) \\
 &=: \eta.
 \end{aligned}$$

Note que, η é um número real. Logo, $\sup f_\gamma(B(x_0, \rho)) \leq \eta$, como queríamos demonstrar.

Ainda, pela Proposição 1.2.1, como g é supercoerciva, então, para todo x em \mathcal{H} , a função $f + g(x - \cdot)$ é coerciva. Além disso, como g é estritamente convexa pelo Exemplo 1.1.1, temos, para todo x em \mathcal{H} , que $f + g(x - \cdot)$ é estritamente convexa. Assim, pelo Teorema 1.2.1 concluímos que $f + g(x - \cdot)$ tem um único minimizador em \mathcal{H} , ou seja, para todo x em \mathcal{H} , o ínfimo em (3.1) é unicamente atingido. ■

A Proposição 3.1.1, com $\gamma = 1$, motiva a definição do principal objeto de estudo deste trabalho.

Definição 3.1.2. *Seja $f \in \Gamma_0(\mathcal{H})$. O operador*

$$\begin{aligned}
 \text{prox}_f : \mathcal{H} &\rightarrow \mathcal{H} \\
 x &\mapsto \arg \min_{y \in \mathcal{H}} f(y) + \frac{1}{2}\|x - y\|^2
 \end{aligned}$$

é o operador proximal de f .

Notemos que prox_f está bem definido, pois pela Proposição 3.1.1 para todo x em \mathcal{H} a função $y \mapsto f(y) + (1/2)\|x - y\|^2$ tem um único minimizador em \mathcal{H} . Convém ressaltar que para $\gamma \in \mathbb{R}_{++}$, temos que

$$\begin{aligned}
 \text{prox}_{\gamma f} x &= \arg \min_{y \in \mathcal{H}} \gamma f(y) + \frac{1}{2}\|x - y\|^2 \\
 &= \arg \min_{y \in \mathcal{H}} \gamma \left(f(y) + \frac{1}{2\gamma}\|x - y\|^2 \right) \\
 &= \arg \min_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma}\|x - y\|^2.
 \end{aligned}$$

A última igualdade é justificada pelo fato que minimizar um múltiplo escalar de uma função é o mesmo que minimizar a função.

Observação 3.1.1. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $x \in \mathcal{H}$. Então, $\text{prox}_{\gamma f}x$ é o único ponto em \mathcal{H} que satisfaz*

$$f_{\gamma}(x) = f(\text{prox}_{\gamma f}x) + \frac{1}{2\gamma}\|x - \text{prox}_{\gamma f}x\|^2. \quad (3.2)$$

Logo, para todo $x \in \mathcal{H}$, o ínfimo em (3.1) é unicamente atingido pelo ponto $\text{prox}_{\gamma f}x$.

Exemplo 3.1.2. *O operador proximal da função nula é o operador identidade. Em outras palavras, o operador identidade é um operador proximal, uma vez que existe uma função $f = 0$ em $\Gamma_0(\mathcal{H})$ tal que $\text{prox}_f x = x$ para todo $x \in \mathcal{H}$, isto é, $\text{prox}_f = \text{I}$.*

Exemplo 3.1.3. *Sejam C um subconjunto não-vazio, fechado e convexo de \mathcal{H} e $\gamma \in \mathbb{R}_{++}$. Se $f = \iota_C$, então*

$$\text{prox}_{\gamma f} = P_C.$$

De fato, para todo x em \mathcal{H} , existe um único ponto P_Cx em C tal que

$$(\forall y \in C) \quad \|x - P_Cx\| \leq \|x - y\|.$$

Disto e do fato que $f(P_Cx) = 0$ segue que

$$(\forall y \in \mathcal{H}) \quad f(P_Cx) + \frac{1}{2\gamma}\|x - P_Cx\|^2 \leq f(y) + \frac{1}{2\gamma}\|x - y\|^2.$$

Então, P_Cx é o minimizador de f_{γ} , mas f_{γ} possui um único minimizador, que é por definição $\text{prox}_{\gamma f}x$. Logo, $P_Cx = \text{prox}_{\gamma f}x$ para todo x em \mathcal{H} , isto é, $\text{prox}_{\gamma f} = P_C$.

Operadores proximais são, portanto, uma generalização de operadores de projeção em conjuntos convexos. Mais exemplos de operadores proximais serão apresentados na Seção 3.4.

Uma caracterização para $\text{prox}_{\gamma f}x$ é dada pelo seguinte resultado.

Proposição 3.1.2. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $x, p \in \mathcal{H}$. Então,*

$$p = \text{prox}_{\gamma f}x \Leftrightarrow x - p \in \gamma \partial f(p). \quad (3.3)$$

Demonstração. Fixe x em \mathcal{H} e seja $p = \text{prox}_{\gamma f}x$. Pela Observação 3.1.1, o ponto p é o minimizador de $y \mapsto f(y) + (1/2\gamma)\|x - y\|^2$. Isto é equivalente a dizer, por (1.20), usando a Proposição 1.5.3 e o Lema

1.5.1, que

$$0 \in \partial \left(f + \frac{1}{2\gamma} \|x - \cdot\|^2 \right) (p) = \partial f(p) + \left\{ \frac{p-x}{\gamma} \right\},$$

isto é, $x - p \in \gamma \partial f(p)$. ■

Operadores proximais possuem propriedades muito atrativas, que os tornam particularmente adequados para iterações de algoritmos de minimização, por exemplo eles são firmemente não-expansivos e seu conjunto de ponto fixo é precisamente o conjunto dos minimizadores de f . Tais propriedades permitem o desenvolvimento de algoritmos com base em operadores proximais.

Proposição 3.1.3. *Seja $f \in \Gamma_0(\mathcal{H})$. Então, os operadores prox_f e $I - \text{prox}_f$ são firmemente não-expansivos.*

Demonstração. Sejam x e y em \mathcal{H} . Defina $p = \text{prox}_f x$ e $q = \text{prox}_f y$. Como p e q estão em \mathcal{H} , segue da equivalência (3.3) e de (1.18) que

$$\langle q - p | x - p \rangle + f(p) \leq f(q) \quad e \quad \langle p - q | y - q \rangle + f(q) \leq f(p).$$

Somando estas duas desigualdades, obtemos

$$\langle q - p | x - p \rangle + \langle p - q | y - q \rangle \leq 0,$$

que é equivalente a

$$\|p - q\|^2 \leq \langle p - q | x - y \rangle.$$

Logo, pela Proposição 2.1.1, os operadores prox_f e $I - \text{prox}_f$ são firmemente não-expansivos. ■

Segue da Proposição 3.1.1, que a regularização de Moreau pertence a $\Gamma_0(\mathcal{H})$ e é contínua. O próximo resultado afirma que ela é na verdade diferenciável em \mathcal{H} .

Proposição 3.1.4. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $\gamma \in \mathbb{R}_{++}$. Então, a função f_γ é diferenciável em \mathcal{H} e*

$$\nabla(f_\gamma) = (I - \text{prox}_{\gamma f})/\gamma. \tag{3.4}$$

Além disso, $\nabla(f_\gamma)$ é $(1/\gamma)$ -Lipschitz contínuo.

Demonstração. Sejam x e y pontos distintos em \mathcal{H} . Defina $p = \text{prox}_{\gamma f} x$ e $q = \text{prox}_{\gamma f} y$. De (3.2) segue

$$f_{\gamma}(x) = f(p) + \|x - p\|^2/(2\gamma) \quad \text{e} \quad f_{\gamma}(y) = f(q) + \|y - q\|^2/(2\gamma). \quad (3.5)$$

Por (3.3) e (1.18) obtemos

$$f(q) - f(p) \geq \langle q - p | x - p \rangle / \gamma \quad (3.6)$$

e

$$f(q) - f(p) \leq -\langle p - q | y - q \rangle / \gamma. \quad (3.7)$$

Com manipulações sobre a norma e o produto escalar, concluímos que

$$\begin{aligned} 2\langle y - x | x - p \rangle + \|y - q - x + p\|^2 &= \\ &= \|y - q\|^2 - \|x - p\|^2 + 2\langle q - p | x - p \rangle \end{aligned} \quad (3.8)$$

e

$$\begin{aligned} 2\langle y - x | y - q \rangle - \|y - q - x + p\|^2 &= \\ &= \|y - q\|^2 - \|x - p\|^2 - 2\langle p - q | y - q \rangle. \end{aligned} \quad (3.9)$$

Usando (3.5), (3.6) e (3.8), temos

$$\begin{aligned} f_{\gamma}(y) - f_{\gamma}(x) &= f(q) - f(p) + (\|y - q\|^2 - \|x - p\|^2)/(2\gamma) \\ &\geq (2\langle q - p | x - p \rangle + \|y - q\|^2 - \|x - p\|^2)/(2\gamma) \\ &= (2\langle y - x | x - p \rangle + \|y - q - x + p\|^2)/(2\gamma) \\ &\geq \langle y - x | x - p \rangle / \gamma. \end{aligned} \quad (3.10)$$

Por outro lado, usando (3.5), (3.7) e (3.9),

$$\begin{aligned} f_{\gamma}(y) - f_{\gamma}(x) &= f(q) - f(p) + (\|y - q\|^2 - \|x - p\|^2)/(2\gamma) \\ &\leq (-2\langle p - q | y - q \rangle + \|y - q\|^2 - \|x - p\|^2)/(2\gamma) \\ &= (2\langle y - x | y - q \rangle - \|y - q - x + p\|^2)/(2\gamma) \\ &\leq \langle y - x | y - q \rangle / \gamma. \end{aligned} \quad (3.11)$$

Combinando as desigualdades (3.10) e (3.11), usando o fato que $\text{prox}_{\gamma f}$ é firmemente não-expansivo e a Proposição 2.1.1(iv), segue que

$$0 \leq f_{\gamma}(y) - f_{\gamma}(x) - \langle y - x | x - p \rangle / \gamma$$

$$\begin{aligned}
&\leq (\langle y-x|y-q\rangle - \langle y-x|x-p\rangle)/\gamma \\
&= (\|y-x\|^2 - \langle x-y|p-q\rangle)/\gamma \\
&\leq (\|y-x\|^2 - \|p-q\|^2)/\gamma \\
&\leq \|y-x\|^2/\gamma.
\end{aligned}$$

Assim,

$$\frac{f_\gamma(y) - f_\gamma(x) - \langle y-x|(x-p)/\gamma\rangle}{\|y-x\|} \leq \frac{\|y-x\|}{\gamma}.$$

Tomando o limite quando y tende a x em ambos os lados da desigualdade acima, obtemos

$$\lim_{y \rightarrow x} \frac{|f_\gamma(y) - f_\gamma(x) - \langle y-x|(x-p)/\gamma\rangle|}{\|y-x\|} = 0.$$

Logo, a função f_γ é diferenciável em \mathcal{H} e $\nabla(f_\gamma)(x) = (x-p)/\gamma$, isto é, $\nabla(f_\gamma) = (\text{I} - \text{prox}_{\gamma f})/\gamma$.

Finalmente, a Lipschitz continuidade do gradiente da f_γ segue da Proposição 3.1.3. ■

Corolário 3.1.1. *Seja C um subconjunto não-vazio, fechado e convexo de \mathcal{H} . Então, a função d_C^2 é diferenciável em \mathcal{H} e*

$$\nabla(d_C^2) = 2(\text{I} - P_C). \quad (3.12)$$

Demonstração. Considerando $f = \iota_C$ e $\gamma = 1/2$, pelo Exemplo 3.1.1, temos que $d_C^2 = f_\gamma$. Aplicando a proposição anterior e usando Exemplo 3.1.3, obtemos o resultado. ■

3.2 CÁLCULO PROXIMAL

Nesta seção, trabalhamos com o cálculo proximal, que segue especialmente da inclusão (3.3). Apresentamos expressões para calcular o operador proximal de algumas funções pertencentes a $\Gamma_0(\mathcal{H})$.

Proposição 3.2.1. *Sejam $f \in \Gamma_0(\mathcal{H})$ e $x \in \mathcal{H}$. Temos:*

(i) *Perturbação quadrática: seja $g = f + \alpha\|\cdot\|^2/2 + \langle \cdot|u\rangle + \beta$, onde*

$u \in \mathcal{H}$, $\alpha \in \mathbb{R}_+$ e $\beta \in \mathbb{R}$. Então,

$$\text{prox}_g x = \text{prox}_{\frac{1}{\alpha+1}f} \left(\frac{x-u}{\alpha+1} \right).$$

(ii) *Translação*: seja $g = f(\cdot - z)$, onde $z \in \mathcal{H}$. Então,

$$\text{prox}_g x = z + \text{prox}_f(x - z).$$

(iii) *Scaling*: seja $g = f(\cdot/\rho)$, onde $\rho \in \mathbb{R} \setminus \{0\}$. Então,

$$\text{prox}_g x = \rho \text{prox}_{\frac{1}{\rho^2}f} \left(\frac{x}{\rho} \right).$$

(iv) *Reflexão*: seja $g : y \mapsto f(-y)$. Então,

$$\text{prox}_g x = -\text{prox}_f(-x).$$

(v) *Regularização de Moreau*: seja $g = f_\gamma$, onde $\gamma \in \mathbb{R}_{++}$. Então,

$$\text{prox}_g x = x + \frac{1}{1+\gamma} (\text{prox}_{(\gamma+1)f} x - x).$$

(vi) *Complemento de Moreau*: seja $g = \|\cdot\|^2/(2\gamma) - f_\gamma$, onde $\gamma \in \mathbb{R}_{++}$. Então,

$$\text{prox}_g x = x - \frac{1}{\gamma} \text{prox}_{\frac{\gamma^2}{\gamma+1}f} \left(\frac{\gamma x}{\gamma+1} \right).$$

Demonstração. Em todos os casos, temos que g está em $\Gamma_0(\mathcal{H})$.

(i) Considere $\varphi = \alpha\|\cdot\|^2/2 + \langle \cdot, u \rangle + \beta$. Temos que o domínio da φ é \mathcal{H} e φ é diferenciável em \mathcal{H} com $\nabla\varphi : x \mapsto \alpha x + u$. Logo, φ é contínua em todos os pontos de \mathcal{H} e

$$\text{dom}f \cap \text{dom}\varphi = \text{dom}f \cap \mathcal{H} = \text{dom}f \neq \emptyset.$$

Assim, pela Proposição 1.5.3, pelo Lema 1.5.1 e por (3.3),

$$\begin{aligned} p = \text{prox}_g x &\Leftrightarrow x - p \in \partial g(p) = \partial(f + \varphi) = \partial f(p) + \{\alpha p + u\} \\ &\Leftrightarrow x - p - \alpha p - u \in \partial f(p) \\ &\Leftrightarrow x - u - (\alpha + 1)p \in \partial f(p) \\ &\Leftrightarrow \frac{x - u}{\alpha + 1} - p \in \frac{1}{\alpha + 1} \partial f(p) \end{aligned}$$

$$\Leftrightarrow p = \operatorname{prox}_{\frac{1}{\alpha+1}f} \left(\frac{x-u}{\alpha+1} \right).$$

(ii) Segue de (3.3) que

$$\begin{aligned} p = \operatorname{prox}_g x &\Leftrightarrow x - p \in \partial g(p) = \partial f(p - z) \\ &\Leftrightarrow (x - z) - (p - z) \in \partial f(p - z) \\ &\Leftrightarrow p - z = \operatorname{prox}_f(x - z) \\ &\Leftrightarrow p = z + \operatorname{prox}_f(x - z). \end{aligned}$$

(iii) Temos que $\partial g = \partial f(\cdot/\rho)$. Primeiro devemos verificar que $\partial f(\cdot/\rho) = (1/\rho)\partial f(\cdot/\rho)$. Observe que

$$\begin{aligned} (f(\cdot/\rho))^*(u) &= \sup_{x \in \mathcal{H}} \langle x|u \rangle - f(x/\rho) \\ &= \sup_{x \in \mathcal{H}} \langle \rho^{-1}x|\rho u \rangle - f(x/\rho) \\ &= \sup_{\frac{x}{\rho} \in \mathcal{H}} \langle \rho^{-1}x|\rho u \rangle - f(x/\rho) \\ &= f^*(\rho u). \end{aligned}$$

Disto e de (1.19) segue

$$\begin{aligned} u \in \partial f(x/\rho) &\Leftrightarrow f(x/\rho) + (f(\cdot/\rho))^*(u) = \langle p|u \rangle \\ &\Leftrightarrow f(x/\rho) + f^*(\rho u) = \langle \rho^{-1}x|\rho u \rangle \\ &\Leftrightarrow \rho u \in \partial f(x/\rho) \\ &\Leftrightarrow u \in \frac{1}{\rho}\partial f(x/\rho). \end{aligned}$$

Usando (3.3) concluímos que

$$\begin{aligned} p = \operatorname{prox}_g x &\Leftrightarrow x - p \in \partial g(p) = \partial f(p/\rho) = \frac{1}{\rho}\partial f(p/\rho) \\ &\Leftrightarrow \frac{x}{\rho} - \frac{p}{\rho} \in \frac{1}{\rho^2}\partial f(p/\rho) \\ &\Leftrightarrow \frac{p}{\rho} = \operatorname{prox}_{\frac{1}{\rho^2}f} \left(\frac{x}{\rho} \right) \\ &\Leftrightarrow p = \rho \operatorname{prox}_{\frac{1}{\rho^2}f} \left(\frac{x}{\rho} \right). \end{aligned}$$

(iv) Basta tomar $\rho = -1$ em (iii).

(v) A função g é diferenciável com $\nabla g = (I - \text{prox}_{\gamma f})/\gamma$, pela Proposição 3.1.4. Disto e de (3.3), segue que

$$\begin{aligned}
p = \text{prox}_g x &\Leftrightarrow x - p \in \partial g(p) = \{(p - \text{prox}_{\gamma f} p)/\gamma\} \\
&\Leftrightarrow x - p = (p - \text{prox}_{\gamma f} p)/\gamma \\
&\Leftrightarrow p - \gamma(x - p) = \text{prox}_{\gamma f} p \\
&\Leftrightarrow p - (p - \gamma(x - p)) \in \gamma \partial f(p - \gamma(x - p)) \\
&\Leftrightarrow x - p \in \partial f(p - \gamma(x - p)) \\
&\Leftrightarrow (\gamma + 1)(x - p) \in (\gamma + 1)\partial f(p - \gamma(x - p)) \\
&\Leftrightarrow x - (p - \gamma(x - p)) \in (\gamma + 1)\partial f(p - \gamma(x - p)) \\
&\Leftrightarrow p - \gamma(x - p) = \text{prox}_{(\gamma+1)f} x \\
&\Leftrightarrow p = \frac{1}{\gamma + 1}(\gamma x + x - x + \text{prox}_{(\gamma+1)f} x) \\
&\Leftrightarrow p = x + \frac{1}{\gamma + 1}(\text{prox}_{(\gamma+1)f} x - x).
\end{aligned}$$

(vi) Primeiramente, vamos mostrar que g está em $\Gamma_0(\mathcal{H})$. Para isto, defina $\psi = f + \|\cdot\|^2/(2\gamma)$. A função ψ está em $\Gamma_0(\mathcal{H})$, uma vez que é soma de duas funções convexas, s.c.i. e próprias tais que a interseção de seus domínios é diferente de vazio. Assim, ψ^* está em $\Gamma_0(\mathcal{H})$. Por (1.13) e (3.1), para todo $u \in \mathcal{H}$, temos que

$$\begin{aligned}
\psi^*(u) &= \sup_{x \in \mathcal{H}} \langle x|u \rangle - \psi(x) \\
&= - \inf_{x \in \mathcal{H}} \psi(x) - \langle x|u \rangle \\
&= - \inf_{x \in \mathcal{H}} f(x) + \frac{\|x\|^2}{2\gamma} - \langle x|u \rangle \\
&= - \inf_{x \in \mathcal{H}} f(x) + \frac{\|\gamma u - x\|^2}{2\gamma} - \frac{\|\gamma u\|^2}{2\gamma} \\
&= \frac{\|\gamma u\|^2}{2\gamma} - \inf_{x \in \mathcal{H}} f(x) + \frac{\|\gamma u - x\|^2}{2\gamma} \\
&= \frac{\|\gamma u\|^2}{2\gamma} - f_\gamma(\gamma u).
\end{aligned}$$

Segue que, $\psi^*(\cdot/\gamma) = g$. Logo, g está em $\Gamma_0(\mathcal{H})$.

Agora, tendo em vista a Proposição 3.1.4, temos que g é diferenciável

com $\nabla g = \text{prox}_{\gamma f}/\gamma$. Consequentemente, segue de (3.3) que

$$\begin{aligned}
 p = \text{prox}_g x &\Leftrightarrow x - p \in \partial g(p) = \{(\text{prox}_{\gamma f} p)/\gamma\} \\
 &\Leftrightarrow \gamma(x - p) = \text{prox}_{\gamma f} p \\
 &\Leftrightarrow p - \gamma(x - p) \in \gamma \partial f(\gamma(x - p)) \\
 &\Leftrightarrow \frac{\gamma}{\gamma + 1}(p - \gamma(x - p) + x - x) \in \frac{\gamma^2}{\gamma + 1} \partial f(\gamma(x - p)) \\
 &\Leftrightarrow \frac{\gamma x}{\gamma + 1} - \gamma(x - p) \in \frac{\gamma^2}{\gamma + 1} \partial f(\gamma(x - p)) \\
 &\Leftrightarrow \gamma(x - p) = \text{prox}_{\frac{\gamma^2}{\gamma + 1} f} \left(\frac{\gamma x}{\gamma + 1} \right) \\
 &\Leftrightarrow p = x - \frac{1}{\gamma} \text{prox}_{\frac{\gamma^2}{\gamma + 1} f} \left(\frac{\gamma x}{\gamma + 1} \right).
 \end{aligned}$$

■

A proposição seguinte estabelece uma fórmula para calcular o operador proximal da composição de uma função com um operador linear, limitado e unitário, em termos do operador proximal dessa função.

Proposição 3.2.2. *Considere \mathcal{K} um espaço de Hilbert real. Sejam $f = g \circ L$, onde $g \in \Gamma_0(\mathcal{K})$ e $L : \mathcal{H} \rightarrow \mathcal{K}$ é um operador linear, limitado e bijetivo tal que $L^{-1} = L^*$, e $x \in \mathcal{H}$. Então,*

$$\text{prox}_f x = L^*(\text{prox}_g(Lx))$$

Demonstração. Segue das hipóteses que $f \in \Gamma_0(\mathcal{H})$. Sejam x e p em \mathcal{H} . Pela Proposição 1.5.2, temos que $\partial f = L^* \circ (\partial g) \circ L$. Assim, segue de (3.3) que

$$\begin{aligned}
 p = \text{prox}_f x &\Leftrightarrow x - p \in L^*(\partial g(Lp)) \\
 &\Leftrightarrow Lx - Lp \in \partial g(Lp) \\
 &\Leftrightarrow Lp = \text{prox}_g Lx \\
 &\Leftrightarrow p = L^*(\text{prox}_g(Lx)).
 \end{aligned}$$

■

O próximo resultado, diz respeito ao cálculo de operadores proximais no espaço produto.

Proposição 3.2.3. *Seja $I = \{1, \dots, m\}$ um conjunto de índices. Seja $(\mathcal{H}_i)_{i \in I}$ uma família de espaços de Hilbert reais. Considere $\mathcal{H}^m = \mathcal{H}_1 \times \dots \times \mathcal{H}_m$. Para todo $i \in I$, sejam $f_i \in \Gamma_0(\mathcal{H}_i)$ e $x_i \in \mathcal{H}_i$. Defina $x = (x_i)_{i \in I}$ e $f : \mathcal{H}^m \rightarrow (-\infty, +\infty]$ por $f(x) = \sum_{i \in I} f_i(x_i)$.*

Então,

$$\text{prox}_f x = (\text{prox}_{f_i} x_i)_{i \in I}.$$

Demonstração. Temos que $f \in \Gamma_0(\mathcal{H}^m)$. Tome $x = (x_i)_{i \in I}$ e $p = (p_i)_{i \in I}$ em \mathcal{H}^m . Segue, de (3.3) e da Proposição 1.5.1, que

$$\begin{aligned} p = \text{prox}_f x &\Leftrightarrow x - p \in \partial f(p) = \partial f_1(p_1) \times \dots \times \partial f_m(p_m) \\ &\Leftrightarrow x_i - p_i \in \partial f_i(p_i) \quad (\forall i \in I) \\ &\Leftrightarrow p_i = \text{prox}_{f_i} x_i \quad (\forall i \in I) \\ &\Leftrightarrow p = (\text{prox}_{f_i} x_i)_{i \in I}. \end{aligned}$$

■

3.3 DECOMPOSIÇÃO DE MOREAU

A decomposição ortogonal linear de um espaço com relação a um subespaço fechado e seu complemento ortogonal (ver [16, Teorema 3.3.4]), pode ser estendida por meio de um princípio não-linear conhecido como princípio de decomposição de Moreau, em que um ponto é decomposto unicamente em termos do operador proximal.

Moreau mostrou em [18] que este princípio de decomposição vale para qualquer função em $\Gamma_0(\mathcal{H})$, usando a regularização de Moreau

$$f_\gamma : x \mapsto \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|x - y\|^2,$$

definida na Seção 3.1.

Teorema 3.3.1. *Sejam $f \in \Gamma_0(\mathcal{H})$, $\gamma \in \mathbb{R}_{++}$ e $x \in \mathcal{H}$. Então*

$$\|x\|^2 = 2\gamma(f_\gamma(x) + (f^*)_{1/\gamma}(x/\gamma)) \quad (3.13)$$

e

$$x = x_\gamma^\oplus + x_\gamma^\ominus, \quad \text{onde} \quad \begin{cases} x_\gamma^\oplus = \text{prox}_{\gamma f} x \\ x_\gamma^\ominus = \gamma \text{prox}_{f^*/\gamma}(x/\gamma). \end{cases} \quad (3.14)$$

Além disso,

$$f(x_\gamma^\oplus) + f^*(x_\gamma^\ominus/\gamma) = \langle x_\gamma^\oplus | x_\gamma^\ominus \rangle / \gamma. \quad (3.15)$$

Demonstração. Defina $g : y \mapsto \|x - y\|^2/(2\gamma)$. Observe que g está em $\Gamma_0(\mathcal{H})$ e em cada $v \in \mathcal{H}$, usando o resultado do Exemplo 1.4.1,

$$\begin{aligned} g^*(v) &= \sup_{y \in \mathcal{H}} \langle y | v \rangle - \frac{1}{2\gamma} \|x - y\|^2 \\ &= \sup_{y \in \mathcal{H}} \langle y - x | v \rangle - \frac{1}{2\gamma} \|x - y\|^2 + \langle x | v \rangle \\ &= \frac{1}{\gamma} \sup_{y \in \mathcal{H}} \left(\langle y - x | \gamma v \rangle - \frac{1}{2} \|x - y\|^2 \right) + \langle x | v \rangle \\ &= \frac{1}{2\gamma} \|\gamma v\|^2 + \langle x | v \rangle \\ &= \frac{\gamma}{2} \|v\|^2 + \langle x | v \rangle. \end{aligned}$$

Como $\text{dom} g = \mathcal{H}$ e g é contínua em \mathcal{H} , podemos usar o Teorema 1.4.1 e assim obtemos

$$\begin{aligned} f_\gamma(x) &= \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|x - y\|^2 \\ &= -\min_{v \in \mathcal{H}} f^*(v) + \frac{\gamma}{2} \|v\|^2 - \langle x | v \rangle \\ &= -\min_{v \in \mathcal{H}} f^*(v) + \frac{\gamma}{2} \|\gamma^{-1}x - v\|^2 - \frac{\gamma}{2} \|\gamma^{-1}x\|^2 \\ &= \frac{1}{2\gamma} \|x\|^2 - \min_{v \in \mathcal{H}} f^*(v) + \frac{\gamma}{2} \|\gamma^{-1}x - v\|^2 \\ &= \frac{1}{2\gamma} \|x\|^2 - (f^*)_{1/\gamma}(x/\gamma), \end{aligned}$$

que fornece (3.13). Agora, para obter (3.14) basta diferenciar (3.13) e usar a Proposição 3.1.4, conforme segue

$$\begin{aligned} \nabla(\|x\|^2/(2\gamma)) &= \nabla(f_\gamma)(x) + \nabla((f^*)_{1/\gamma})(x/\gamma) \\ \frac{2\|x\|}{2\gamma} \cdot \frac{x}{\|x\|} &= \frac{1}{\gamma}(x - \text{prox}_{\gamma f}x) + \gamma \left(\frac{x}{\gamma} - \text{prox}_{f^*/\gamma}(x/\gamma) \right) \cdot \frac{1}{\gamma} \\ \frac{x}{\gamma} &= \frac{x}{\gamma} - \frac{\text{prox}_{\gamma f}x}{\gamma} + \frac{x}{\gamma} - \text{prox}_{f^*/\gamma}(x/\gamma) \\ x &= \text{prox}_{\gamma f}x + \gamma \text{prox}_{f^*/\gamma}(x/\gamma). \end{aligned}$$

Finalmente, observe que (3.3), (3.14) e a definição de subdiferencial implicam que

$$\begin{aligned} x_\gamma^\oplus = \text{prox}_{\gamma f} x &\Leftrightarrow x - x_\gamma^\oplus \in \gamma \partial f(x_\gamma^\oplus) \\ &\Leftrightarrow x_\gamma^\ominus \in \gamma \partial f(x_\gamma^\oplus) \\ &\Leftrightarrow x_\gamma^\ominus / \gamma \in \partial f(x_\gamma^\oplus) \\ &\Leftrightarrow f(x_\gamma^\oplus) + f^*(x_\gamma^\ominus / \gamma) = \langle x_\gamma^\oplus | x_\gamma^\ominus \rangle / \gamma, \end{aligned}$$

que fornece (3.15). ■

De acordo com o Teorema 3.3.1, dada qualquer função f convexa, própria e s.c.i. em \mathcal{H} , cada ponto x em \mathcal{H} pode ser decomposto com respeito a f em uma soma $x = x_\gamma^\oplus + x_\gamma^\ominus$ tal que $x_\gamma^\ominus \in \partial f(x_\gamma^\oplus)$.

A simetria quando $\gamma = 1$ no Teorema 3.3.1 merece ser observada. Uma vez que, com $\gamma = 1$ o Teorema 3.3.1 fornece

$$f_1 + (f^*)_1 = \frac{\|\cdot\|^2}{2} \quad \text{e} \quad \text{prox}_f + \text{prox}_{f^*} = \text{I}, \quad (3.16)$$

usando a Proposição 3.1.4, obtemos o gradiente da conjugada de Fenchel de f_1 :

$$\nabla(f_1^*) = \text{I} - \nabla(f_1) = \text{prox}_f.$$

Se $f = \iota_K$, onde K é um cone não-vazio, fechado e convexo em \mathcal{H} , obtemos a decomposição cônica de Moreau. Lembremos que $K \subset \mathcal{H}$ é um cone convexo se $K + K = \{x + y \mid x \in K, y \in K\} \subset K$ e $\alpha K = \{\alpha x \mid x \in K\} \subset K$ para todo $\alpha \in \mathbb{R}_{++}$.

O cone polar de K é o cone não-vazio, fechado e convexo em \mathcal{H} definido por $K^\ominus = \{u \in \mathcal{H} \mid (\forall x \in K) \langle x | u \rangle \leq 0\}$. Por exemplo, se K é um subespaço vetorial, então seu cone polar é o seu complemento ortogonal, denotado por K^\perp .

Corolário 3.3.1. *Sejam K um cone não-vazio, fechado e convexo em \mathcal{H} com cone polar K^\ominus e $x \in \mathcal{H}$. Então,*

(i) $x = P_K x + P_{K^\ominus} x.$

(ii) $\langle P_K x | P_{K^\ominus} x \rangle = 0.$

(iii) $\|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x).$

Demonstração. Considere $f = \iota_K$ e $\gamma = 1$ no Teorema 3.3.1. Por (1.14)

$$\begin{aligned} f^*(u) &= \sup_{x \in K} \langle x|u \rangle \\ &= \begin{cases} 0, & \text{se } u \in K^\ominus \\ +\infty, & \text{se } u \notin K^\ominus \end{cases} \\ &= \iota_{K^\ominus}(u), \end{aligned}$$

para todo $u \in \mathcal{H}$, isto é, $f^* = \iota_{K^\ominus}$. Pelo Exemplo 3.1.1, temos que $f_1 = d_K^2/2$ e $(f^*)_1 = d_{K^\ominus}^2/2$, substituindo em (3.13) obtemos (iii). Pelo Exemplo 3.1.3, temos que $\text{prox}_f = P_K$ e $\text{prox}_{f^*} = P_{K^\ominus}$, substituindo em (3.14) obtemos (i). Por fim, substituindo em (3.15) segue que

$$f(P_K x) + f^*(P_{K^\ominus} x) = \langle P_K x | P_{K^\ominus} x \rangle. \quad (3.17)$$

Como $P_K x \in K$ e $P_{K^\ominus} x \in K^\ominus$, então $f(P_K x) = 0$ e $f^*(P_{K^\ominus} x) = 0$. Logo, (3.17) fornece (ii). \blacksquare

Se $K = V$, onde V um subespaço vetorial fechado de \mathcal{H} , o Teorema 3.3.1 se reduz à conhecida decomposição ortogonal.

Corolário 3.3.2. *Seja V um subespaço vetorial não-vazio e fechado de \mathcal{H} com complemento ortogonal V^\perp e $x \in \mathcal{H}$. Então,*

$$(i) \quad x = P_V x + P_{V^\perp} x.$$

$$(ii) \quad \langle P_V x | P_{V^\perp} x \rangle = 0.$$

$$(iii) \quad \|x\|^2 = d_V^2(x) + d_{V^\perp}^2(x).$$

Demonstração. Basta considerar $K = V$ no Corolário 3.3.1, neste caso $K^\ominus = V^\perp$ conforme já mencionado. \blacksquare

Observação 3.3.1. *Uma forma alternativa de escrever (3.14) é*

$$x = x_\gamma^\oplus + x_\gamma^\ominus, \quad \text{onde} \quad \begin{cases} x_\gamma^\oplus = \text{prox}_{\gamma f} x \\ x_\gamma^\ominus = \text{prox}_{\gamma f^*}(\cdot/\gamma)x. \end{cases} \quad (3.18)$$

Basta tomar $g = \gamma f^(\cdot/\gamma)$ na Proposição 3.2.1(iii).*

3.4 EXEMPLOS DE OPERADORES PROXIMAIS

Nesta seção, fornecemos alguns exemplos de operadores proximais que são de interesse em problemas de recuperação de sinal, como veremos na Seção 4.2.

Exemplo 3.4.1. *A translação $x \mapsto x - u$ é um operador proximal. Para verificar esta afirmação devemos mostrar que existe uma função $g \in \Gamma_0(\mathcal{H})$ tal que $\text{prox}_g x = x - u$. De fato, basta tomar $f = 0$ e $\alpha = 0$ na Proposição 3.2.1(i) e usar o Exemplo 3.1.2. Assim, obtemos a função $g = \langle \cdot | u \rangle + \beta$, onde $u \in \mathcal{H}$ e $\beta \in \mathbb{R}$, cujo operador proximal é a translação $x \mapsto x - u$.*

Exemplo 3.4.2. *A transformação $x \mapsto \kappa x$ é um operador proximal, com $\kappa \in (0, 1]$. Basta tomar $f = 0$, $u = 0$ e $\alpha = 1/\kappa - 1$ na Proposição 3.2.1(i) e usar o Exemplo 3.1.2. Assim, obtemos a função $g = (1 - \kappa)\|x\|^2/(2\kappa) + \beta$, onde $\beta \in \mathbb{R}$, tal que seu operador proximal é a transformação $x \mapsto \kappa x$.*

Do ponto de vista numérico, a decomposição de Moreau fornece uma forma alternativa para calcular $x_\gamma^\oplus = \text{prox}_{\gamma f} x$. Visto que, por (3.14) podemos encontrar x_γ^\oplus resolvendo $x_\gamma^\oplus = x - \gamma \text{prox}_{f^*/\gamma}(x/\gamma)$, ou, por (3.18)

$$x_\gamma^\oplus = x - \text{prox}_{\gamma f^*(\cdot/\gamma)} x. \quad (3.19)$$

Isto é especialmente importante em situações em que pode ser difícil obter $\text{prox}_{\gamma f} x$ diretamente mas em que $\text{prox}_{f^*/\gamma}(x/\gamma)$ é mais fácil de ser obtido. O seguinte exemplo ilustra este ponto.

Exemplo 3.4.3. *Seja $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ definida por*

$$f : x \mapsto \sup_{y \in D} \langle x | Ly \rangle,$$

onde $L : \mathcal{K} \rightarrow \mathcal{H}$ é um operador linear limitado, \mathcal{K} é um espaço de Hilbert e D é um subconjunto não-vazio de \mathcal{K} . Então, $x_\gamma^\oplus = x - P_{\gamma C}$.

De fato, a função f está em $\Gamma_0(\mathcal{H})$. Seja C o menor conjunto convexo fechado que contém $L(D)$. Usando (1.14), podemos escrever

$$f : x \mapsto \sup_{u \in C} \langle x | u \rangle = \sigma_C(x).$$

Por (1.15), temos que $f^* = \sigma_C^* = \iota_C$ e, pelo Exemplo 3.1.3, que $\text{prox}_{\gamma \iota_C} = P_C$. Logo, podemos calcular x_γ^\oplus através de uma operação

projeção para todo $x \in \mathcal{H}$, já que (3.19) torna-se

$$x_{\gamma}^{\oplus} = x - \text{prox}_{\gamma f^*(\cdot/\gamma)}x = x - \text{prox}_{\gamma \iota_C(\cdot/\gamma)}x = x - P_{\gamma C}x.$$

Notemos que o Exemplo 3.4.3 fornece o operador proximal da função suporte de C , isto é, $\text{prox}_{\sigma_C} = I - P_C$.

Agora, vamos apresentar uma aplicação no espaço produto da Proposição 3.2.3.

Exemplo 3.4.4. *Sejam $\gamma \in \mathbb{R}_{++}$ e $I = \{1, \dots, m\}$ um conjunto de índices. Para todo $i \in I$ defina ϕ_i em $\Gamma_0(\mathbb{R})$ por*

$$\phi_i : x_i \mapsto \begin{cases} -\ln(x_i), & \text{se } x_i > 0 \\ +\infty, & \text{se } x_i \leq 0 \end{cases}$$

e f em $\Gamma_0(\mathbb{R}^m)$ por

$$f : x \mapsto \sum_{i=1}^m \phi_i(x_i), \quad x = (x_i)_{i \in I}.$$

Então,

$$(\forall x \in \mathbb{R}^m) \quad \text{prox}_{\gamma f}x = \frac{1}{2} \left(x_i + \sqrt{x_i^2 + 4\gamma} \right)_{1 \leq i \leq m}.$$

Demonstração. Fixe $i \in I$ e seja $x_i \in \mathbb{R}$. Observemos que

$$\partial\phi_i(x_i) = \begin{cases} \left\{ -\frac{1}{x_i} \right\}, & \text{se } x_i > 0 \\ \emptyset, & \text{se } x_i \leq 0. \end{cases}$$

Por (3.3), temos

$$\begin{aligned} p = \text{prox}_{\gamma\phi_i}x_i &\Leftrightarrow x_i - p \in \gamma\partial\phi_i(p) = \left\{ -\frac{\gamma}{p} \right\} \\ &\Leftrightarrow x_i - p = -\frac{\gamma}{p} \\ &\Leftrightarrow p^2 - x_i p - \gamma = 0 \\ &\Leftrightarrow p = \frac{1}{2} \left(x_i + \sqrt{x_i^2 + 4\gamma} \right). \end{aligned}$$

Logo, o resultado segue pela Proposição 3.2.3. ■

Já vimos no Exemplo 3.1.3 que operadores de projeção em conjuntos convexos são operadores proximais. De uma forma mais geral, temos o seguinte exemplo.

Exemplo 3.4.5. *Sejam C um subconjunto não-vazio, fechado e convexo de \mathcal{H} , $\gamma \in \mathbb{R}_{++}$ e $x \in \mathcal{H}$. Então,*

$$\text{prox}_{d_C^2/(2\gamma)}x = x + \frac{1}{\gamma + 1}(P_Cx - x). \quad (3.20)$$

Vimos no Exemplo 3.1.1 e no Exemplo 3.1.3 que se $f = \iota_C$, então $f_\gamma = d_C^2/(2\gamma)$ e $\text{prox}_{f_\gamma} = P_C$. Logo, substituindo na Proposição 3.2.1(v), obtemos o resultado.

A proposição seguinte fornece o exemplo de uma transformação que pode ser expressa por meio de um operador proximal, e que por sua vez é o operador proximal da função distância.

Proposição 3.4.1. *Sejam C um subconjunto não-vazio, fechado e convexo de \mathcal{H} , $\gamma \in \mathbb{R}_{++}$ e $x \in \mathcal{H}$. Então,*

$$\text{prox}_{\gamma d_C}x = \begin{cases} x + \frac{\gamma}{d_C(x)}(P_Cx - x), & \text{se } d_C(x) > \gamma \\ P_Cx, & \text{se } d_C(x) \leq \gamma. \end{cases} \quad (3.21)$$

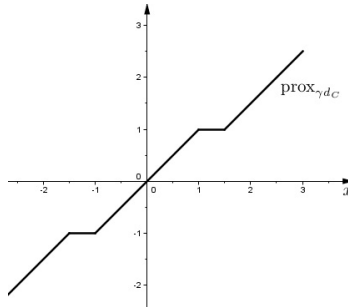


Figura 5: Gráfico de $\text{prox}_{\gamma d_C}$, considerando $\mathcal{H} = \mathbb{R}$, $C = [-1, 1]$ e $\gamma = 1/2$.

Demonstração. Como já observamos a função d_C está em $\Gamma_0(\mathcal{H})$. Seja $p = \text{prox}_{\gamma d_C}x$. Por (3.3), $x - p \in \gamma \partial d_C(p)$. Tendo em vista (1.21), segue de (1.22) que:

(i) Se $p \in C$, então $x - p \in N_C(p) \cap B(0, \gamma)$. Disto, usando (3.3) e Exemplo 3.1.3, temos

$$x - p \in N_C(p) = \partial \iota_C(p) \Leftrightarrow p = \text{prox}_{\iota_C} x = P_C x. \quad (3.22)$$

E, sabendo (1.1),

$$x - p \in B(0, \gamma) \Rightarrow \|x - p\| \leq \gamma \Rightarrow d_C(x) \leq \gamma. \quad (3.23)$$

(ii) Se $p \notin C$, então

$$x - p = \gamma \left(\frac{p - P_C p}{d_C(p)} \right).$$

Somando e subtraindo $P_C p$ no lado esquerdo da igualdade acima, obtemos

$$x - P_C p = \left(1 + \frac{\gamma}{d_C(p)} \right) (p - P_C p) \in N_C(P_C p), \quad (3.24)$$

uma vez que $P_C p \in C$ e $\langle y - P_C p | (1 + \gamma/d_C(p))(p - P_C p) \rangle \leq 0$ para todo $y \in C$, por (1.3). Assim,

$$x - P_C p \in N_C(P_C p) = \partial \iota_C(P_C p) \Leftrightarrow P_C p = \text{prox}_{\iota_C} x = P_C x$$

por (3.3) e Exemplo 3.1.3. Consequentemente, substituindo $P_C p$ no lado esquerdo da igualdade (3.24) por $P_C x$,

$$x - P_C x = \left(1 + \frac{\gamma}{d_C(p)} \right) (p - P_C p). \quad (3.25)$$

Tomando a norma em ambos os lados da igualdade (3.25), ficamos com

$$\|x - P_C x\| = \left(1 + \frac{\gamma}{d_C(p)} \right) \|p - P_C p\|,$$

isto é, por (1.2),

$$d_C(x) = d_C(p) + \gamma. \quad (3.26)$$

Por fim, reescrevendo (3.25), usando (3.26) e o fato que $P_C p = P_C x$, obtemos

$$p = x + \frac{\gamma}{d_C(x)} (P_C x - x). \quad (3.27)$$

Agora, suponha que $d_C(x) > \gamma$, então $p \notin C$, pois caso contrário (3.23) daria $d_C x \leq \gamma$, que é absurdo. A expressão de p é, portanto, fornecida por (3.27).

Em seguida, suponha que $d_C(x) \leq \gamma$, então $p \in C$, pois por (3.26) se

$$p \notin C \Rightarrow d_C(p) = d_C(x) - \gamma \leq 0 \Rightarrow d_C(p) = 0 \Rightarrow p \in \bar{C} = C,$$

que é absurdo. A expressão de p é, portanto, fornecida por (3.22). ■

Na Proposição 3.4.1, C pode ser considerado como um conjunto de sinais que possuem uma certa propriedade (ver [10, 25] para exemplos de conjuntos convexos fechados modelando restrições em problemas de recuperação de sinais). Se o sinal x está próximo o suficiente para satisfazer a propriedade em questão, então $\text{prox}_{\gamma d_C} x$ é simplesmente a projeção de x em C , caso contrário, $\text{prox}_{\gamma d_C} x$ é obtido através de uma relaxação desta projeção. Um caso especial é descrito no exemplo a seguir.

Exemplo 3.4.6. *Suponha que $C = \{0\}$ na Proposição 3.4.1. Então, (3.21) torna-se*

$$\text{prox}_{\gamma \|\cdot\|} x = \begin{cases} \left(1 - \frac{\gamma}{\|x\|}\right)x, & \text{se } \|x\| > \gamma \\ 0, & \text{se } \|x\| \leq \gamma, \end{cases} \quad (3.28)$$

uma vez que $P_C x = 0$ para todo x em \mathcal{H} .

Em particular, se $\mathcal{H} = \mathbb{R}$, temos que

$$\text{prox}_{\gamma|\cdot|} x = \text{sign}(x) \cdot \begin{cases} |x| - \gamma, & \text{se } |x| - \gamma > 0 \\ 0, & \text{se } |x| - \gamma \leq 0, \end{cases}$$

onde

$$\text{sign}(x) = \frac{x}{|x|} = \begin{cases} 1, & \text{se } x > 0 \\ -1, & \text{se } x \leq 0. \end{cases}$$

Então, (3.28) se reduz a

$$\text{prox}_{\gamma|\cdot|} x = \text{sign}(x) \max\{|x| - \gamma, 0\}. \quad (3.29)$$

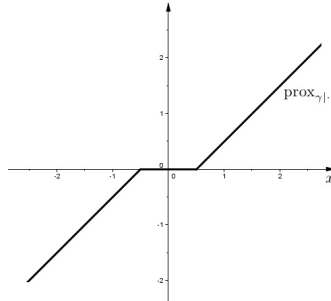


Figura 6: Gráfico de $\text{prox}_{\gamma|\cdot|}$ com $\gamma = 1/2$.

A proposição abaixo será utilizada na Seção 4.2.

Proposição 3.4.2. *Considere $(e_k)_{k \in \mathbb{N}}$ uma base ortonormal de \mathcal{H} . Sejam $(\phi_k)_{k \in \mathbb{N}}$ funções em $\Gamma_0(\mathbb{R})$ tais que*

$$(\forall k \in \mathbb{N}) \quad \phi_k \geq 0 \quad e \quad \phi_k(0) = 0, \quad (3.30)$$

e seja $\psi : \mathcal{H} \rightarrow (-\infty, +\infty] : x \mapsto \sum_{k \in \mathbb{N}} \phi(\langle x | e_k \rangle)$. Então:

(i) $\psi \in \Gamma_0(\mathcal{H})$.

(ii) $(\forall x \in \mathcal{H}) \quad \text{prox}_{\psi} x = \sum_{k \in \mathbb{N}} (\text{prox}_{\phi_k} \langle x | e_k \rangle) e_k$.

Demonstração. Para verificar este exemplo, vamos introduzir um operador

$$L : \mathcal{H} \rightarrow \ell^2(\mathbb{N}) : x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}}$$

e uma função

$$\varphi : \ell^2(\mathbb{N}) \rightarrow (-\infty, +\infty] : (x_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \phi_k(x_k).$$

Note que L está bem definido, pois usando a desigualdade de Bessel temos que $\sum_{k \in \mathbb{N}} |\langle x | e_k \rangle|^2 \leq \|x\|^2 < +\infty$, logo $(\langle x | e_k \rangle)_{k \in \mathbb{N}}$ está em $\ell^2(\mathbb{N})$ para todo x em \mathcal{H} . Além disso, L é um operador linear, limitado e inversível com

$$L^{-1} = L^* : \ell^2(\mathbb{N}) \rightarrow \mathcal{H} : (x_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} x_k e_k. \quad (3.31)$$

(i) Observe que $\psi = \varphi \circ L$. Para mostrar que ψ está em $\Gamma_0(\mathcal{H})$, pelas

propriedades de L , é suficiente mostrar que φ está em $\Gamma_0(\ell^2(\mathbb{N}))$. Para isto, defina

$$(\forall M \in \mathbb{N}) \quad \varphi_M((x_l)_{l \in \mathbb{N}}) = \sum_{k=0}^M \phi_k(x_k).$$

Então, a função φ_M é s.c.i. e convexa em $\Gamma_0(\ell^2(\mathbb{N}))$, pois é uma soma finita de tais funções. Consequentemente, os conjuntos $(\text{epi}\varphi_M)_{M \in \mathbb{N}}$ são fechados e convexos em $\ell^2(\mathbb{N}) \times \mathbb{R}$. Por hipótese temos que as funções $(\varphi_M)_{M \in \mathbb{N}}$ são não-negativas, então $\varphi = \sup_{M \in \mathbb{N}} \varphi_M$. Além disso, usando (1.6) temos que

$$\begin{aligned} \text{epi}\varphi &= \{(x, \lambda) \in \ell^2(\mathbb{N}) \times \mathbb{R} \mid \varphi(x) \leq \lambda\} \\ &= \left\{ (x, \lambda) \in \ell^2(\mathbb{N}) \times \mathbb{R} \mid \sup_{M \in \mathbb{N}} \varphi_M(x) \leq \lambda \right\} \\ &= \{(x, \lambda) \in \ell^2(\mathbb{N}) \times \mathbb{R} \mid \varphi_M(x) \leq \lambda, \forall M \in \mathbb{N}\} \\ &= \bigcap_{M \in \mathbb{N}} \text{epi}\varphi_M. \end{aligned}$$

Portanto, o conjunto $\text{epi}\varphi$ é fechado e convexo como uma interseção de conjuntos fechados e convexos. Com isto concluímos que φ é s.c.i. e convexa. Ainda, φ é própria, pois por hipótese $\phi_k(0) = 0$ para todo $k \in \mathbb{N}$, então $\varphi(0) = 0 < +\infty$ que implica $\text{dom}\varphi \neq \emptyset$. Assim, segue o resultado.

(ii) Fixe $(x_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})$. Defina $p = \text{prox}_{\varphi} x$ e $q = (\pi_k)_{k \in \mathbb{N}}$, onde $\pi_k = \text{prox}_{\phi_k} x_k$ para todo $k \in \mathbb{N}$. Então, tendo em vista a Proposição 3.2.2, devemos mostrar que $p = q$. Primeiro observe que, para todo $k \in \mathbb{N}$, (3.30) implica que 0 minimiza ϕ_k , ou seja, que $0 \in \partial\phi_k(0)$ por (1.20), logo $\text{prox}_{\phi_k} 0 = 0$ por (3.3). Consequentemente, segue da não-expansividade dos operadores $(\text{prox}_{\phi_k})_{k \in \mathbb{N}}$ vista na Proposição 3.1.3 que

$$\sum_{k \in \mathbb{N}} |\pi_k|^2 = \sum_{k \in \mathbb{N}} |\text{prox}_{\phi_k} x_k - \text{prox}_{\phi_k} 0|^2 \leq \sum_{k \in \mathbb{N}} |x_k - 0|^2 = \|x\|^2 < +\infty.$$

Disto $q \in \ell^2(\mathbb{N})$. Agora, seja $y = (y_k)_{k \in \mathbb{N}}$ um ponto arbitrário em $\ell^2(\mathbb{N})$. Segue de (3.3) e (1.18) que p é o único ponto em $\ell^2(\mathbb{N})$ que satisfaz

$$\langle y - p \mid x - p \rangle + \varphi(p) \leq \varphi(y).$$

Por outro lado, usando a mesma caracterização para cada ponto em

$(\pi_k)_{k \in \mathbb{N}}$ temos que

$$(\forall k \in \mathbb{N})(y_k - \pi_k)(x_k - \pi_k) + \phi_k(\pi_k) \leq \phi_k(y_k).$$

Tomando o somatório sobre $k \in \mathbb{N}$ nesta última desigualdade, obtemos

$$\sum_{k \in \mathbb{N}} (y_k - \pi_k)(x_k - \pi_k) + \sum_{k \in \mathbb{N}} \phi_k(\pi_k) \leq \sum_{k \in \mathbb{N}} \phi_k(y_k),$$

isto é,

$$\langle y - q | x - q \rangle + \varphi(q) \leq \varphi(y).$$

Mas, vimos que p é o único ponto em $\ell^2(\mathbb{N})$ que satisfaz isto, logo $p = q$. Portanto, pela Proposição 3.2.2 e por (3.31), temos que

$$\begin{aligned} (\forall x \in \mathcal{H}) \quad \text{prox}_{\psi} x &= L^*(\text{prox}_{\varphi}(Lx)) \\ &= L^*(\text{prox}_{\varphi}(\langle x | e_k \rangle)_{k \in \mathbb{N}}) \\ &= L^*(\text{prox}_{\phi_k} \langle x | e_k \rangle)_{k \in \mathbb{N}} \\ &= \sum_{k \in \mathbb{N}} (\text{prox}_{\phi_k} \langle x | e_k \rangle) e_k. \end{aligned}$$

■

O seguinte exemplo é um caso especial da Proposição 3.4.2.

Exemplo 3.4.7. Considere $(e_k)_{k \in \mathbb{N}}$ uma base ortonormal de \mathcal{H} . Sejam (w_k) uma seqüência em \mathbb{R}_{++} , $\psi : \mathcal{H} \rightarrow (-\infty, +\infty] : x \mapsto \sum_{k \in \mathbb{N}} w_k |\langle x | e_k \rangle|$ e $x \in \mathcal{H}$. Então,

$$\text{prox}_{\psi} x = \sum_{k \in \mathbb{N}} \pi_k e_k,$$

onde

$$(\forall k \in \mathbb{N}) \quad \pi_k = \text{sign}(\langle x | e_k \rangle) \max\{|\langle x | e_k \rangle| - w_k, 0\}.$$

De fato, basta tomar $\phi_k = w_k |\cdot|$ na Proposição 3.4.2 e usar (3.29).

4 MINIMIZANDO A SOMA DE DUAS FUNÇÕES CONVEXAS

Minimizar uma função pode ser extremamente difícil quando não conhecemos suas propriedades de regularidade, como a diferenciabilidade, por exemplo. Mas, este trabalho pode se tornar mais fácil se conseguirmos decompor a função em uma soma de duas funções cujas propriedades de regularidade são conhecidas.

Neste capítulo, consideramos o problema de minimizar a soma de duas funções, que surge no seguinte formato:

$$(P) \quad \underset{x \in \mathcal{H}}{\text{minimize}} \quad f_1(x) + f_2(x),$$

onde, $f_1 : \mathcal{H} \rightarrow (-\infty, +\infty]$ e $f_2 : \mathcal{H} \rightarrow \mathbb{R}$ são duas funções em $\Gamma_0(\mathcal{H})$ tal que f_2 é diferenciável em \mathcal{H} com gradiente $(1/\beta)$ -Lipschitz contínuo, para algum $\beta \in \mathbb{R}_{++}$. Denotamos o conjunto das soluções deste problema por G . Investigamos propriedades e solução numérica do Problema (P) e apresentamos uma aplicação em recuperação de sinais.

4.1 ALGORITMO FORWARD-BACKWARD

Iniciamos esta seção estudando propriedades como existência, unicidade e caracterização de soluções para o Problema (P), em seguida analisamos a convergência de um algoritmo forward-backward para resolvê-lo.

Proposição 4.1.1.

- (i) *Existência:* O Problema (P) possui pelo menos uma solução se $f_1 + f_2$ é coerciva.
- (ii) *Unicidade:* O Problema (P) possui no máximo uma solução se $f_1 + f_2$ é estritamente convexa.
- (iii) *Caracterização:* Sejam $x \in \mathcal{H}$ e $\gamma \in \mathbb{R}_{++}$. Então, são equivalentes:
 - (a) $x \in G$, isto é, x é solução de (P).
 - (b) $x = \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x))$.

Demonstração.

(i) e (ii) A função $f_1 + f_2$ é soma de duas funções convexas, s.c.i. e próprias. Como $\text{dom} f_2 = \mathcal{H}$, tem-se

$$\text{dom} f_1 \cap \text{dom} f_2 = \text{dom} f_1 \cap \mathcal{H} = \text{dom} f_1 \neq \emptyset.$$

Logo, a função $f_1 + f_2$ está em $\Gamma_0(\mathcal{H})$. Portanto, as afirmações (i) e (ii) seguem do Teorema 1.2.1.

(iii) Segue de (1.20), Proposição 1.5.3, Lema 1.5.1 e (3.3) que

$$\begin{aligned} x \in G &\Leftrightarrow 0 \in \partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) = \partial f_1(x) + \{\nabla f_2(x)\} \\ &\Leftrightarrow -\nabla f_2(x) \in \partial f_1(x) \\ &\Leftrightarrow -\gamma \nabla f_2(x) \in \gamma \partial f_1(x) \\ &\Leftrightarrow (x - \gamma \nabla f_2(x)) - x \in \gamma \partial f_1(x) \\ &\Leftrightarrow x = \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x)). \end{aligned} \tag{4.1}$$

■

A caracterização de ponto fixo fornecida pela Proposição 4.1.1 (iii)(b) sugere resolver o Problema (P) através da iteração de ponto fixo

$$x_{n+1} = \text{prox}_{\gamma f_1}(x_n - \gamma \nabla f_2(x_n)) \tag{4.2}$$

para um valor conveniente do parâmetro γ .

Esta iteração, que é chamada em otimização de algoritmo forward-backward, consiste em dois passos. Primeiro, realiza um passo explícito envolvendo somente f_2 para calcular $x_{n+1/2} = x_n - \gamma \nabla f_2(x_n)$, depois realiza um passo implícito envolvendo somente f_1 para calcular $x_{n+1} = \text{prox}_{\gamma f_1} x_{n+1/2}$. Formalmente, este segundo passo equivale a resolver a inclusão

$$x_{n+1} = \text{prox}_{\gamma f_1} x_{n+1/2} \Leftrightarrow x_{n+1/2} - x_{n+1} \in \gamma \partial f(x_{n+1}),$$

por isso sua natureza implícita.

Notemos que, por um lado, quando $f_1 = 0$, (4.2) se reduz ao método do gradiente

$$x_{n+1} = x_n - \gamma \nabla f_2(x_n)$$

para minimizar uma função com gradiente Lipschitz contínuo. Por outro lado, quando $f_2 = 0$, (4.2) se reduz ao chamado algoritmo de

ponto proximal (ver [3, Teorema 27.1])

$$x_{n+1} = \text{prox}_{\gamma f_1} x_n$$

para minimizar uma função não diferenciável. O algoritmo forward-backward pode, portanto, ser considerado uma combinação desses dois métodos básicos.

O próximo teorema fornece uma iteração um pouco mais geral em que é introduzida uma sequência de relaxação (λ_n) , que pode ser usada para melhorar a taxa de convergência na implementação numérica do algoritmo.

Considere f_1, f_2, β e G conforme definidos no Problema (P).

Teorema 4.1.1 (Forward-Backward). *Suponha que $G \neq \emptyset$. Sejam $\gamma \in (0, 2\beta)$ e $\delta = 1/2 + \min\{1, \beta/\gamma\}$. Considere (λ_n) uma sequência tal que $0 < \delta\lambda \leq \lambda_n \leq \delta\bar{\lambda} < 1$ para todo $n \in \mathbb{N}$. Fixe $x_0 \in \mathcal{H}$ e defina*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \left(\text{prox}_{\gamma f_1} (x_n - \gamma \nabla f_2(x_n)) - x_n \right). \quad (4.3)$$

Então, a sequência (x_n) converge fracamente para um ponto x em G .

Demonstração. Defina $T = \text{prox}_{\gamma f_1} \circ (I - \gamma \nabla f_2)$. Então, a expressão (4.3) se reduz a

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n).$$

Usando (4.1) temos que

$$x \in G \Leftrightarrow x = \text{prox}_{\gamma f_1} (I - \gamma \nabla f_2)(x) = Tx \Leftrightarrow x \in \text{Fix } T.$$

Logo, $G = \text{Fix } T$. Assim, basta verificar as hipóteses do Teorema 2.4.4 para concluir que (x_n) converge fracamente para um ponto x em G .

Observe que $\text{Fix } T \neq \emptyset$, pois $G \neq \emptyset$, por hipótese.

Sabemos que $\text{prox}_{\gamma f_1}$ é firmemente não-expansivo, pela Proposição 3.1.3, e assim é $1/2$ -ponderado, pela Proposição 2.2.1(iii). Ainda, como f_2 é diferenciável em \mathcal{H} e ∇f_2 é $1/\beta$ -Lipschitz contínuo, por hipótese, então ∇f_2 é β -cocoercivo, pela Proposição 2.1.2. Logo, $I - \gamma \nabla f_2$ é $\gamma/2\beta$ -ponderado, pela Proposição 2.2.4.

Segue pela Proposição 2.2.3 que T é α -ponderado, onde

$$\alpha = \frac{2 \max\{1/2, \gamma/2\beta\}}{1 + \max\{1/2, \gamma/2\beta\}}$$

que é equivalente a

$$\alpha = \frac{2}{1 + \min\{2, 2\beta/\gamma\}} = \frac{1}{1/2 + \min\{1, \beta/\gamma\}} = \frac{1}{\bar{\delta}},$$

e assim, para todo $n \in \mathbb{N}$, temos que $0 < \underline{\lambda}/\alpha \leq \lambda_n \leq \bar{\lambda}/\alpha < 1$.

Portanto, pelo Teorema 2.4.4, a sequência (x_n) converge fracamente para um ponto x em G . ■

4.2 APLICAÇÃO EM PROBLEMAS DE RECUPERAÇÃO DE SINAIS

Um problema de recuperação de sinais é descrito como o problema de deduzir um sinal \bar{x} em um espaço de Hilbert \mathcal{H} das observações de um sinal z em um espaço de Hilbert \mathcal{Z} . Por exemplo, na restauração de imagens [1], o objetivo é recuperar a forma original de uma imagem \bar{x} pela observação de uma versão borrada z , e portanto $\mathcal{H} = \mathcal{Z}$. Por outro lado, em reconstrução de sinais, os dados z são indiretamente relacionados a \bar{x} e portanto \mathcal{H} e \mathcal{Z} são frequentemente espaços diferentes. Assim, em tomografia [14], um sinal deve ser recuperado a partir de uma coleção de medições de sinais com dimensões menores.

Nesta seção, consideramos o modelo de formação dos dados em que z está relacionado com \bar{x} da seguinte forma

$$z = T\bar{x} + w,$$

onde $T : \mathcal{H} \rightarrow \mathcal{Z}$ é um operador linear limitado e $w \in \mathcal{Z}$ representa uma perturbação. Este modelo abrange numerosos problemas de reconstrução de sinais e imagens (ver [1, 14, 24, 25]).

O problema em questão é o seguinte:

$$(P1) \quad \underset{x \in \mathcal{H}}{\text{minimize}} \quad f(Lx) + \frac{1}{2} \|Tx - z\|^2,$$

onde

- (i) \mathcal{K} é um espaço de Hilbert real;
- (ii) $T : \mathcal{H} \rightarrow \mathcal{Z}$ é um operador linear limitado;
- (iii) $L : \mathcal{H} \rightarrow \mathcal{K}$ é um operador linear limitado e bijetivo tal que $L^{-1} = L^*$;

(iv) $f \in \Gamma_0(\mathcal{K})$.

O conjunto das soluções deste problema é denotado por G .

No Problema (P1), o termo $\|Tx - z\|^2/2$ é chamado termo de fidelidade dos dados que tenta refletir a contribuição do modelo de formação, enquanto o termo $f(Lx)$ fornece alguma informação a-priori sobre o sinal original \bar{x} .

Proposição 4.2.1.

(i) O Problema (P1) possui pelo menos uma solução se f é coerciva.

(ii) O Problema (P1) possui no máximo uma solução se uma das seguintes condições são satisfeitas:

(a) f é estritamente convexa.

(b) T é injetivo.

(iii) O Problema (P1) possui exatamente uma solução se T é limitado inferiormente, isto é, existe um $\kappa \in \mathbb{R}_{++}$ tal que

$$(\forall x \in \mathcal{H}) \quad \|Tx\| \geq \kappa\|x\|.$$

(iv) Sejam $x \in \mathcal{H}$ e $\gamma \in \mathbb{R}_{++}$. Então, são equivalentes:

(a) $x \in G$, isto é, x é solução de (P1).

(b) $x = (L^* \circ \text{prox}_{\gamma f} \circ L)(x - \gamma T^*(z - Tx))$.

Demonstração. Defina $f_1 = f \circ L$ e $f_2 : x \mapsto \|Tx - z\|^2/2$. Então, segue das hipóteses (i)-(iv) do Problema (P1) que f_1 e f_2 estão em $\Gamma_0(\mathcal{H})$ e que f_2 é diferenciável com $\nabla f_2 : x \mapsto T^*(Tx - z)$. Consequentemente, para todo $x, y \in \mathcal{H}$,

$$\begin{aligned} \|\nabla f_2(x) - \nabla f_2(y)\| &= \|T^*(Tx - z) - T^*(Ty - z)\| \\ &= \|T^*T(x - y)\| \\ &= \|T\|^2\|x - y\|, \end{aligned}$$

isto é, ∇f_2 é Lipschitz contínuo com constante $\|T\|^2$. Portanto, o Problema (P1) é um caso especial do Problema (P) com $\beta = 1/\|T\|^2$.

(i) Tendo em vista a Proposição 4.1.1(i), é suficiente mostrar que $f_1 + f_2$ é coerciva. Temos que

$$f_1 + f_2 = f \circ L + \frac{\|Tx - z\|^2}{2} \geq f \circ L.$$

Além disso, como f é coerciva e L é linear limitado, segue que $f \circ L$ é coerciva. Logo,

$$(f_1 + f_2)(x) \geq (f \circ L)(x) \rightarrow +\infty$$

quando $\|x\| \rightarrow +\infty$, ou seja, $f_1 + f_2$ é coerciva.

(ii) Pela Proposição 4.1.1(ii), devemos mostrar que $f_1 + f_2$ é estritamente convexa, para isto é suficiente mostrar que f_1 ou f_2 é estritamente convexa. (a) Como f é estritamente convexa e L é injetivo, segue que f_1 é estritamente convexa. (b) Pelo Exemplo 1.1.1, temos que $\|\cdot - z\|^2/2$ é estritamente convexa, então f_2 é estritamente convexa pela injetividade de T .

(iii) Seja $x \in \mathcal{H}$. Vamos mostrar que T é injetivo, isto é, se $Tx = 0$ então $x = 0$. Suponha que $Tx = 0$. Como T é limitado inferiormente, temos

$$\kappa\|x\| \leq \|Tx\| = 0,$$

ou seja, $x = 0$. Portanto, por (ii)(b), existe no máximo uma solução. Quanto a existência, pela Proposição 4.1.1(i), devemos mostrar que $f_1 + f_2$ é coerciva. Uma vez que $f \in \Gamma_0(\mathcal{K})$, pelo Lema 1.1.1, ela é limitada inferiormente por uma função afim, digamos $\langle \cdot | u \rangle + \eta/2$, onde $u \in \mathcal{K} \setminus \{0\}$ e $\eta \in \mathbb{R}$. Disso, usando a desigualdade de Cauchy-Schwarz e o fato que T é limitado inferiormente, obtemos que, para todo $x \in \mathcal{H}$,

$$\begin{aligned} 2(f_1(x) + f_2(x)) &\geq 2\langle Lx | u \rangle + \eta + \|Tx - z\|^2 \\ &= 2\langle x | L^*u \rangle + \eta + \|Tx\|^2 - 2\langle Tx | z \rangle + \|z\|^2 \\ &= 2\langle x | L^*u \rangle + \eta + \|Tx\|^2 - 2\langle x | T^*z \rangle + \|z\|^2 \\ &= 2\langle x | L^*u - T^*z \rangle + \|Tx\|^2 + \|z\|^2 + \eta \\ &= \|x + L^*u - T^*z\|^2 + (\|Tx\|^2 - \|x\|^2) \\ &\quad - \|L^*u - T^*z\|^2 + \|z\|^2 + \eta \\ &\geq (\|x\| - \|L^*u - T^*z\|)^2 + (\kappa^2 - 1)\|x\|^2 \\ &\quad - \|L^*u - T^*z\|^2 + \|z\|^2 + \eta \\ &\geq (\kappa\|x\| - \|L^*u - T^*z\|/\kappa)^2 - \|L^*u - T^*z\|^2/\kappa^2 \\ &\quad + \|z\|^2 + \eta \\ &\rightarrow +\infty \end{aligned}$$

quando $\|x\| \rightarrow +\infty$. Portanto, $f_1 + f_2$ é coerciva.

(iv) Usando a Proposição 4.1.1(iii) e a Proposição 3.2.2,

$$\begin{aligned} x &= \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x)) \\ &= (L^* \circ \text{prox}_{\gamma f} \circ L)(x + \gamma T^*(z - Tx)). \end{aligned}$$

■

No contexto do Problema (P1), o algoritmo forward-backward (4.3) assume a seguinte forma.

Teorema 4.2.1. *Suponha que $G \neq \emptyset$. Sejam $\gamma \in (0, 2/\|T\|^2)$ e $\delta = 1/2 + \min\{1, 1/(\gamma\|T\|^2)\}$. Considere (λ_n) uma sequência tal que $0 < \delta \underline{\lambda} \leq \lambda_n \leq \delta \bar{\lambda} < 1$ para todo $n \in \mathbb{N}$. Fixe $x_0 \in \mathcal{H}$ e defina, para todo $n \in \mathbb{N}$,*

$$x_{n+1} = x_n + \lambda_n \left((L^* \circ \text{prox}_{\gamma f} \circ L)(x_n - \gamma T^*(Tx_n - z)) - x_n \right). \quad (4.4)$$

Então, a sequência (x_n) converge fracamente para um ponto x em G .

Demonstração. Tendo em vista a demonstração da Proposição 4.2.1, segue que (4.4) é um caso especial de (4.3) com $\text{prox}_{\gamma f_1} = L^* \circ \text{prox}_{\gamma f} \circ L$, $\nabla f_2 : x \mapsto T^*(Tx - z)$ e $\beta = 1/\|T\|^2$. Logo, pelo Teorema 4.1.1 a sequência (x_n) converge fracamente para um ponto x em G . ■

Um caso notável do Problema (P1) é quando $\mathcal{K} = \mathcal{H}$, $L = I$ e $f = \iota_C$, com C um subconjunto não-vazio, fechado e convexo de \mathcal{H} . Este caso é conhecido como problema de mínimos quadrados restrito, onde buscamos uma solução para

$$(P2) \quad \underset{x \in C}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2,$$

que, pela Proposição 4.2.1 e pelo Exemplo 3.1.3, é caracterizada por $x = P_C(x - \gamma T^*(z - Tx))$.

Obtemos o seguinte corolário direto do Teorema 4.2.1, usado para encontrar soluções para o Problema (P2).

Corolário 4.2.1. *Suponha que $G \neq \emptyset$. Sejam $\gamma \in (0, 2/\|T\|^2)$ e $\delta = 1/2 + \min\{1, 1/(\gamma\|T\|^2)\}$. Considere (λ_n) uma sequência tal que $0 < \delta \underline{\lambda} \leq \lambda_n \leq \delta \bar{\lambda} < 1$ para todo $n \in \mathbb{N}$. Fixe $x_0 \in \mathcal{H}$ e defina, para*

todo $n \in \mathbb{N}$,

$$x_{n+1} = x_n + \lambda_n \left(P_C(x_n - \gamma T^*(Tx_n - z)) - x_n \right). \quad (4.5)$$

Então, a sequência (x_n) converge fracamente para um ponto x em G .

Demonstração. Segue do Teorema 4.2.1 e do Exemplo 3.1.3. ■

Uma idéia interessante em recuperação de sinais é decompor uma função em uma base ortonormal e transformar os coeficientes da decomposição para construir aproximações ou estimativas do sinal original (ver [5, 6, 12, 17]). Uma formulação deste conceito, conhecido por problema de regularização esparsa, é a seguinte:

$$(P3) \quad \underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2 + \sum_{k \in \mathbb{N}} \phi_k(\langle x | e_k \rangle),$$

onde

- (i) $T : \mathcal{H} \rightarrow \mathcal{Z}$ é um operador linear limitado;
- (ii) $(e_k)_{k \in \mathbb{N}}$ é uma base ortonormal de \mathcal{H} ;
- (iii) $(\phi_k)_{k \in \mathbb{N}}$ são funções em $\Gamma_0(\mathbb{R})$ tal que $\phi_k \geq 0$ e $\phi_k(0) = 0$, para todo $n \in \mathbb{N}$.

Esse tipo de formulação surge em problemas de recuperação de sinal em que z é o sinal observado e o sinal original tem uma representação na base ortonormal $(e_k)_{k \in \mathbb{N}}$.

Como podemos ver na demonstração da Proposição 3.4.2, o Problema (P3) é um caso especial do Problema (P1) com $\mathcal{K} = \ell^2(\mathbb{N})$, $L : x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}}$ e $f : (x_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \phi_k(x_k)$.

Segue da Proposição 4.2.1 e da Proposição 3.4.2, a seguinte caracterização para as soluções do Problema (P3):

$$(\forall k \in \mathbb{N}) \quad \langle x | e_k \rangle = \text{prox}_{\gamma \phi_k}(x - \gamma T^*(Tx - z) | e_k).$$

Por exemplo, considerando $T = I$ e $\gamma = 1$. Definindo $\phi_k = w_k |\cdot|$ para todo $k \in \mathbb{N}$, onde (w_k) é uma sequência em \mathbb{R}_{++} , temos

$$f \circ L : x \mapsto \sum_{k \in \mathbb{N}} w_k |\langle x | e_k \rangle|.$$

Neste caso, pela Proposição 4.2.1 e pelo Exemplo 3.4.7, obtemos a seguinte caracterização para a solução:

$$x = \text{prox}_{f \circ L} z = \sum_{k \in \mathbb{N}} \pi_k e_k,$$

onde

$$(\forall k \in \mathbb{N}) \quad \pi_k = \text{sign}(\langle z | e_k \rangle) \max\{|\langle z | e_k \rangle| - w_k, 0\}.$$

Quanto a solução numérica do Problema (P3), temos o seguinte corolário.

Corolário 4.2.2. *Suponha que $G \neq \emptyset$. Sejam $\gamma \in (0, 2/\|T\|^2)$ e $\delta = 1/2 + \min\{1, 1/(\gamma\|T\|^2)\}$. Considere (λ_n) uma sequência tal que $0 < \delta\lambda \leq \lambda_n \leq \delta\bar{\lambda} < 1$ para todo $n \in \mathbb{N}$. Fixe $x_0 \in \mathcal{H}$ e defina, para todo $n \in \mathbb{N}$,*

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{N}} \text{prox}_{\gamma\phi_k} \langle x_n - \gamma T^*(Tx_n - z) | e_k \rangle e_k - x_n \right). \quad (4.6)$$

Então, a sequência (x_n) converge fracamente para um ponto x em G .

Demonstração. Segue do Teorema 4.2.1, usando a Proposição 3.2.2 e a Proposição 3.4.2. ■

CONSIDERAÇÕES FINAIS

Neste trabalho, apresentamos a estrutura matemática por trás do problema de minimizar a soma de duas funções convexas, próprias e semicontínuas inferiormente, quando uma delas é diferenciável. Vimos que a solução deste problema pode ser caracterizada por equações de ponto fixo envolvendo operadores proximais. Uma vez que operadores proximais são firmemente não-expansivos, eles podem ser usados para elaborar o algoritmo forward-backward, que converge fracamente para a solução do problema.

A importância deste estudo está no fato de que esta formulação do problema torna possível obter existência, unicidade e caracterização dos resultados de uma maneira uniforme e padrão para uma classe de problemas aparentemente diferentes e fornecer uma análise simplificada para uma variedade de métodos iterativos existentes, como podemos ver em [11].

Para estudos futuros, podemos investigar a convergência forte do algoritmo forward-backward e aperfeiçoá-lo (ver [8]), explorando a teoria geral de operadores monótonos. Uma iteração mais geral pode ser fornecida em que o coeficiente γ dependa da iteração e erros sejam permitidos na avaliação dos operadores $\text{prox}_{\gamma f_1}$ e ∇f_2 . Os erros permitem uma tolerância na implementação numérica do algoritmo, enquanto a flexibilidade introduzida pelo parâmetro γ pode ser usada para melhorar sua taxa de convergência.

REFERÊNCIAS

- [1] ANDREWS H. C.; HUNT, B. R. *Digital Image Restoration*. NJ: Prentice-Hall, Englewood Cliffs, 1977.
- [2] AUJOL, J.-F. et al. Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vision*, v. 22, p. 71–88, 2005.
- [3] BAUSCHKE H. H.; COMBETTES, P. L. *Convex analysis and monotone operator theory in Hilbert spaces*. New York: Springer, 2011.
- [4] BREZIS, H. *Functional analysis, Sobolev spaces and partial differential equations*. New York: Springer, 2011.
- [5] CHAMBOLLE, A. et al. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Trans. Image Process.*, v. 7, p. 319–335, 1998.
- [6] COHEN, A. *Numerical analysis of wavelet methods*. Amsterdam: North-Holland Publishing Co., 2003.
- [7] COMBETTES, P. L. Inconsistent signal feasibility problems: Least-squares solutions in a product space. *IEEE Trans. Signal Process.*, v. 42, p. 2955–2966, 1994.
- [8] COMBETTES, P. L. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, v. 53, p. 475–504, 2004.
- [9] COMBETTES P. L.; BONDON, P. Hard-constrained inconsistent signal feasibility problems. *IEEE Trans. Signal Process.*, v. 47, p. 2460–2468, 1999.
- [10] COMBETTES P. L.; PESQUET, J.-C. Wavelet-constrained image restoration. *Int. J. Wavelets Multiresolut. Inf. Process.*, v. 2, p. 371–389, 2004.
- [11] COMBETTES P. L.; WAJS, V. R. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, p. 1168–1200 (electronic), 2005.

- [12] DAUBECHIES, I.; DEFRISE, M.; MOL, C. D. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, v. 57, p. 1413–1457, 2004.
- [13] EKELAND I.; TÉMAM, R. *Convex analysis and variational problems*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1999.
- [14] HERMAN, G. T. *Image reconstruction from projections, the fundamentals of computerized tomography*. New York: Academic Press Inc., 1980.
- [15] LANNES, A.; ROQUES, S.; CASANOVE, M. J. Stabilized reconstruction in signal and image processing. part 1: Partial deconvolution and spectral extrapolation with limited field. *J. Modern Opt.*, v. 34, p. 161–226, 1987.
- [16] KREYSZIG, E. *Introductory functional analysis with applications*. New York-London-Sydney: John Wiley & Sons, 1978.
- [17] MOL C.; DEFRISE, M. D. *A note on wavelet-based inversion algorithms*. Providence, RI: Amer. Math. Soc., 2002.
- [18] MOREAU, J. J. Fonctions convexes duales et points proximaux dans un espace hilbertien. *C. R. Acad. Sci. Paris*, v. 255, p. 2897–2899, 1962.
- [19] MOREAU, J. J. Propriétés des applications ‘prox’. *C. R. Acad. Sci. Paris*, v. 256, p. 1069–1071, 1963.
- [20] MOREAU, J. J. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, v. 93, p. 273–299, 1965.
- [21] NESTEROV, Y. *Introductory lectures on convex optimization*. Boston: Kluwer Academic Publishers, 2004.
- [22] PESQUET J.-C.; COMBETTES, P. L. Wavelet synthesis by alternating projections. *IEEE Trans. Signal Process.*, v. 44, p. 728–732, 1996.
- [23] SABHARWAL A.; POTTER, L. C. Convexly constrained linear inverse problems: iterative least-squares and regularization. *IEEE Trans. Signal Process.*, v. 46, p. 906–916, 1998.
- [24] STARK, H. *Image recovery: theory and application*. Orlando: Academic Press Inc., 1987.

[25] STARK H.; YANG, Y. *Vector Space Projections : A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics*. New York: Wiley, 1998.