



UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

RUI MARQUES CARVALHO

Uma Abordagem Quasi-Newton às Equações Generalizadas com Restrições

Campinas

2021

Rui Marques Carvalho

Uma Abordagem Quasi-Newton às equações Generalizadas com Restrições

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Matemática Aplicada.

Orientador: Roberto Andreani

Coorientador: Gilson do Nascimento Silva

Este trabalho corresponde à versão final da Tese defendida pelo aluno Rui Marques Carvalho e orientada pelo Prof. Dr. Roberto Andreani.

Campinas

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

C253a Carvalho, Rui Marques, 1990-
Uma abordagem quasi-Newton às equações generalizadas com restrições /
Rui Marques Carvalho. – Campinas, SP : [s.n.], 2021.

Orientador: Roberto Andreani.

Coorientador: Gilson do Nascimento Silva.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de
Matemática, Estatística e Computação Científica.

1. Equações generalizadas com restrições. 2. Método quase-Newton. 3.
Projeção inexata (Otimização). 4. Atualização de Broyden. I. Andreani,
Roberto, 1961-. II. Silva, Gilson do Nascimento. III. Universidade Estadual de
Campinas. Instituto de Matemática, Estatística e Computação Científica. IV.
Título.

Informações para Biblioteca Digital

Título em outro idioma: A quasi-Newton approach to generalized constrained equations

Palavras-chave em inglês:

Constrained generalized equations

Quasi-Newton method

Inexact projection (Optimization)

Broyden updates

Área de concentração: Matemática Aplicada

Titulação: Doutor em Matemática Aplicada

Banca examinadora:

Roberto Andreani [Orientador]

Paulo Sergio Marques dos Santos

Orizon Pereira Ferreira

Leonardo Delarmelina Secchin

Estevão Esmi Laureano

Data de defesa: 19-02-2021

Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-2151-8733>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4212734609363457>

**Tese de Doutorado defendida em 19 de fevereiro de 2021 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). ROBERTO ANDREANI

Prof(a). Dr(a). PAULO SERGIO MARQUES DOS SANTOS

Prof(a). Dr(a). ORIZON PEREIRA FERREIRA

Prof(a). Dr(a). LEONARDO DELARMELINA SECCHIN

Prof(a). Dr(a). ESTEVÃO ESMI LAUREANO

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Inside, so deep inside... You will die if you don't dream anymore.

Agradecimentos

Agradeço a Deus por me proporcionar mais uma conquista.

Agradeço a minha esposa e companheira Marina Rocha, por me acompanhar nessa jornada tão difícil.

Agradeço aos meus Pais e Avós pelas ligações e acompanhamentos diários durante essa jornada.

Agradeço também a toda minha família que me acompanhou durante todos os momentos, em especial aos que vieram até Campinas.

Agradeço aos poucos e bons amigos que fiz na Unicamp.

Agradeço a todos os meus amigos do Piauí que sempre acompanharam minha jornada torcendo a cada conquista.

Agradeço ao IFPI pelo financiamento.

Resumo

Equações Generalizadas vêm sendo estudadas por muitos matemáticos há várias décadas, mas o estudo do problema com restrições é bem recente. Nesta Tese, faremos uma abordagem Quasi-Newton às Equações Generalizadas com Restrições. Essa abordagem foi feita de duas formas diferentes, primeiro o Método Quasi-Newton com Projeção Inexata, em seguida o Método Quasi-Newton Inexato com Projeção Inexata. De início, provamos os teoremas exigindo apenas uma condição para a atualização B_k e depois, para ambos os casos, enunciamos um novo resultado usando a atualização de Broyden. Alguns testes numéricos foram feitos.

Palavras-chave: Equações Generalizadas com Restrições. Quasi-Newton. Projeção Inexata. Broyden.

Abstract

Generalized equations have been studied by many mathematicians for several decades, however the study of the constrained problem is quite recent. In this Thesis, we will use the Quasi-Newton method for Constrained Generalized Equations. This approach was developed in two different ways, firstly using the Quasi Newton Method with Inexact Projection, secondly using the Quasi Newton Inexact Method with Inexact Projection. At first, we proved the theorems by requiring only one condition for the Bk update, and then, for both cases, we stated a new result using the Broyden update. Some numerical tests were made.

Keywords: Constrained Generalized Equations. Quasi-Newton. Inexact Projection. Broyden.

Lista de ilustrações

Figura 1 – Cone Normal	17
Figura 2 – Cone Dual	17
Figura 3 – O gráfico de $x \mapsto S(x) + B_{l_{\ x'-x\ }}(0)$	22
Figura 4 – Satisfaz Aubin.	22
Figura 5 – Não satisfaz Aubin.	22
Figura 6 – Projecção Factível Inexata	25

Lista de abreviaturas e siglas

KKT	<i>Karush-Kuhn-Tucker</i>
EQ	<i>Equação Generalizada</i>
EQR	<i>Equação Generalizada com Restrições</i>
PNL	<i>Problema de Programação Não linear</i>
CVIP	<i>Constrained Variational Inequality Problem</i>
SVIP	<i>Split Variational Inequality Problem</i>

Lista de símbolos

$\ \cdot \ $	Norma Euclidiana ou a norma matricial induzida
$N_C(x^*)$	Cone Normal em x^* com relação a C
K^*	Cone Dual do cone K
$\langle x, y \rangle$	Produto interno usual entre $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$
$B_\delta(x)$	Bola aberta de centro x e raio δ
$B_\delta[x]$	Bola fechada de centro x e raio δ
$d(x, D)$	Distância do ponto x ao conjunto D
$e(C, D)$	Excesso de C além de D .
$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$	O espaço das transformações lineares de \mathbb{R}^n em \mathbb{R}^m
$\text{gfr}F$	Gráfico de F
$\text{dom}F$	Domínio de F
$\text{img}F$	Imagem de F
$F^{-1}(u)$	Inversa da aplicação ponto-conjunto F

Sumário

1	INTRODUÇÃO	13
2	PRELIMINARES	16
2.1	Motivação	16
2.2	Resultados Técnicos	19
2.3	A Projeção Inexata	24
3	CONVERGÊNCIA DO ALGORITMO	26
3.1	Convergência para a atualização de Broyden	35
4	ALGORITMO INEXATO	38
5	APLICAÇÕES E TESTES NUMÉRICOS	46
5.1	Aplicações	46
5.1.1	Desigualdade Variacional com Restrição	46
5.1.2	Desigualdade Variacional Split	48
5.2	Testes Numéricos	49
5.2.1	Problema 1	51
5.2.2	Problema 2	51
5.2.3	Problema 3	52
6	CONSIDERAÇÕES FINAIS	55
	REFERÊNCIAS	56

1 Introdução

Várias famílias de problemas em otimização podem ser remodelados na forma de uma Equação Generalizada e um método que proponha solução pode nos trazer uma forma padronizada para resolvermos vários desses problemas. Problemas de equilíbrio, desigualdades variacionais e problemas lineares e não lineares de otimização podem ser adaptados na forma de uma equação generalizada, mesmo que tal reformulação não torne o problema mais simples. Mostraremos em mais detalhes como essa associação pode ser feita em alguns momentos do trabalho. A nossa abordagem principal, nesse trabalho, dá-se sobre o problema das Equações Generalizadas com Restrições e buscamos uma contribuição ao que propôs Oliveira, Ferreira e Silva em [1]. Lá buscou-se resolver o problema com uma abordagem Newton e sobre outras hipóteses de regularidade que ainda serão comentadas nas seções posteriores. O problema em interesse é

$$x \in C, \quad f(x) + F(x) \ni 0, \quad (1.1)$$

Onde $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função continuamente diferenciável, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ é uma aplicação ponto conjunto com o gráfico fechado e não vazio e $C \subset \mathbb{R}^n$ é um conjunto convexo fechado.

O estudo de Equações Generalizadas começou por volta dos anos 1970 com S.M. Robinson ¹ e naquele momento trabalhava-se o problema sem restrições. Uma grande contribuição, na época, foi feita por Josephy ² em sua tese de doutorado [2]. Em seu trabalho, Josephy relacionou muitos exemplos e contextualizações foram feitas e serviram de base para quase todos os trabalhos futuros em torno das Equações Generalizadas. Com o passar do tempo muitas contribuições e novas aplicações foram feitas, desde aplicações computacionais até aplicações na medicina. O problema abordado era:

$$x \in \mathbb{R}^n, \quad f(x) + F(x) \ni 0, \quad (1.2)$$

Ao longo deste trabalho faremos várias menções, pelo menos dos resultados mais relacionados, sobre a teoria desenvolvida em torno do estudo de Equações Generalizadas, principalmente no tocante à métodos Quasi-Newton e Equações Generalizadas com Restrições.

As nossas contribuições serão divididas em dois segmentos, um deles trata-se de uma nova abordagem ao problema trabalhado em [1] fazendo uma relação ao que propôs

¹ Professor Emérito da Universidade de Wisconsin-Madison.

² Norman H. Josephy, Orientado em seu trabalho de doutorado por M. S. Robinson.

Artacho, Dontchev, Belyakov e Lopez em [3]. Trata-se de uma abordagem Quasi-Newton, com projeção inexata pra resolver o problema 1.1. A saber, a abordagem no primeiro momento é:

$$f(x_k) + B_k(y_k - x_k) + F(y_k) \ni 0 \quad k = 0, 1, \dots, \quad (1.3)$$

Muito já foi estudado a cerca das Equações Generalizadas e uma parte desse estudo como o feito em [4], [2], [5], [6] e [7] serviram de base para a abordagem Newton das Equações Generalizadas. Mais estudos foram feitos, alterando propriedades das funções f e F e abordagens diferentes como as propostas em [3], [8], [9], [10], [11], [12], [13] e [14]. Estes últimos, mesclam abordagens Newton e Quasi-Newton e apoiam-se em diversas condições de regularidade diferentes das quais trataremos posteriormente.

No segundo segmento, atacamos o mesmo problema usando um método inexato explanando a ideia proposta por Dontchev e Rockafellar em [11]. O método Quasi-Newton Inexato proposto é descrito por:

$$(f(x_k) + B_k(y_k - x_k) + F(y_k)) \cap R_k(x_k, y_k) \neq \emptyset, \quad k = 0, 1, \dots, \quad (1.4)$$

Onde B_k é uma sequência de Matrizes e R_k uma sequência de aplicações ponto-conjunto com gráfico fechado. Aqui, como no item anterior, y_k é o candidato a x_{k+1} .

Nos próximos capítulos falaremos sobre os elementos desse método, bem como todas as hipóteses de regularidade que foram exigidas para a convergência do algoritmo proposto. Vale a pena salientar, também, que houve muito estudo a respeito do tema e muitos deles, como [9], [15], [16], [17] e [18].

Neste trabalho, como já mencionamos, trabalharemos sempre com Equações Generalizadas com Restrições e para obter uma didática melhor teremos um capítulo destinado às noções preliminares. Neste capítulo exporemos de forma organizada os principais resultados que servirão de base para as demonstrações de convergências dos algoritmos propostos. E ainda, trataremos de forma ampla a projeção inexata que é a ferramenta chave para trabalharmos o problema com restrição.

No Capítulo 3, faremos uma breve retrospectiva do problema e suas abordagens anteriores, afim de que possamos entender a importância do que já foi feito sobre as EGR e conseguirmos expor quais foram nossas contribuições para o problema. Neste capítulo enunciaremos e demonstraremos a convergencia de um algoritmo Quasi-Newton para a EGR.

No Capítulo 4, assim como no anterior, faremos uma retrospectiva do que já foi estudado acerca dos algoritmos inexatos para as EG. Nele, também, enunciaremos e demonstraremos a convergência de um algoritmo Quasi-Newton inexato para as EGR que foram as nossas contribuições para essa abordagem.

No Capítulo 5, será explanado aplicações práticas das EGR e apresentaremos o tratamento que os problemas tiveram ao ser estudado, mesmo antes de adaptá-los a uma EGR, afim de evidenciarmos uma nova forma de abordar tais problemas.

Por fim faremos algumas considerações finais a respeito do trabalho e falaremos, também, sobre algumas perspectivas futuras sobre o tema.

2 Preliminares

Neste capítulo, apresentamos uma motivação para o estudo das Equações Generalizadas e apresentaremos vários resultados teóricos que servirão de base para os demais capítulos. Faremos uma revisão histórica do problema, passando pelo problema irrestrito, onde alguns resultados serão apenas referenciados. Aqui explicitaremos apenas as demonstração dos resultados que serão utilizados nos nossos algoritmo.

2.1 Motivação

Nesta seção procuraremos mostrar alguns problemas em Matemática Aplicada que podem ser reformulados na forma de uma Equação Generalizada (EG). Faremos aqui abordagens tanto para o caso restrito quanto para o caso irrestrito a fim de acompanhar a evolução da abordagem do problema. Para começar, olharemos para o problema de programação não linear:

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{S.a} \quad & h(x) = 0, \\ & g(x) \leq 0, \end{aligned} \tag{PNL}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}, h : \mathbb{R}^n \rightarrow \mathbb{R}^m, g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 e o $\Omega = \{x \in \mathbb{R}^n | h(x) = 0, g(x) \leq 0\}$ é o conjunto viável.

Definição 2.1. Dizemos que $x^* \in \Omega$ é um minimizador local de f em Ω quando existe $\delta > 0$, tal que $f(x^*) \leq f(x)$ para todo $x \in B_\delta(x^*) \cap \Omega$.

A condição de KKT é uma condição necessária de otimalidade para o problema (PNL) se alguma qualificação de restrição é válida. Nesse trabalho não entraremos no mérito das condições de qualificação e assumiremos, sempre que não for especificado, que temos regularidade no ponto, ou seja, independência linear dos gradientes das restrições ativas. Temos o seguinte teorema.

Teorema 2.1. Seja $x^* \in \Omega$ um minimizador local do problema (PNL) e suponha alguma condição de regularidade no ponto x^* . Então existem vetores λ^* e μ^* tais que

$$\begin{aligned} -\nabla f(x^*) &= \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla g_i(x^*), \\ \mu_i^* &\geq 0, \quad i = 1, \dots, p, \\ \mu_i^* g_i(x^*) &= 0, \quad i = 1, \dots, p. \end{aligned}$$

Conhecendo tal fato, podemos olhar com uma nova ótica o PNL a fim de associarmos a uma EG. Antes, precisamos de uma definição auxiliar.

Definição 2.2. Um conjunto $K \subset \mathbb{R}^n$ chama-se Cone quando

$$d \in K \implies td \in K \quad \forall t \in \mathbb{R}_+.$$

Pela definição, se K é um cone não vazio, necessariamente $0 \in K$. Abaixo, definimos cones que são importantes para nosso desenvolvimento.

Definição 2.3. Sejam $C \subset \mathbb{R}^n$ um conjunto convexo e $x^* \in C$. O Cone Normal no ponto x^* em relação ao conjunto C é dado por

$$N_C(x^*) = \{d \in \mathbb{R}^n \mid \langle d, x - x^* \rangle \leq 0 \quad \forall x \in C\} \tag{2.1}$$

E o Cone Dual de um cone $K \subset \mathbb{R}^n$ é definido por

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, d \rangle \leq 0 \quad \forall d \in K\}. \tag{2.2}$$

Figura 1 – Cone Normal

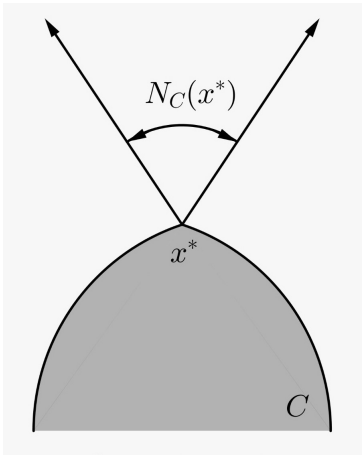
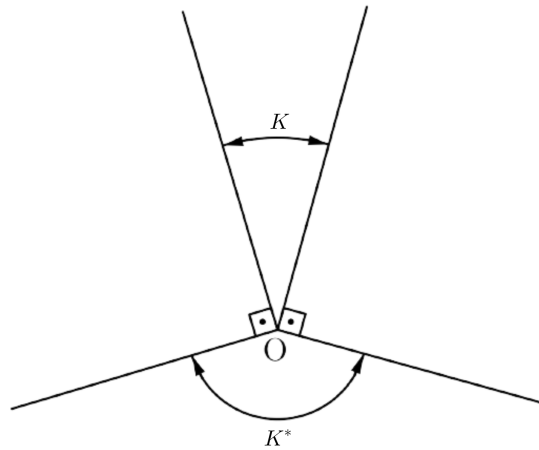


Figura 2 – Cone Dual



Nas figuras acima temos uma ilustração simples dos cones definidos. Aqui não entraremos a fundo na teoria de cones mas podemos encontrar em [28].

Exemplo 2.1. Definiremos a Lagrangiana do PNL como :

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{i=1}^p \mu_i g_i(x)$$

e denotaremos o gradiente de L com respeito a x por $\nabla_x L$. Dizemos que um terço (x^*, λ^*, μ^*) é KKT se, e somente se, $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$, $\mu^* \geq 0$, $g(x^*) \leq 0$, $\langle \mu^*, g(x^*) \rangle =$

0 e $h(x^*) = 0$. Seja $K : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+^p$. Observe que seu cone dual é $K^* = \{0\}^n \times \{0\}^m \times \mathbb{R}_+^p$. Definamos agora $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ por:

$$F(x, \lambda, \mu) = \begin{bmatrix} \nabla_x L(x, \lambda, \mu) \\ -h(x) \\ -g(x) \end{bmatrix}.$$

Então o terno (x^*, λ^*, μ^*) é KKT se, e somente se, (x^*, λ^*, μ^*) é solução da EG:

$$0 \in F(x, \lambda, \mu) + N_K(x, \lambda, \mu).$$

Antes de mostrar um problema que se remodela como um EGR, daremos mais um exemplo de problema que se remodela como uma EG.

Exemplo 2.2. Considere o problema de complementaridade não linear em \mathbb{R}^n :

$$f(x) \geq 0, x \geq 0, \langle x, f(x) \rangle = 0.$$

Pelo exemplo anterior é equivalente a resolver :

$$0 \in f(x) + N_K(x),$$

onde $K = \mathbb{R}_+^n$. Olharemos agora para a equação abaixo:

$$0 \in f(x^*) + f'(x^*)(x - x^*) + N_K(x). \quad (2.3)$$

Reagrupando os vetores constantes e reorganizando de forma conveniente, temos:

$$a := f(x^*) - f'(x^*)x^*, M := f'(x^*)$$

e sua linearização pode ser reescrita como:

$$0 \in a + Mx + N_K(x),$$

que é equivalente ao problema de complementaridade linear

$$a + Mx \geq 0, x \geq 0, \langle x, a + Mx \rangle = 0.$$

Assim, a linearização aplicada a um problema não linear de complementaridade resulta num problema linear de complementaridade. Logo, ao utilizar o método de Newton para resolver um problema não linear de complementaridade consistirá em resolver uma sequência de problemas lineares de complementaridade.

Observação 2.1. A equação 2.3 será aparecerá com frequência no nosso estudo e será vista em breve.

Daremos agora um exemplo que traz a Equação Generalizada com Restrições. Mais sobre tais exemplos pode ser visto em [1], [29], [30], [31] entre outros vários artigos como veremos nos próximos capítulos. Vamos expor aqui o problema de desigualdade variacional restrito (CVIP) que trata:

Exemplo 2.3. *Sejam U e Ω conjuntos fechados e convexos em \mathbb{R}^n e $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função contínua. O problema de desigualdade variacional restrito é definido como:*

$$\text{Encontrar } x^* \in U \cap \Omega \text{ tal que } \langle h(x^*), x - x^* \rangle \geq 0, \forall x \in U. \quad (2.4)$$

Note agora que o problema é equivalente a encontrar $x^ \in U \cap \Omega$ tal que $0 \in h(x^*) + N_U(x^*)$. Assim o problema pode ser abordado como uma equação generalizada com restrição.*

2.2 Resultados Técnicos

Nesta seção apresentaremos os principais resultados teóricos a serem usados durante toda a tese. Muitas das fundamentações deste tópico podem ser encontradas com mais detalhes em [4]. Aqui estabeleceremos algumas notações, apresentaremos as principais noções de distância e convergência para aplicações ponto-conjunto (usadas sobre F) e apresentaremos, também, o que chamamos de condição de regularidade para essas aplicações.

Antes de expor as principais condições que nossa aplicação ponto-conjunto terá que satisfazer, temos que fixar algumas notações e expor algumas definições que serão cruciais. Antes de tudo, usaremos a notação de bola aberta e fechada como as seguintes, respectivamente:

$$B_\delta(x) := \{y \in \mathbb{R}^n : \|x - y\| < \delta\}, \quad B_\delta[x] := \{y \in \mathbb{R}^n : \|x - y\| \leq \delta\}.$$

Usaremos a norma usual de matrizes definida por:

$$\|A\| := \sup \left\{ \frac{\|Ax\|}{\|x\|} : \|x\| \in \mathbb{R}^n \text{ com } x \neq 0 \right\}$$

desde que $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, que denotaremos pelo espaço das transformações lineares de \mathbb{R}^n em \mathbb{R}^m . O gráfico, domínio, imagem e inversa da nossa aplicação ponto-conjunto F são definidas seguindo o padrão usado em funções, fixaremos uma denotação para cada conjunto e mostraremos suas definições como as seguintes, respectivamente:

$$\text{gr } F := \{(x, u) \in \mathbb{R}^n \times \mathbb{R}^m : u \in F(x)\}$$

$$\text{dom } F := \{x \in \mathbb{R}^n : F(x) \neq \emptyset\}$$

$$\text{img } F := \{u \in \mathbb{R}^m : u \in F(x) \text{ para algum } x \in \mathbb{R}^n\}$$

$$F^{-1}(u) := \{x \in \mathbb{R}^n : u \in F(x)\}$$

Para a função f seguiremos a notação acima, com sua devida adaptação e usaremos $f'(x)$ para representar a matriz Jacobiana de f em x . Posteriormente, necessitaremos usar uma generalização do teorema do ponto fixo, para isso precisaremos de algumas definições e também de algumas noções de distancia que serão dadas abaixo.

Definição 2.4. *Seja $f : \Omega \rightarrow \mathbb{R}^m$ e $F : \Omega \rightrightarrows \mathbb{R}^m$ com $\Omega \subseteq \mathbb{R}^n$, a linearização parcial de $f + F$ em $x \in \Omega$ é a aplicação ponto-conjunto denotada por $L_{f+F}(x, \cdot) : \Omega \rightrightarrows \mathbb{R}^m$ e definida como:*

$$L_{f+F}(x_*, y) := f(x_*) + f'(x_*)(y - x_*) + F(y). \quad (2.5)$$

Definição 2.5. *Para conjuntos C e D em \mathbb{R}^n , a distância de x a D e o excesso de C sobre D são respectivamente definidos por:*

$$d(x, D) := \inf_{y \in D} \|x - y\|, \quad e(C, D) := \sup_{x \in C} d(x, D), \quad (2.6)$$

Convencionalmente adotamos que $d(x, D) = +\infty$, quando $D = \emptyset$, $e(\emptyset, D) = 0$ quando $D \neq \emptyset$, e $e(\emptyset, \emptyset) = +\infty$.

De posse dessas notações e definições fixadas acima, somos capazes de definir as condições de regularidade que serão exigidas para nossa aplicação F , como faremos a seguir.

Definição 2.6. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e não vazio. Dizemos que a aplicação ponto-conjunto $G : \Omega \rightrightarrows \mathbb{R}^m$ é Metricamente Regular em $\bar{x} \in \Omega$ para $\bar{u} \in \mathbb{R}^m$ quando $\bar{u} \in G(\bar{x})$, o gráfico de G é localmente fechado em (\bar{x}, \bar{u}) , e existem constantes $\kappa \geq 0$, $a > 0$ e $b > 0$ tais que $B_a[\bar{x}] \subset \Omega$ e*

$$d(x, G^{-1}(u)) \leq \kappa[d(u, G(x))], \text{ para todo } (x, u) \in B_a[\bar{x}] \times B_b[\bar{u}]. \quad (2.7)$$

Além do mais, se a aplicação $B_b[\bar{u}] \ni u \mapsto G^{-1}(u) \cap B_a[\bar{x}]$ é ponto a ponto, então G é chamado Fortemente Metricamente Regular em $\bar{x} \in \Omega$ para $\bar{u} \in \mathbb{R}^m$, com as constantes $\kappa \geq 0$, $a > 0$ e $b > 0$ associadas.

Observação 2.2. *Se G é fortemente metricamente regular em $\bar{x} \in \Omega$ para $\bar{u} \in \mathbb{R}^m$ com constantes $\kappa \geq 0$, $a > 0$, e $b > 0$, então a aplicação $B_b[\bar{u}] \ni u \mapsto G^{-1}(u) \cap B_a[\bar{x}]$ é ponto a ponto e lipschitz contínua em $B_b[\bar{u}]$ com constante de Lipschitz κ , isto é, $\|G^{-1}(u) \cap B_a[\bar{x}] - G^{-1}(v) \cap B_a[\bar{x}]\| \leq \kappa\|u - v\|$ para todo $u, v \in B_b[\bar{u}]$. Veja [4, Proposition 3G.1, p. 193].*

Enunciaremos a seguinte proposição que será útil na demonstração de um dos principais resultados do trabalho, sua demonstração pode ser encontrada em [11].

Proposição 2.1. *Seja $\zeta > 0$ e assumamos que a aplicação $f + F$ é metricamente regular em \bar{x} para 0 com constantes λ e $\lambda\zeta < 1$. Seja $u \in X$ e considere a aplicação*

$$X \ni x \mapsto L_{f+F}(u, x) = f(u) + B_u(x - u) + F(x). \quad (2.8)$$

Se $\|B_u - f'(\bar{x})\| \leq \zeta$ para algum $\zeta > 0$ então para cada $\kappa > \lambda/(1 - \lambda\zeta)$, existem números positivos a e b tais que

$$e(L_{f+F}^{-1}(u, y) \cap \mathcal{B}_a[\bar{x}], L_{f+F}^{-1}(u, y')) \leq \kappa\|y - y'\| \quad \text{para todo } u \in \mathcal{B}_a[\bar{x}] \text{ e } y, y' \in \mathcal{B}_b[0]. \quad (2.9)$$

Em vários trabalhos são usadas a propriedade de Aubin para caracterização de regularidade da aplicação ponto-conjunto para entendermos melhor falaremos um pouco a respeito, dando sua devida relação com nossa caracterização quando possível. Para fins de curiosidade muitas definições nesse aspecto podem ser encontradas no trabalho de Dontchev e Rockafellar [32]. Por hora traremos apenas a definição de aplicação Aubin Contínua que será usada posteriormente.

Definição 2.7. *A aplicação $S : Y \rightrightarrows X$ é dita ser Aubin Contínua, ou ter a propriedade de Aubin, em \bar{y} para \bar{x} se $\bar{x} \in S(\bar{y})$ e existem $\lambda > 0$ juntos com uma vizinhança U de \bar{x} e V de \bar{y} tais que*

$$e(S(y) \cap U, S(y')) \leq \lambda\|y - y'\| \quad \text{para todo } y, y' \in V.$$

A mesma definição pode ser vista de forma equivalente como a seguir. .

Definição 2.8. *A aplicação ponto-conjunto $S : Y \rightrightarrows X$ tem a propriedade de Aubin em \bar{x} para $\bar{y} \in S(\bar{x})$ se $\text{gfr} S$ é localmente fechado em (\bar{x}, \bar{y}) e para algum $l > 0$ existem $\rho, \delta > 0$ tais que*

$$S(x') \cap B_\rho(\bar{y}) \subset S(x) + B_{l\|x'-x\|}(0), \quad (x', x) \in B_\delta(\bar{x}) \quad (2.10)$$

A figura nos ajuda a entender melhor o conjunto $S(x) + B_{l\|x'-x\|}(0)$ e isso nos ajudará bastante a entender como se comporta, graficamente, as aplicações que satisfazem a propriedade de Aubin.

De posse dessa figura, conseguimos entender como se comporta graficamente uma aplicação que satisfaz a propriedade de Aubin.

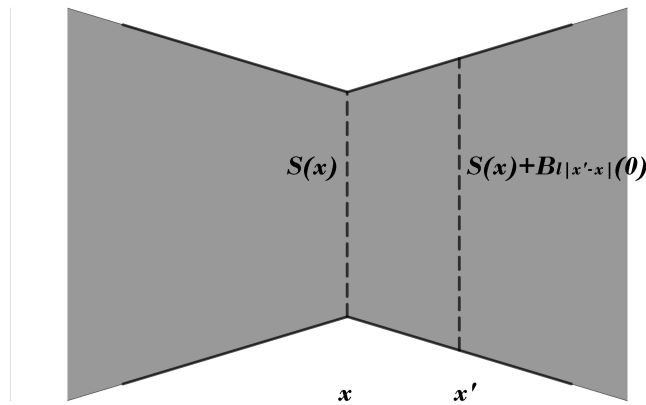


Figura 3 – O gráfico de $x \mapsto S(x) + B_{l_{\|x'-x\|}}(0)$

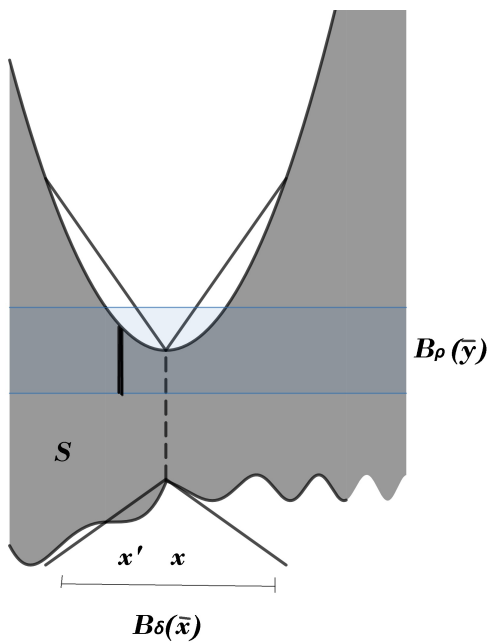


Figura 4 – Satisfaz Aubin.

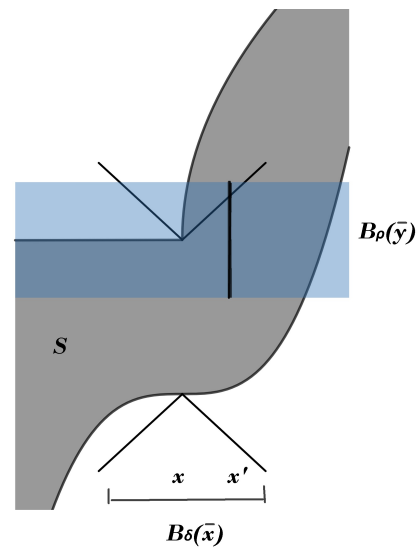


Figura 5 – Não satisfaz Aubin.

Temos também a versão parcial para a propriedade de Aubin para aplicações de duas variáveis. Dizemos que a aplicação $T : P \times X \rightrightarrows Y$ é parcialmente Aubin Contínua em \bar{t} para \bar{x} uniformemente em p em torno de \bar{p} se $\bar{x} \in T(\bar{y}, \bar{p})$ e existem $\lambda > 0$ e vizinhanças U de \bar{x} , V de \bar{y} e Q de \bar{p} tais que

$$e(T(p, y) \cap U, T(p, y')) \leq \lambda \|y - y'\| \text{ para todo } y, y' \in V \text{ e todo } p \in Q.$$

Observação 2.3. Vale a pena ressaltar, como feito em [34], que a Propriedade de Aubin, ou ser Aubin Contínua é também conhecida como Pseudo-Lipschitz. Então, podemos ver algumas associações dessa forma na literatura.

Um resultado que associa as condições impostas as aplicações ponto-conjunto em todo o texto será anunciado abaixo, muito mais sobre essas equivalências podem ser encontradas em [35] mas não as trataremos porque elas fogem do mérito do trabalho.

Teorema 2.2. *Seja X e Y espaços de Banach e seja $F : X \rightrightarrows Y$ uma aplicação Ponto-Conjunto com gráfico fechado. Então as seguintes afirmações são equivalentes*

- F é metricamente regular em $(\bar{x}, \bar{y}) \in \text{gfr } F$
- F^{-1} satisfaz a propriedade de Aubin em $(\bar{x}, \bar{y}) \in \text{gfr } F^{-1}$

Um resultado bastante importante e que é usado como principal ferramenta para demonstração do resultado de convergência do algoritmo, no próximo capítulo, é uma generalização do teorema das contrações (ou teorema do ponto fixo) para aplicações ponto-conjunto. Devido sua grande importância, mesmo não sendo o objetivo do trabalho, trataremos sua demonstração que segue o mesmo padrão da prova encontrada em [4, Theorem 5E.2, p. 313] com suas devidas adaptações ao nosso problema e notações. O resultado é o seguinte:

Teorema 2.3. *Sejam $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ uma aplicação ponto-conjunto e $\bar{x} \in \mathbb{R}^n$. Suponha que existem escalares $\rho > 0$ e $\lambda \in (0, 1)$ tais que o conjunto $\text{grf } \Phi \cap (B_\rho[\bar{x}] \times B_\rho[\bar{x}])$ é fechado e valem as seguintes condições:*

- (i) $d(\bar{x}, \Phi(\bar{x})) \leq \rho(1 - \lambda)$;
- (ii) $e(\Phi(p) \cap B_\rho[\bar{x}], \Phi(q)) \leq \lambda\|p - q\|$ para todo $p, q \in B_\rho[\bar{x}]$.

Então, Φ tem um ponto fixo em $B_\rho[\bar{x}]$. Ou seja, existe $y \in B_\rho[\bar{x}]$ tal que $y \in \Phi(y)$.

Demonstração. Assumindo (i), existe $x_1 \in \Phi(\bar{x})$ tal que $d(x_1, \bar{x}) < \rho(1 - \lambda)$. Procedendo, por indução, seja $x_0 = \bar{x}$ e suponha que existe $x_{k+1} \in \Phi(x_k) \cap B_\rho[\bar{x}]$ para $k = 0, 1, \dots, j - 1$ com

$$d(x_{k+1}, x_k) < \rho(1 - \lambda)\lambda^k$$

Por assumir (ii),

$$d(x_j, \Phi(x_j)) \leq e(\Phi(x_{j-1}) \cap B_\rho[\bar{x}], \Phi(x_j)) \leq \lambda d(x_j, x_{j-1}) < \rho(1 - \lambda)\lambda^j$$

isto implica que existe um $x_{j+1} \in \Phi(x_j)$ tal que

$$d(x_{j+1}, x_j) < \rho(1 - \lambda)\lambda^j.$$

Por desigualdade triangular

$$d(x_{j+1}, \bar{x}) \leq \sum_{i=0}^j d(x_{i+1}, x_i) < \rho(1 - \lambda) \sum_{i=0}^j \lambda^i < \rho.$$

Assim $x_{j+1} \in \Phi(x_j) \cap B_\rho[\bar{x}]$ e o passo de indução fica completo. Para qualquer $k > m > 1$ teremos então

$$d(x_k, x_m) \leq \sum_{i=m}^{k-1} d(x_{i+1}, x_i) < \rho(1 - \lambda) \sum_{i=m}^{k-1} \lambda^i < \rho\lambda^m.$$

Assim, x_k é uma sequência de Cauchy e consequentemente converge para algum $x \in B_\rho[\bar{x}]$. Como sabemos que $(x_{k-1}, x_k) \in \text{grf}\Phi \cap (B_\rho[\bar{x}] \times B_\rho[\bar{x}])$ que é um conjunto fechado, concluímos que $x \in \Phi(x)$.

□

Encerraremos esta seção apresentando resultados sobre a projeção factível inexata. Estes resultados serão de grande valia uma vez que trabalharemos com um problema restrito e a cada passo do algoritmo essa factibilidade pode ser perdida. Vale a pena ressaltar que existem outras formas de definir projeções inexatas mas não serão utilizadas.

2.3 A Projeção Inexata

Existem diversas maneiras de definir uma projeção factível inexata. Apresentaremos abaixo a que usaremos nos algoritmos.

Definição 2.9. *Seja $C \subset \mathbb{R}^n$ um conjunto fechado e convexo, com $x \in C$ e $\theta \geq 0$. A aplicação projeção factível inexata relativa a x com tolerância θ , denotada por $P_C(\cdot, x, \theta) : \mathbb{R}^n \rightrightarrows C$ é a aplicação ponto-conjunto definida por:*

$$P_C(y, x, \theta) := \{w \in C : \langle y - w, z - w \rangle \leq \theta \|y - x\|^2, \quad \forall z \in C\}.$$

Cada ponto $w \in P_C(y, x, \theta)$ é chamado de uma projeção factível inexata de y em C com respeito a x e com erro de tolerância θ .

A figura abaixo serve pra dar uma visão geral da atuação da projeção factível inexata.

Observação 2.4. *Como $C \subset \mathbb{R}^n$ é um conjunto convexo e fechado, [36, Proposição 2.1.3, p. 201] garante que para cada $y \in \mathbb{R}^n$ temos $P_C(y) \in P_C(y, x, \theta)$, onde P_C denota a aplicação projeção exata. Portanto, $P_C(y, x, \theta) \neq \emptyset$ para todo $y \in \mathbb{R}^n$ e $x \in C$. Se $\theta = 0$ na Definição 2.9, então $P_C(y, x, 0) = \{P_C(y)\}$ para todo $y \in \mathbb{R}^n$ e $x \in C$. Usamos $P_C(y, x, 0) = P_C(y)$ em vez de $P_C(y, x, 0) = \{P_C(y)\}$.*

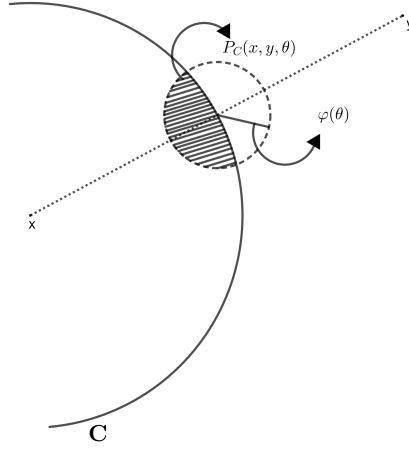


Figura 6 – Projeção Factível Inexata

O próximo resultado é de muita importância para o restante do trabalho e mesmo se tratando de um resultado bastante técnico traremos sua demonstração devido a sua grande usabilidade no capítulo seguinte.

Lema 2.1. *Sejam $y, \tilde{y} \in \mathbb{R}^n$, $x, \tilde{x} \in C$, e $\theta \geq 0$. Então, para algum $w \in P_C(y, x, \theta)$, temos:*

$$\|w - P_C(\tilde{y}, \tilde{x}, 0)\| \leq \|y - \tilde{y}\| + \sqrt{2\theta}\|y - x\|.$$

Demonstração. Para simplificar um pouco a notação faremos $\tilde{w} = P_C(\tilde{y}, \tilde{x}, 0)$, e tomamos $w \in P_C(y, x, \theta)$. Primeiro, note que $\|y - \tilde{y}\|^2 = \|(y - w) - (\tilde{y} - \tilde{w})\|^2 + \|w - \tilde{w}\|^2 + 2\langle(y - w) - (\tilde{y} - \tilde{w}), w - \tilde{w}\rangle$, e isso implica que:

$$\|w - \tilde{w}\|^2 \leq \|y - \tilde{y}\|^2 + 2\langle y - w, \tilde{w} - w \rangle + 2\langle \tilde{y} - \tilde{w}, w - \tilde{w} \rangle.$$

Como $\tilde{w} = P_C(\tilde{y}, \tilde{x}, 0)$ e $w \in P_C(y, x, \theta)$, usando a Definição 2.9 e o fato que $\tilde{w}, w \in C$, podemos concluir que $\langle y - w, \tilde{w} - w \rangle \leq \theta\|y - x\|^2$ e $\langle \tilde{y} - \tilde{w}, w - \tilde{w} \rangle \leq 0$. Assim, ao combinarmos essas três últimas desigualdades ganhamos $\|w - \tilde{w}\|^2 \leq \|y - \tilde{y}\|^2 + 2\theta\|y - x\|^2$, e então $\|w - \tilde{w}\| \leq \|y - \tilde{y}\| + \sqrt{2\theta}\|y - x\|$, que resulta na desigualdade desejada. \square

Observação 2.5. *O resultado apresentado acima é uma propriedade básica da projeção factível inexata e um resultado similar pode ser encontrado em [37, Lema 4], a título de curiosidade.*

3 Convergência do Algoritmo

Neste capítulo usaremos o método Quasi-Newton para resolver o problema (1.1). Aqui também estudaremos as propriedades de convergência local de uma sequência gerada através do método. Faremos uma análise assumindo algumas condições de regularidade, que já foram trabalhadas no capítulo anterior, sobre a aplicação ponto-conjunto $f + F$ e assumiremos, em alguns momentos, algumas condições sobre a f como Lipschitz continuidade da sua derivada f' . Como trabalharemos com um problema restrito será necessário nos preocuparmos com a factibilidade dos termos da sequência, para isso usaremos alguns conceitos e propriedades da projeção inexata, que também foram trabalhados no capítulo anterior.

Antes de abordar nosso problema, entenderemos as bases do nosso estudo e o que nos levou a abordar dessa forma as Equações Generalizadas com Restrição. Antes de tudo, revisaremos o que fez [3] no que se trata a uma abordagem Quasi-Newton às Equações Generalizadas sem restrição. Os Teoremas a seguir são de grande importância para o nosso estudo e, com certeza, são de grande importância para o estudo das EG.

Teorema 3.1. (*Perturbação da Regularidade Métrica*) *Seja (X, ρ) um espaço métrico completo e (Y, ρ) um espaço métrico linear com uma invariante de deslocamento métrico. Considere uma aplicação $H : X \rightrightarrows Y$ com gráfico fechado e ponto $(\tilde{x}, \tilde{y}) \in \text{gph } H$ onde cada H é metricamente regular, isto é, existem constantes positivas a, b e k tal que*

$$d(x, H^{-1}(y)) \leq kd(y, H(x)) \text{ para todo } (x, y) \in \mathbb{B}_a(\tilde{x}) \times \mathbb{B}_b(\tilde{y}) \quad (3.1)$$

Seja $\mu > 0$ tal que $k\mu < 1$ e seja $k' > k$. Então para cada positivo α e β tal que

$$\alpha \leq a/2, \mu\alpha + 2\beta \leq b \text{ e } 2k'\beta \leq \alpha(1 - k\mu) \quad (3.2)$$

e para cada função $h : X \rightarrow Y$ satisfazendo

$$\rho(h(\tilde{x}), 0) \leq \beta \quad (3.3)$$

e

$$\rho(h(x), h(\tilde{x})) \leq \mu\rho(x, \tilde{x}) \text{ para cada } x, x' \in \mathbb{B}_\alpha(\tilde{x}) \quad (3.4)$$

a aplicação $h + H$ tem as seguintes propriedades: para cada $y, y' \in \mathbb{B}_\beta(\tilde{y})$ e cada $x \in (h + H)^{-1}(y) \cap \mathbb{B}_\alpha(\tilde{x})$ existe $x' \in (h + H)^{-1}(y')$ tal que

$$\rho(x, x') \leq \frac{k'}{1 - k\mu} \rho(y, y') \quad (3.5)$$

Além disso, se a aplicação H é fortemente metricamente regular em \tilde{x} para \tilde{y} ; especificamente, a aplicação $y \mapsto H^{-1}(y) \cap \mathbb{B}_\alpha(\tilde{x})$ é ponto a ponto e Lipschitz contínua em $\mathbb{B}_b(\tilde{y})$ com constante Lipschitz k , então para $\mu k', \alpha$ e β como citado e alguma função h satisfazendo (3.3) e (3.4), a aplicação $y \mapsto (h + H)^{-1}(y) \cap \mathbb{B}_\alpha(\tilde{x})$ é uma função Lipschitz contínua em $\mathbb{B}_\beta(\tilde{y})$ com constante Lipschitz $k'/(1 - k\mu)$.

Várias suposições foram feitas, no artigo citado, a respeito da regularidade de $h + H$ tanto para quando satisfaz a propriedade de Aubin, quanto quando é fortemente metricamente regular. Foram, também, analisados alguns casos particulares para a função h desde sua nulidade ou outros casos específicos, mas não entraremos em detalhes. O Teorema a seguir garante a convergência do algoritmo.

Teorema 3.2. *Suponha que $f + F$ seja metricamente regular em \tilde{x} para 0 com constante λ . Então, em particular, \tilde{x} é uma solução de (1.2). Considere o método Quasi-Newton (3.14) e assuma que*

$$\|B_0 - f'(\tilde{x})\| < 1/(2\lambda) \quad (3.6)$$

Além disso, assuma que existe uma constante $c > 0$ e uma vizinhança U de \tilde{x} tal que, para $k = 0, 1, \dots$, e para cada B_k e cada $x_k, x_{k+1} \in U$, $x_k \neq x_{k+1}$, satisfazendo (3.14) o operador B_{k+1} é escolhido de modo que

$$\|B_{k+1} - f'(\tilde{x})\| \leq \|B_k - f'(\tilde{x})\| + c(\|x_k - \tilde{x}\| + \|x_{k+1} - \tilde{x}\|) \quad (3.7)$$

Então existe uma vizinhança O de \tilde{x} tal que para cada $x_0 \in O$ exista uma sequência $\{x_k\}$ começando em x_0 e gerada por (3.14) que permanece em O e chega a uma solução de (1.2) em O após uma quantidade finita de passos ou converge q -linearmente para \tilde{x} . Se além disso $f + F$ é fortemente metricamente regular em \tilde{x} para 0 então para cada $x_0 \in O$ existe uma única sequência $\{x_k\}$ em O começando em x_0 e gerada de acordo com (3.14), e essa sequência converge q -linearmente para \tilde{x} .

O próximo Teorema é bem semelhante ao enunciado acima, basta notar que agora a atualização B_k é Broyden. Explicitaremos, em breve, tal atualização que por ora será apenas citada.

Teorema 3.3. *Considere a Equação Generalizada (1.2) na configuração dos espaços X e Y Hilbert com soluções \tilde{x} e suponha que a aplicação derivada $f'(x)$ é lipschitz contínua em torno de \tilde{x} . Em seguida, suponha que $f + F$ seja metricamente regular em \tilde{x} para 0 com constante λ . Considere o método Quasi-Newton (3.14) aplicado em (1.2) com as atualizações de Broyden e com B_0 satisfazendo (3.6). Então existe uma vizinhança O de \tilde{x} tal que para algum $x_0 \in O$ e uma sequência $\{x_k\}$ começando de x_0 , gerada por (3.14) que fica em O e chega a uma solução de (1.2) em uma quantidade finita de passos ou converge q -linearmente para \tilde{x} . Se além disso $f + F$ é fortemente metricamente regular em \tilde{x} para 0*

então para todo $x_0 \in O$ existe uma única sequência $\{x_k\}$ em O começando de x_0 e gerada de acordo com (3.14), e esta sequência converge q -linearmente para \tilde{x} .

Observação 3.1. *Embora não tenhamos definidos todos os termos e expressões, enunciaremos o teorema anterior de modo que sigamos a leitura do texto em ordem cronológica.*

Antes de partir para o nosso problema, que se trata de uma abordagem Quasi-Newton ao problema restrito, vamos mostrar como foi a abordagem Newton ao problema, feita em [1]. Afim de que entendamos nossa base teórica, explicitaremos o algoritmo e o teorema de convergência, mas não traremos sua demonstração e nem os casos particulares uma vez que eles foram bem feitos e detalhados no artigo citado.

O algoritmo abaixo é bem parecido com o que faremos e pra ficar claro suas alterações o traremos em uma estrutura parecida.

Newton-InexP method

Passo 0. Seja $x_0 \in C$ e $\{\theta_k\} \subset [0, +\infty)$ dados, e tome $k = 0$.

Passo 1. Se $f(x_k) + F(x_k) \ni 0$, então **pare**; caso contrário, compute $y_k \in \mathbb{R}^n$ tal que

$$f(x_k) + f'(x_k)(y_k - x_k) + F(y_k) \ni 0. \quad (3.8)$$

Passo 2. Se $y_k \in C$, tome $x_{k+1} = y_k$; caso contrário tome algum $x_{k+1} \in C$ satisfazendo

$$x_{k+1} \in P_C(y_k, x_k, \theta_k). \quad (3.9)$$

Passo 3. Faça $k \leftarrow k + 1$, e vá para o **Passo 1**.

Teorema 3.4. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto, $f : \Omega \rightarrow \mathbb{R}^m$ continuamente diferenciável em Ω , e $F : \Omega \rightrightarrows \mathbb{R}^m$ uma aplicação ponto-conjunto com gráfico fechado. Assuma que $C \subset \Omega$ é um conjunto fechado e convexo, $x_* \in C$, $f(x_*) + F(x_*) \ni 0$, exista $L > 0$ tal que*

$$\|f'(x) - f'(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Omega, \quad (3.10)$$

e a aplicação ponto-conjunto $\Omega \ni y \mapsto L_{f+F}(x_, y)$ seja metricamente regular em x_* para 0, com constantes $\kappa > 0$, $a > 0$, e $b > 0$. Seja $r := \sup \{t \in \mathbb{R} : B_t(x_*) \subset \Omega\}$, $\{\theta_k\} \subset [0, 1/2)$*

e

$$r_* := \min \left\{ r, \frac{2(1 - \sqrt{2\tilde{\theta}})}{(3 - \sqrt{2\tilde{\theta}})\kappa L}, a, \sqrt{\frac{2b}{3L}} \right\}, \quad \tilde{\theta} := \sup_k \theta_k < \frac{1}{2}. \quad (3.11)$$

Então, para cada $x_0 \in C \cap B_{r_*}(x_*) \setminus \{x_*\}$, existe uma sequência $\{x_k\}$ gerada pelo algoritmo 'Newton-InexP method' que resolve (1.1), associado ao $\{\theta_k\}$ e iniciando em x_0 , que está contido em $B_{r_*}(x_*) \cap C$ e converge para x_* com a seguinte taxa de convergência:

$$\|x_* - x_{k+1}\| \leq \left[\left(1 + \sqrt{2\theta_k}\right) \frac{\kappa L \|x_* - x_k\|}{2(1 - \kappa L \|x_* - x_k\|)} + \sqrt{2\theta_k} \right] \|x_* - x_k\|, \quad k = 0, 1, \dots \quad (3.12)$$

Como consequência, se $\lim_{k \rightarrow +\infty} \theta_k = 0$ então $\{x_k\}$ converge para x_* superlinearmente. Em particular, se $\theta_k = 0$ para todo $k = 0, 1, \dots$, então

$$\|x_* - x_{k+1}\| \leq \frac{3\kappa L}{2} \|x_* - x_k\|^2, \quad k = 0, 1, \dots, \quad (3.13)$$

e $\{x_k\}$ converge para x_* Q -quadraticamente. Além do mais, se a aplicação $L_{f+F}(x_*, \cdot)$ é fortemente metricamente regular em x_* para 0, então x_* é a única solução de (1.1) em $B_{r_*}(x_*)$, e cada sequência gerada pelo algoritmo 'Newton-InexP method' associado a $\{\theta_k\}$ e iniciando em $x_0 \in C \cap B_{r_*}(x_*) \setminus \{x_*\}$ satisfaz (3.12) e converge para x_* .

Os teoremas acima foram bastante importantes para a nossa contribuição para as Equações Generalizadas com Restrição. Agora, ao pensarmos no que foi abordado em [3] e [1] obtemos um método Quasi-Newton para as Equações Generalizadas com restrições.

O método Quasi-Newton que propomos e chamaremos de Quasi-Newton-InexP para resolver a Equação Generalizada com restrição (1.1), com a projeção inexata e com $\{\theta_k\} \subset [0, +\infty)$ e $x_0 \in C$ como ponto inicial é descrito formalmente como o seguinte:

Quasi-Newton-InexP

Step 0. Seja $x_0 \in C$ e $\{\theta_j\} \subset [0, +\infty)$ dados, tome $k = 0$.

Step 1. Se $f(x_k) + F(x_k) \ni 0$, então **pare**; caso contrário, busque $y_k \in \mathbb{R}^n$ tal que

$$f(x_k) + B_k(y_k - x_k) + F(y_k) \ni 0. \quad (3.14)$$

Step 2. Se $y_k \in C$, tome $x_{k+1} = y_k$; caso contrário escolha algum $x_{k+1} \in C$ satisfazendo

$$x_{k+1} \in P_C(y_k, x_k, \theta_k). \quad (3.15)$$

Step 3. Faça $k \leftarrow k + 1$, e vá para **Step 1**.

Observemos que no momento em que x_k é solução do problema (1.1) nosso método pára ainda no passo 1. Caso contrário computamos y_k através da inclusão (3.14). Um fato importante é que y_k pode não pertencer ao conjunto C e caso isso aconteça faremos a projeção inexata sobre C e essa projeção será nosso próximo iterando. Se y_k pertencer ao conjunto ele próprio será o próximo iterando.

A projeção usada nos exemplos é um problema de otimização $\min_{z \in C} \{\|z - y_k\|^2/2\}$, satisfazendo $\langle y_k - x_{k+1}, z - x_{k+1} \rangle \leq \theta_k \|y_k - x_k\|^2$ para qualquer $z \in C$. A escolha da tolerância θ_k será importante na demonstração do teorema de convergência como veremos em breve.

A seguir enunciaremos o principal teorema de convergência do método proposto acima. Em seguida faremos sua demonstração que segue de uma combinação de resultados que serão todos expostos ou referenciados quando for o caso.

Teorema 3.5. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto, $f : \Omega \rightarrow \mathbb{R}^m$ uma função definida em Ω , e $F : \Omega \rightrightarrows \mathbb{R}^m$ uma aplicação ponto-conjunto com gráfico fechado e $f + F$ metricamente regular, em x_* para 0, com constantes $\kappa > 0$, $a > 0$ e $b > 0$. Assuma que $C \subset \Omega$ é um conjunto fechado e convexo, $x_* \in C$, $0 \in f(x_*) + F(x_*)$, existe $\epsilon > 0$ tal que*

$$\|f(u) - f(v) - f'(x_*)(u - v)\| \leq \epsilon \|u - v\|, \quad \text{para todo } u, v \in B_{\bar{a}}(x_*). \quad (3.16)$$

Considere o método Quasi-Newton-InexP com $\theta_k \geq 0$ e assuma que

$$\delta := \|B_0 - f'(x_*)\| < \frac{1}{2\kappa}. \quad (3.17)$$

Além do mais, assumimos que existem uma constante $c > 0$ e uma vizinhança U de \bar{x} tal que, para $k = 0, 1, \dots$, e para cada B_k e cada $x_k, y_k \in U$, $x_k \neq y_k$, todos satisfazendo (3.14), o operador B_{k+1} é escolhido

$$\|B_{k+1} - f'(x_*)\| \leq \|B_k - f'(x_*)\| + c(\|x_k - x_*\| + \|y_k - x_*\|). \quad (3.18)$$

Então, existe uma vizinhança \mathcal{U} de x_* , tal que para cada $x_0 \in C \cap \mathcal{U} \setminus \{x_*\}$, existe uma sequência $\{x_k\}$ gerada pelo algoritmo Quasi-Newton-Inexp que resolve (1.1), associado a $\{\theta_k\}$ e iniciando em x_0 , que está contido em $C \cap \mathcal{U}$ e converge linearmente para x_* .

Para demonstrarmos este teorema, precisaremos antes de alguns lemas técnicos, que serão expostos a seguir, começando por:

Lema 3.1. *Se $\theta \geq 0$, $x \in C \cap B_{r_*}(x_*) \setminus \{x_*\}$ e $y \in \Phi_x(y)$, vale que:*

$$\|x_* - w\| \leq (1 + \sqrt{2\theta})\|x_* - y\| + \sqrt{2\theta}\|x_* - x\|, \quad \forall w \in P_C(y, x, \theta).$$

Demonstração. Tome $w \in P_C(y, x, \theta)$. Então, aplicando o Lema 2.1 com $\tilde{y} = x_*$ e $\tilde{x} = x_*$, temos:

$$\begin{aligned} \|P_C(x_*, x_*, 0) - w\| &\leq \|x_* - y\| + \sqrt{2\theta}(\|x_* - x\| + \|x_* - y\|) \\ &= (1 + \sqrt{2\theta})\|x_* - y\| + \sqrt{2\theta}\|x_* - x\|. \end{aligned}$$

□

Com isso, podemos demonstrar o Teorema.

Demonstração. Para mostrarmos o resultado, usaremos indução em n com o propósito de mostrar que

$$\|x_* - x_{n+1}\| \leq \left[(1 + \sqrt{2\theta_n}) \frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right)}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)} + \sqrt{2\theta_n} \right] \|x_* - x_n\| < \gamma \|x_* - x_n\|.$$

Para isso, faremos algumas suposições que serão necessárias para obtenção de tal resultado.

Seja $0 < \gamma < 1$. Fazemos \bar{a} pequeno suficiente e $\kappa' > \kappa$ tal que $B_{\bar{a}}(x_*) \subset B_a(x_*)$,

$$\frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right)}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)} < \frac{1 - \sqrt{2\theta_n}}{1 + \sqrt{2\theta_n}}, \quad (3.19)$$

e

$$\gamma = \max \left\{ \left[(1 + \sqrt{2\theta_0}) \frac{\kappa'(\epsilon + \delta)}{1 - \kappa\delta} + \sqrt{2\theta_0} \right], \left[(1 + \sqrt{2\theta_n}) \frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right)}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)} + \sqrt{2\theta_n} \right] \right\}. \quad (3.20)$$

Seja $x_0 \in C \cap B_{r_*}(x_*) \setminus \{x_*\}$ e $n = 0$, onde

$$r_* := \min \left\{ \bar{a}, \frac{b}{\epsilon + 2\delta}, \frac{b(1-\gamma)}{2[(1-\gamma)(\epsilon + \delta) + 2c\bar{a}]} \right\}.$$

Considere a seguinte aplicação

$$\Phi_{x_0}(y) := L_{f+F}(x_*, f(x_*) - f(x_0) - B_0(y - x_0) + f'(x_*)(y - x_*))^{-1}.$$

Segue da definição de Φ_{x_0} que

$$d(x_*, \Phi_{x_0}(x_*)) = d(x_*, L_{f+F}(x_*, f(x_*) - f(x_0) - B_0(x_* - x_0))^{-1}).$$

Usando (3.29), (3.30) e r_* obtemos que

$$\begin{aligned} \|f(x_*) - f(x_0) - B_0(x_* - x_0)\| &\leq \|f(x_*) - f(x_0) - f'(x_*)(x_* - x_0)\| + \\ &\quad \|f'(x_*) - B_0\| \|x_* - x_0\| \\ &\leq (\epsilon + \delta) \|x_* - x_0\| \leq b. \end{aligned}$$

Então usando a Definição 2.6 e levando em conta que $0 \in L_{f+F}(x_*, x_*)$ obtemos

$$\begin{aligned} d(x_*, \Phi_{x_0}(x_*)) &\leq \kappa' d(f(x_*) - f(x_0) - B_0(x_* - x_0), L_{f+F}(x_*, x_*)) \\ &\leq \kappa' \|f(x_*) - f(x_0) - B_0(x_* - x_0)\| \\ &\leq \kappa' \|f(x_*) - f(x_0) - f'(x_*)(x_* - x_0)\| + \kappa' \|f'(x_*) - B_0\| \|x_* - x_0\| \\ &\leq \kappa'(\epsilon + \delta) \|x_* - x_0\| = \frac{\kappa'(\epsilon + \delta) \|x_* - x_0\|}{1 - \kappa\delta} (1 - \kappa\delta) = \rho(1 - \kappa\delta), \end{aligned}$$

onde

$$\rho := \frac{\kappa'(\epsilon + \delta) \|x_* - x_0\|}{1 - \kappa\delta}.$$

Agora tomamos $p, q \in B_\rho[x_*]$. Levando em conta (3.19) e que $\|x_* - x_0\| < r_*$, podemos verificar que $\rho < r_*$. Assim, para $s = p$ ou $s = q$ obtemos que

$$\begin{aligned} \|f(x_*) - f(x_0) - B_0(s - x_0) + f'(x_*)(s - x_*)\| &\leq \|f(x_*) - f(x_0) - f'(x_*)(x_* - x_0)\| \\ + \|f'(x_*)(x_* - x_0) - f'(x_*)(x_* - s) - B_0(s - x_0)\| &\leq (\epsilon + 2\delta) \|x_* - x_0\| \leq b. \end{aligned}$$

Uma vez que $e(\emptyset, \Phi_{x_0}(q)) = 0$, podemos assumir sem perda de generalidade que $\Phi_{x_0}(p) \cap B_a[x_*] \neq \emptyset$ para todo $p \in B_\rho[x_*]$. Seja $z \in \Phi_{x_0}(p) \cap B_a[x_*]$. Então, pela Definição 2.6 com $\bar{x} = x_*$, $\bar{u} = 0$, $x_0 = z$, $u = f(x_*) - f(x_0) - B_0(q - x_0) + f'(x_*)(q - x_*)$, e $G = L_{f+F}(x_*, \cdot)$, temos

$$d(z, \Phi_{x_0}(q)) \leq \kappa d(f(x_*) - f(x_0) - B_0(q - x_0) + f'(x_*)(q - x_*), L_{f+F}(x_*, z)).$$

Uma vez que $z \in \Phi_{x_0}(p)$ isto implica que $f(x_*) - f(x_0) - B_0(p - x_0) + f'(x_*)(p - x_*) \in L_{f+F}(x_*, z)$, usando a definição de distância dada em (2.6), obtemos

$$d(f(x_*) - f(x_0) - B_0(q - x_0) + f'(x_*)(q - x_*), L_{f+F}(x_*, z)) \leq \|[B_0 - f'(x_*)](p - q)\|.$$

Assim, combinando as duas últimas desigualdades podemos concluir que

$$d(z, \Phi_{x_0}(q)) \leq \kappa \|[B_0 - f'(x_*)](p - q)\|.$$

Tomando o supremo com respeito a $z \in \Phi_{x_0}(p) \cap B_a[x_*]$ na última desigualdade e usando a definição de excesso (2.6), temos

$$e(\Phi_{x_0}(p) \cap B_a[x_*], \Phi_{x_0}(q)) \leq \kappa \|[B_0 - f'(x_*)](p - q)\|.$$

Como $\rho < r_* \leq a$, temos $e(\Phi_{x_0}(p) \cap B_\rho[x_*], \Phi_{x_0}(q)) \leq e(\Phi_{x_0}(p) \cap B_a[x_*], \Phi_{x_0}(q))$. Assim, da última desigualdade, (3.30) e das propriedades de norma, obtemos

$$e(\Phi_{x_0}(p) \cap B_\rho[x_*], \Phi_{x_0}(q)) \leq \kappa \|B_0 - f'(x_*)\| \|p - q\| \leq \kappa\delta \|p - q\|.$$

Como $\kappa\delta < 1$, por (3.30), podemos aplicar o Teorema 2.3 com $\Phi = \Phi_{x_0}$, $\bar{x} = x_*$, e $\lambda = \kappa\delta$ para concluir que existe $y_0 \in \Phi_{x_0}(y_0)$ tal que

$$\|y_0 - x_*\| \leq \frac{\kappa'(\epsilon + \delta)}{1 - \kappa\delta} \|x_0 - x_*\|.$$

Se $y_0 \in C$, então tome $x_1 = y_0 \in C \cap B(x_*, r_*)$. Caso contrário, se $y_0 \notin C$ então tome $x_1 \in P_C(y_0, x_0, \theta_0)$. Aplicando o Lema 3.1 com $w = x_1$, $y = y_0$, $x = x_0$ e $\theta = \theta_0$, obtemos depois de algumas manipulações

$$\begin{aligned} \|x_1 - x_*\| &\leq (1 + \sqrt{2\theta_0})\|x_* - y_0\| + \sqrt{2\theta_0}\|x_* - x_0\| \\ &\leq \left[(1 + \sqrt{2\theta_0})\frac{\kappa'(\epsilon + \delta)}{1 - \kappa\delta} + \sqrt{2\theta_0} \right] \|x_* - x_0\| < \gamma \|x_0 - x_*\|. \end{aligned}$$

Uma vez que $\gamma < 1$, é verdade que $x_1 \in C \cap B(x_*, r_*)$. Por indução, suponha que existe um número inteiro $n > 1$, pontos y_1, \dots, y_n e x_1, \dots, x_n com $y_n \in B(x_*, r_*)$ e $x_n \in C \cap B_{r_*}(x_*)$, e ainda

$$\|y_{i-1} - x_*\| \leq \gamma \|x_{i-1} - x_*\| \quad \text{for } i = 1, \dots, n, \quad (3.21)$$

$$\|x_i - x_*\| \leq \gamma \|x_{i-1} - x_*\| \quad \text{for } i = 1, \dots, n. \quad (3.22)$$

Definamos a aplicação

$$\Phi_{x_n}(z) := L_{f+F}(x_*, f(x_*) - f(x_n) - B_n(z - x_n) + f'(x_*)(z - x_*))^{-1}.$$

Segue da definição de Φ_{x_n} que

$$d(x_*, \Phi_{x_n}(x_*)) = d(x_*, L_{f+F}(x_*, f(x_*) - f(x_n) - B_n(x_* - x_n))^{-1}).$$

Agora mostraremos que $\|f(x_*) - f(x_n) - B_n(x_* - x_n)\| \leq b$. Primeiramente, usando a desigualdade triangular temos

$$\begin{aligned} \|f(x_*) - f(x_n) - B_n(x_* - x_n)\| &\leq \|f(x_*) - f(x_n) - f'(x_*)(x_* - x_n)\| \\ &\quad + \|f'(x_*) - B_n\| \|x_* - x_n\| \end{aligned} \quad (3.23)$$

$$\leq \epsilon \|x_* - x_n\| + \|f'(x_*) - B_n\| \|x_* - x_n\| \quad (3.24)$$

Por outro lado, depois de simples manipulações obtemos de (3.18)

$$\begin{aligned} \|f'(x_*) - B_n\| &\leq \|f'(x_*) - B_{n-1}\| + c(\|x_{n-1} - x_*\| + \|y_{n-1} - x_*\|) \\ &\leq \|f'(x_*) - B_{n-2}\| + c(\|x_{n-2} - x_*\| + \|y_{n-2} - x_*\|) \\ &\quad + c(\|x_{n-1} - x_*\| + \|y_{n-1} - x_*\|) \\ &\quad \vdots \\ &\leq \|f'(x_*) - B_0\| + c \left(\sum_{k=0}^{n-1} \|x_k - x_*\| + \sum_{k=1}^n \|y_{k-1} - x_*\| \right). \end{aligned}$$

Combinando (3.21) e (3.22), a última desigualdade transforma-se em

$$\begin{aligned} \|f'(x_*) - B_n\| &\leq \|f'(x_*) - B_0\| + c \left(\sum_{k=0}^{n-1} \|x_k - x_*\| + \sum_{k=1}^n \|y_{k-1} - x_*\| \right) \\ &\leq \delta + 2c \sum_{k=0}^n \gamma^k \|x_0 - x_*\| \\ &\leq \delta + \frac{2c r_*}{1 - \gamma} \leq \delta + \frac{2ca}{1 - \gamma}. \end{aligned} \quad (3.25)$$

Assim, temos de (3.23), (3.25) e da condição em r_* que

$$\|f(x_*) - f(x_n) - B_n(x_* - x_n)\| \leq \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right) r_* \leq b.$$

Por usar a Definição 2.6 e levar em conta que $0 \in L_{f+F}(x_*, x_*)$ obtemos

$$\begin{aligned} d(x_*, \Phi_{x_n}(x_*)) &\leq \kappa' d(f(x_*) - f(x_n) - B_n(x_* - x_n), L_{f+F}(x_*, x_*)) \\ &\leq \kappa' \|f(x_*) - f(x_n) - B_n(x_* - x_n)\| \\ &\leq \kappa' \|f(x_*) - f(x_n) - f'(x_*)(x_* - x_n)\| + \kappa' \|f'(x_*) - B_n\| \|x_* - x_n\| \\ &\leq \kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right) \|x_* - x_n\| \\ &= \bar{\rho} \left[1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right) \right], \end{aligned}$$

onde

$$\bar{\rho} = \frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right) \|x_* - x_n\|}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)}.$$

Agora tomamos $p, q \in B_{\bar{\rho}}[x_*]$. Levando em conta a segunda desigualdade em (3.20) e $\|x_* - x_n\| < r_*$, podemos verificar que $\bar{\rho} < r_*$. Assim, não fica difícil verificar que para $v = p$ ou $v = q$ temos $\|f(x_*) - f(x_n) - B_n(v - x_n) + f'(x_*)(v - x_*)\| < b$.

Como $e(\emptyset, \Phi_{x_n}(q)) = 0$, podemos assumir que $\Phi_{x_n}(p) \cap B_a[x_*] \neq \emptyset$ para todo $p \in B_{\bar{\rho}}[x_*]$. Seja $y \in \Phi_{x_n}(p) \cap B_a[x_*]$. Então, pela Definição 2.6 com $\bar{x} = x_*$, $\bar{u} = 0$, $x_n = y$, $u = f(x_*) - f(x_n) - B_n(q - x_n) + f'(x_*)(q - x_*)$, e $G = L_{f+F}(x_*, \cdot)$, temos

$$d(y, \Phi_{x_n}(q)) \leq \kappa d(f(x_*) - f(x_n) - B_n(q - x_n) + f'(x_*)(q - x_*), L_{f+F}(x_*, y)).$$

Uma vez que $y \in \Phi_{x_n}(p)$, isto implica que $f(x_*) - f(x_n) - B_n(p - x_n) + f'(x_*)(p - x_*) \in L_{f+F}(x_*, y)$. Usando a definição de distância dada em (2.6), obtemos

$$d(f(x_*) - f(x_n) - B_n(q - x_n) + f'(x_*)(q - x_*), L_{f+F}(x_*, y)) \leq \|[B_n - f'(x_*)](p - q)\|.$$

Assim, combinando as duas últimas desigualdades podemos concluir que

$$d(y, \Phi_{x_n}(q)) \leq \kappa \|[B_n - f'(x_*)](p - q)\|.$$

Tomando o supremo com respeito a $z \in \Phi_{x_n}(p) \cap B_a[x_*]$ na última desigualdade e usando a definição de excesso dada em (2.6), temos

$$e(\Phi_{x_n}(p) \cap B_a[x_*], \Phi_{x_n}(q)) \leq \kappa \|[B_n - f'(x_*)](p - q)\|.$$

Como $\bar{\rho} < r_* \leq a$, Temos $e(\Phi_{x_n}(p) \cap B_{\bar{\rho}}[x_*], \Phi_{x_n}(q)) \leq e(\Phi_{x_n}(p) \cap B_a[x_*], \Phi_{x_n}(q))$. Assim, da última desigualdade e algumas propriedades da norma, obtemos

$$e(\Phi_{x_n}(p) \cap B_{\bar{\rho}}[x_*], \Phi_{x_n}(q)) \leq \kappa \|B_n - f'(x_*)\| \|p - q\| \leq \kappa \left(\delta + \frac{2ca}{1-\gamma} \right) \|p - q\|.$$

Por (3.19) temos $\kappa \left(\delta + \frac{2ca}{1-\gamma} \right) < 1$, então podemos aplicar o Teorema 2.3 com $\Phi = \Phi_{x_n}$, $\bar{x} = x_*$, e $\lambda = \kappa' \left(\delta + \frac{2ca}{1-\gamma} \right)$ para concluir que existe $y_n \in \Phi_{x_n}(y_n)$ e

$$\|y_n - x_*\| \leq \frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right)}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)} \|x_* - x_n\|.$$

Se $y_n \in C$, então tomamos $x_{n+1} = y_n \in C \cap B(x_*, r_*)$. Caso contrário, se $y_n \notin C$ então tomamos $x_{n+1} \in P_C(y_n, x_n, \theta_n)$. Aplicando o Lema 3.1 com $w = x_{n+1}$, $y = y_n$, $x = x_n$ e $\theta = \theta_n$, obtemos

$$\|x_* - x_{n+1}\| \leq \left[(1 + \sqrt{2\theta_n}) \frac{\kappa' \left(\epsilon + \delta + \frac{2ca}{1-\gamma} \right)}{1 - \kappa \left(\delta + \frac{2ca}{1-\gamma} \right)} + \sqrt{2\theta_n} \right] \|x_* - x_n\| < \gamma \|x_* - x_n\|.$$

e isso conclui a nossa demonstração. \square

3.1 Convergência para a atualização de Broyden

O Teorema 3.5 demonstrado acima, prova a convergência do algoritmo dando uma condição de deterioração limitada para B_k mas, até o momento, não fazemos nenhuma alusão a como tomaremos essas sequências de matrizes. Nesta seção consideraremos a atualização de Broyden definida por:

$$B_{k+1} := B_k + \frac{(z_k - B_k s_k) \langle s_k, \cdot \rangle}{\|s_k\|^2} \quad (3.26)$$

onde $z_k := f(y_k) - f(x_k)$ e $s_k := y_k - x_k$. Usualmente, B_0 é tomado como $f'(x_0)$.

Consideramos $\langle \cdot, \cdot \rangle$ o produto escalar usual de \mathbb{R}^n . Antes de enunciar o resultado relacionado ao Teorema utilizando a atualização de Broyden, será necessário enunciar alguns resultados técnicos.

Lema 3.2. *Seja $A \in \mathcal{L}(X, Y)$, onde X e Y são espaços de Hilbert. Se $x \in X \setminus \{0\}$, então*

$$\left\| A - \frac{\langle x, \cdot \rangle Ax}{\|x\|^2} \right\| = \begin{cases} 0, & \text{if } \dim X = 1; \\ \|A\|, & \text{if } \dim X > 1. \end{cases} \quad (3.27)$$

A prova do Lema, por ser muito técnica, será omitida mas pode ser encontrada em [3].

Proposição 3.1. *Suponha que a derivada Fréchet da função f é lipschitz contínua com constante L em uma vizinhança U de um ponto x_* . Dada $B_k \in \mathcal{L}(X, Y)$ e $x_k, y_k \in U$, com $y_k \neq x_k$, se B_{k+1} é definida como (3.26), então*

$$\|B_{k+1} - f'(x_*)\| \leq \|B_k - f'(x_*)\| + \frac{L}{2}(\|y_k - x_*\| + \|x_* - x_k\|). \quad (3.28)$$

Demonstração. Como assumimos,

$$\|f'(u) - f'(v)\| \leq L\|u - v\| \quad \text{para todo } u, v \in U.$$

Se $y_k, x_k \in U$, $x_k \neq y_k$ e B_{k+1} for definida como (3.26). Então

$$\begin{aligned} B_{k+1} - f'(x_*) &= B_k - f'(x_*) + \frac{(z_k - B_k s_k) \langle s_k, \cdot \rangle}{\|s_k\|^2} \\ &= B_k - f'(x_*) - \frac{(B_k - f'(x_*)) s_k \langle s_k, \cdot \rangle}{\|s_k\|^2} + \frac{(z_k - f'(x_*) s_k) \langle s_k, \cdot \rangle}{\|s_k\|^2}. \end{aligned}$$

Assim,

$$\|B_{k+1} - f'(x_*)\| \leq \left\| B_k - f'(x_*) - \frac{(B_k - f'(x_*)) s_k \langle s_k, \cdot \rangle}{\|s_k\|^2} \right\| + \frac{\|z_k - f'(x_*) s_k\|}{\|s_k\|}.$$

Pelo Lema 3.2,

$$\left\| B_k - f'(x_*) - \frac{(B_k - f'(x_*)) s_k \langle s_k, \cdot \rangle}{\|s_k\|^2} \right\| \leq \|B_k - f'(x_*)\|.$$

Utilizando o Teorema do Valor Médio, obtemos

$$\begin{aligned} \|z_k - f'(x_*) s_k\| &= \|f(y_k) - f(x_k) - f'(x_*) s_k\| \\ &= \left\| \int_0^1 [f'(x_k + t(y_k - x_k))(y_k - x_k) - f'(x_*) s_k] dt \right\| \\ &\leq \|s_k\| \int_0^1 \|f'(x_k + t(y_k - x_k)) - f'(x_*)\| dt \\ &\leq L \|s_k\| \int_0^1 ((1-t)\|x_k - x_*\| + t\|y_k - x_*\|) dt \\ &= \frac{L}{2} \|s_k\| (\|y_k - x_*\| + \|x_k - x_*\|) \end{aligned}$$

que conclui a demonstração. □

Agora podemos usar o Teorema 3.5, combinado com os resultados acima, e enunciar o seguinte resultado:

Teorema 3.6. *Se $\Omega \subset \mathbb{R}^n$ um conjunto aberto, $f : \Omega \rightarrow \mathbb{R}^m$ uma função definida em Ω , e $F : \Omega \rightrightarrows \mathbb{R}^m$ uma aplicação ponto-conjunto com gráfico fechado e $f + F$ metricamente regular, em x_* para 0, com constantes $\kappa > 0$, $a > 0$ e $b > 0$. Assuma que $C \subset \Omega$ é um conjunto fechado e convexo, $x_* \in C$, $0 \in f(x_*) + F(x_*)$, existe $\epsilon > 0$ tal que*

$$\|f(u) - f(v) - f'(x_*)(u - v)\| \leq \epsilon \|u - v\|, \quad \text{para todo } u, v \in B_{\bar{a}}(x_*). \quad (3.29)$$

Suponha que a derivada Fréchet da função f é Lipschitz contínua com constante L em uma vizinhança U de um ponto x_ , ou seja,*

$$\|f'(u) - f'(v)\| \leq L \|u - v\| \quad \text{para todo } u, v \in U.$$

Considere o método Quasi-Newton-Inexp com $\theta_n \geq 0$ e assuma que

$$\delta := \|B_0 - f'(x_*)\| < \frac{1}{2\kappa}. \quad (3.30)$$

Além do mais, assumimos que existem uma constante $c > 0$ e uma vizinhança U de \bar{x} tal que, para $k = 0, 1, \dots$, e para cada B_k e cada $x_k, y_k \in U$, $x_k \neq y_k$, todos satisfazendo (3.14), o operador B_{k+1} é escolhido tal que

$$B_{k+1} := B_k + \frac{(z_k - B_k s_k) \langle s_k, \cdot \rangle}{\|s_k\|^2} \quad (3.31)$$

onde $z_k := f(y_k) - f(x_k)$ e $s_k := y_k - x_k$. Então, existe uma vizinhança \mathcal{U} de x_ , tal que para cada $x_0 \in C \cap \mathcal{U} \setminus \{x_*\}$, existe uma sequência $\{x_k\}$ gerada pelo algoritmo Quasi-Newton-Inexp que resolve (1.1), associado a $\{\theta_k\}$ e iniciando em x_0 , que está contido em $C \cap \mathcal{U}$ e converge linearmente para x_* .*

4 Algoritmo inexato

Ao mostrar a convergência do algoritmo do capítulo anterior, pensamos imediatamente em ver, se na prática, o algoritmo funcionaria bem. Mas o primeiro ponto é pensar que a exatidão do algoritmo anterior possa ser um impasse para a máquina que sempre trabalha com tolerâncias. Pensando nisso e fazendo adaptações e contribuições para o trabalho [11], proporemos nesse capítulo um algoritmo inexato para resolver o problema (1.1).

Para entender melhor o que fizemos, vamos relembrar o Método de Newton Inexato para resolver Equações Generalizadas que foi introduzido por Dembo, Eisenstat e Steihaug em [15]. Tal método consiste em resolver aproximadamente a equação $f(x) = 0$ para $X = Y = \mathbb{R}$ da seguinte maneira: Dados uma sequência de escalares positivos η_k e iniciando em um ponto x_0 o iterando $(k + 1)$ é escolhido para satisfazer a condição

$$\|f(x_k) + f'(x_k)(x_{k+1} - x_k)\| \leq \eta_k \|f(x_k)\| \quad (4.1)$$

Seguindo a teoria desenvolvida em [16] e ajustando ao que fez [15] podemos reescrever a equação acima como

$$(f(x_k) + f'(x_k)(x_{k+1} - x_k)) \cap B_{\eta_k \|f(x_k)\|}(0) \neq \emptyset \quad (4.2)$$

E o que foi investigado, no final, foi o uso do método de Newton Inexato para Equações Generalizadas da seguinte forma

$$(f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset, \quad k = 0, 1, \dots, \quad (4.3)$$

E seguindo o que viemos fazendo até agora, sem esquecer que estamos trabalhando com o problema com restrições, propomos o método Quasi-Newton Inexato descrito por:

$$(f(x_k) + B_k(y_k - x_k) + F(y_k)) \cap R_k(x_k, y_k) \neq \emptyset, \quad k = 0, 1, \dots, \quad (4.4)$$

onde $\{B_k\}$ é uma sequência de matrizes e $R_k(x_k, y_k)$ é uma sequência de aplicações ponto-conjunto com gráfico fechado. Para ilustrar a flexibilidade desses critérios de inexatidão, observamos que quando F é a aplicação nula, η_k é uma sequência de escalares positivos e $R_k(x_k, y_k) := B_{\eta_k \|f(x_k)\|}(0)$, é recuperada uma instância do método Quasi-Newton inexato estudado por [15] para solucionar equações. Também, se $R_k(x_k, y_k) = \{-r_k(x_k)\}$ onde $\{r_k\}$

é uma sequência de funções que representa a inexatidão, nosso método proposto coincide com uma instância do método Quasi-Newton inexato estudado por [17].

Analogamente ao capítulo anterior, propomos um algoritmo formal para o Método Quasi-Newton Inexato com Projeção Inexata que é dado a seguir.

IQNM-InexP Method

Step 0. Seja $x_0 \in C$ e $\{\theta_k\} \subset [0, +\infty)$ dados, e tome $k = 0$.

Step 1. Se $f(x_k) + F(x_k) \ni 0$, então **pare**; caso contrário, compute $y_k \in \mathbb{R}^n$ tal que

$$(f(x_k) + B_k(y_k - x_k) + F(y_k)) \cap R_k(x_k, y_k) \neq \emptyset. \quad (4.5)$$

Step 2. Se $y_k \in C$, tome $x_{k+1} = y_k$; caso contrário escolha algum $x_{k+1} \in C$ satisfazendo

$$x_{k+1} \in P_C(y_k, x_k, \theta_k). \quad (4.6)$$

Step 3. Faça $k \leftarrow k + 1$, e vá para o **Step 1**.

Como no capítulo anterior, analisaremos aqui a sequência gerada pelo algoritmo. Aqui, também, usaremos uma ideia análoga para mostrar sua convergência mas, como tratamos com inexatidão, em vez do Teorema das Contrações, ou Teorema do Ponto Fixo Generalizado usaremos o Teorema da Coincidência, o qual enunciaremos a seguir e também pode ser encontrado em [4].

Teorema 4.1. *Sejam X e Y dois espaços métricos e considere as aplicações ponto-conjunto $\Phi : X \rightrightarrows Y$ e $\Gamma : Y \rightrightarrows X$. Se $\bar{x} \in X$ e $\bar{y} \in Y$ e sejam c, κ e μ escalares positivos tais que $\mu\kappa < 1$. Suponha que um dos conjuntos $\text{gph } \Phi \cap (B_c[\bar{x}] \times B_{c/\mu}[\bar{y}])$ e $\text{gph } \Gamma \cap (B_{c/\mu}[\bar{y}] \times B_c[\bar{x}])$ é fechado enquanto o outro é completo, ou ambos os conjuntos $\text{gph } (\Phi \circ \Gamma) \cap (B_c[\bar{x}] \times B_c[\bar{x}])$ e $\text{gph } (\Gamma \circ \Phi) \cap (B_{c/\mu}[\bar{y}] \times B_{c/\mu}[\bar{y}])$ são completos. Também, suponha que valham as seguintes condições:*

$$(i) \quad d(\bar{y}, \Phi(\bar{x})) < c(1 - \kappa\mu)/(2\mu);$$

$$(ii) \quad d(\bar{x}, \Gamma(\bar{y})) < c(1 - \kappa\mu)/2;$$

$$(iii) \quad e(\Phi(p) \cap B_{c/\mu}[\bar{y}], \Phi(q)) \leq \mu\rho(p, q) \text{ para todo } p, q \in B_c[\bar{x}] \text{ tais que } \rho(p, q) \leq c(1 - \kappa\mu)/\mu ;$$

(iv) $e(\Gamma(p) \cap B_c[\bar{x}], \Gamma(q)) \leq \kappa\rho(p, q)$ para todo $p, q \in B_{c/\mu}[\bar{y}]$ tais que $\mu(p, q) \leq c(1 - \kappa\mu)$

Então existem $\hat{x} \in B_c[\bar{x}]$ e $\hat{y} \in B_{c/\mu}[\bar{y}]$ tais que $\hat{y} \in \Phi(\hat{x})$ e $\hat{x} \in \Gamma(\hat{y})$. Se a aplicação $B_c[\bar{x}] \ni x \mapsto \Phi(x) \cap B_{c/\mu}[\bar{y}]$ e $B_{c/\mu}[\bar{y}] \ni y \mapsto \Gamma(y) \cap B_c[\bar{x}]$ são ponto a ponto, então os pontos \hat{x} e \hat{y} são únicos em $B_c[\bar{x}]$ e $B_{c/\mu}[\bar{y}]$, respectivamente.

Aplicando o Teorema 4.1 e outros resultados trabalhados nos capítulos anteriores, podemos enunciar um resultado de convergência para a sequência gerada pelo algoritmo acima.

Teorema 4.2. *Seja λ e μ duas constantes positivas tais que $\lambda\mu < 1$. Suponha que a aplicação $f + F$ é metricamente regular em \bar{x} para 0 com constante λ . Considere o método IQNM-InexP com θ_k uma sequência de números reais não crescente e assuma que*

$$\hat{\theta} := \sup \theta_k < \frac{1}{2}, \quad \|B_0 - f'(\bar{x})\| < 1/(2\lambda). \quad (4.7)$$

Além disso, assuma que existe uma constante positiva ι e uma vizinhança U de \bar{x} tal que, para $k = 0, 1, \dots$, e para cada B_k e cada $x_k, y_k \in U$, $x_k \neq y_k$, satisfazendo (4.4), o operador B_{k+1} é escolhido de maneira que

$$\|B_{k+1} - f'(\bar{x})\| \leq \|B_k - f'(\bar{x})\| + \iota(\|x_k - \bar{x}\| + \|y_k - \bar{x}\|). \quad (4.8)$$

Também, suponha que para cada $k = 0, 1, \dots$, a aplicação $(u, x) \mapsto R_k(u, x)$ é parcialmente Aubin Contínua com respeito a x em \bar{x} para 0 uniformemente em u em torno de \bar{x} com constante μ . E ainda, suponha que existam escalares positivos $\gamma < (1 - \lambda\mu)/2\mu$ e β tais que

$$d(0, R_k(u, \bar{x})) \leq \gamma\|u - \bar{x}\| \quad \text{para todo } u \in \mathcal{B}_\beta(\bar{x}) \quad \text{e todo } k = 0, 1, \dots \quad (4.9)$$

Então existe uma vizinhança \mathcal{O} de \bar{x} tal que, para cada ponto inicial $x_0 \in \mathcal{O}$ existe uma sequência $\{x_k\}$ gerada pelo método IQNM-InexP, que está contida em \mathcal{O} e que converge q -linearmente para \bar{x} .

Demonstração. Tome $t \in \left(0, \frac{1 - \sqrt{2\hat{\theta}}}{1 + \sqrt{2\hat{\theta}}}\right)$ tal que $0 < \gamma < t(1 - \lambda\mu)/2\mu$ e seja

$$\delta := \|B_0 - f'(\bar{x})\|.$$

Sejam a e b constantes em (2.9) e que elas satisfazem

$$e(R_k(u, x) \cap \mathcal{B}_b[0], R_k(u, x')) \leq \mu\|x - x'\| \quad \text{para todo } u, x, x' \in \mathcal{B}_a[\bar{x}]. \quad (4.10)$$

Uma vez que f é estritamente diferenciável em \bar{x} , podemos tomar a tão pequeno, se necessário, a fim de que

$$\|f(x) - f(\bar{x}) - f'(\bar{x})(x - \bar{x})\| \leq \epsilon\|x - \bar{x}\|, \quad \text{para todo } x \in \mathcal{B}_a[\bar{x}], \quad (4.11)$$

e

$$\left(\delta + \frac{2\iota a}{1 - \tilde{t}_0} \right) \lambda < 1, \quad (4.12)$$

para algum real positivo ϵ e $\tilde{t}_k := [(1 + \sqrt{2\theta_k})t + \sqrt{2\theta_k}]$, para todo $k \geq 0$. Não é difícil ver que $0 < \tilde{t}_k \leq \tilde{t}_0 < 1$. Também, podemos escolher uma constante positiva κ tal que

$$\begin{aligned} \kappa &> \frac{\lambda}{1 - \lambda \left(\delta + \frac{2\iota a}{1 - \tilde{t}_0} \right)} := \varpi, \\ \epsilon + \delta + \frac{2\iota a}{1 - \tilde{t}_0} &< \frac{t(1 - \kappa\mu)}{2\kappa}. \end{aligned} \quad (4.13)$$

e $\gamma < t(1 - \kappa\mu)/2\mu$. Em seguida, aplicaremos a Proposição 2.1. Supomos que

$$G_{\bar{x}}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + F(x), \quad (4.14)$$

para $x \in X$. Também, considere a função

$$g(x_0, x) = f(x_0) + B_0(x - x_0) - f(\bar{x}) - f'(\bar{x})(x - \bar{x}).$$

Fixe $r_* > 0$ satisfazendo

$$r_* \leq \min \left\{ a, \frac{b}{\left(\epsilon + \delta + \frac{2\iota a}{1 - \tilde{t}_0} \right)}, \beta, b\mu \right\}. \quad (4.15)$$

Denote $\mathcal{O} := \mathcal{B}_{r_*}[\bar{x}]$ e escolha algum $x_0 \in \mathcal{O}$ tal que $x_0 \neq \bar{x}$. O próximo passo é verificar as condições do Teorema 4.1 com $\Phi(x) = R_0(x_0, x)$, $\Gamma = G_{x_0}^{-1}$, $\bar{y} := 0$ e $c := t\|x_0 - \bar{x}\|$. Então, utilizando que $x_0 \in \mathcal{B}_{r_*}[\bar{x}]$, $r_* \leq a$, e que $(\epsilon + \delta)r_* \leq b$ obtemos de (4.25) que

$$\begin{aligned} \|f(\bar{x}) - f(x_0) - B_0(\bar{x} - x_0)\| &\leq \|f(\bar{x}) - f(x_0) - f'(\bar{x})(\bar{x} - x_0)\| + \|f'(\bar{x}) - B_0\|\|\bar{x} - x_0\| \\ &\leq (\epsilon + \delta)r_* \leq b. \end{aligned}$$

Uma vez que $0 \in f(\bar{x}) + F(\bar{x})$ temos

$$-f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0) \in L_{f+F}(x_0, \bar{x}) \cap \mathcal{B}_b[0]. \quad (4.16)$$

Então, como $\kappa > \varpi$ e $\lambda\delta = \lambda\|B_0 - f'(\bar{x})\| < 1$, podemos aplicar a Proposição 2.1 com $u := x_0$, $\zeta := \delta$, $y := -f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0)$, $y' := 0$, e $a := r_*$ para concluir que

$$\begin{aligned} e \left(L_{f+F}(x_0, -f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0))^{-1} \cap \mathcal{B}_{r_*}[\bar{x}], L_{f+F}(x_0, 0)^{-1} \right) &\leq \\ &\kappa \| -f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0) \|. \end{aligned}$$

Agora, usamos a definição de excesso em (2.6) e (4.16) para obter

$$d(\bar{x}, L_{f+F}(x_0, 0)^{-1}) \leq \kappa \| -f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0) \|.$$

Depois de algumas manipulações e usando (4.13), (4.25) e (4.11), a desigualdade acima implica em

$$\begin{aligned} d(\bar{x}, L_{f+F}(x_0, 0)^{-1}) &\leq \kappa \| -f(\bar{x}) + f(x_0) + B_0(\bar{x} - x_0) \| \\ &\leq \kappa \| f(\bar{x}) - f(x_0) - f'(\bar{x})(\bar{x} - x_0) \| + \kappa \| f'(\bar{x}) - B_0 \| \| \bar{x} - x_0 \| \\ &\leq \kappa(\epsilon + \delta) \| \bar{x} - x_0 \| \\ &< \frac{t(1 - \kappa\mu) \| \bar{x} - x_0 \|}{2} = \frac{c(1 - \kappa\mu)}{2}. \end{aligned}$$

Por outro lado, como $\|x_0 - \bar{x}\| \leq r_* \leq \beta$, $t \in (0, 1)$, $\lambda\mu < 1$ e $\gamma < (1 - \kappa\mu)/(2\mu)$ obtemos de (4.27) que

$$d(0, R_0(x_0, \bar{x})) \leq \gamma \|x_0 - \bar{x}\| < \frac{t(1 - \kappa\mu)}{2\mu} \|x_0 - \bar{x}\| = \frac{c(1 - \kappa\mu)}{2\mu}. \quad (4.17)$$

Assim, as condições (i) e (ii) do Teorema 4.1 são satisfeitas. Agora, uma vez que $x_0 \in \mathcal{B}_{r_*}[\bar{x}]$, obtemos que $c \leq r_* \leq a$ e $c/\mu \leq b$. Assim, temos $\mathcal{B}_{c/\mu}[0] \subset \mathcal{B}_b[0]$ e $\mathcal{B}_c[\bar{x}] \subset \mathcal{B}_a[\bar{x}]$. Dessa forma, para qualquer $x, x' \in \mathcal{B}_a[\bar{x}]$ a condição (4.10) nos dá

$$e(R_0(x_0, x) \cap \mathcal{B}_{c/\mu}[0], R_0(x_0, x')) \leq e(R_0(x_0, x) \cap \mathcal{B}_b[0], R_0(x_0, x')) \leq \mu \|x - x'\|,$$

ou seja, a condição (iii) do Teorema 4.1 é satisfeita. Além do mais, de (2.9) concluímos que

$$\begin{aligned} e(L_{f+F}(x_0, x)^{-1} \cap \mathcal{B}_c[\bar{x}], L_{f+F}(x_0, x')^{-1}) &\leq e(L_{f+F}(x_0, x)^{-1} \cap \mathcal{B}_a[\bar{x}], L_{f+F}(x_0, x')^{-1}) \\ &\leq \kappa \|x - x'\| \end{aligned}$$

para todo $x, x' \in \mathcal{B}_{c/\mu}[0] \subset \mathcal{B}_b[0]$. Ou seja, a condição (iv) do Teorema 4.1 vale. Portanto, podemos aplicar o Teorema 4.1 obtendo que existe $y_0 \in \mathcal{B}_c[\bar{x}]$ e $v_1 \in \mathcal{B}_{c/\mu}[0]$ tais que $y_0 \in L_{f+F}(x_0, v_1)^{-1}$ e $v_1 \in R_0(x_0, y_0)$, ou seja,

$$v_1 \in L_{f+F}(x_0, y_0) \cap R_0(x_0, y_0) = (f(x_0) + B_0(y_0 - x_0) + F(y_0)) \cap R_0(x_0, y_0). \quad (4.18)$$

Além do mais, o Teorema 4.1 implica que $\|y_0 - \bar{x}\| \leq t\|x_0 - \bar{x}\|$. A inclusão (4.18) valida o fato de y_0 satisfazer (4.4) para $k = 0$. Em particular, desde que $t < 1$, $y_0 \in \mathcal{O} = \mathcal{B}_{r_*}[\bar{x}]$. Se $y_0 \in C \cap \mathcal{B}_c[\bar{x}]$ então tome $x_1 = y_0$. Caso contrário, tome $x_1 \in P_C(y_0, x_0, \theta_0)$. Assim, usando o Lema 3.1 com $w = x_1$, $y = y_0$, $x = x_0$ e $\theta = \theta_0$, obtemos:

$$\begin{aligned} \|x_1 - \bar{x}\| &\leq (1 + \sqrt{2\theta_0})\|\bar{x} - y_0\| + \sqrt{2\theta_0}\|\bar{x} - x_0\| \\ &= [(1 + \sqrt{2\theta_0})t + \sqrt{2\theta_0}]\|\bar{x} - x_0\| = \tilde{t}_0\|x_0 - \bar{x}\|. \end{aligned}$$

Como $\tilde{t}_0 < 1$, $x_1 \in C \cap \mathcal{B}_c[\bar{x}]$.

Por indução, supomos que existe um natural $k > 1$ e pontos $x_1, x_2, \dots, x_k \in \mathcal{B}_{r_*}[\bar{x}] \cap C$ e $y_1, y_2, \dots, y_k \in \mathcal{B}_{r_*}[\bar{x}]$, todos eles satisfazendo

$$\|y_{j-1} - \bar{x}\| \leq t\|x_{j-1} - \bar{x}\|, \quad \forall j = 1, 2, \dots, k. \quad (4.19)$$

$$\|x_j - \bar{x}\| \leq \tilde{t}_0 \|x_{j-1} - \bar{x}\|, \quad \forall j = 1, 2, \dots, k. \quad (4.20)$$

Sem perda de generalidade, assumimos que $y_{j-1} \neq x_{j-1} \neq \bar{x}$. Observe que, como $x_k \in \mathcal{B}_{r_*}[\bar{x}]$ e $r_* \leq \beta$, podemos repetir o mesmo argumento de (4.17) com x_0 substituindo por x_k . Assim, obtemos

$$d(0, R_k(x_k, \bar{x})) \leq \frac{c(1 - \kappa\mu)}{2\mu}. \quad (4.21)$$

Ao aplicar a Proposição 2.1 no passo de indução, primeiramente precisamos mostrar que $\|B_k - f'(\bar{x})\| \leq \zeta$ para algum escalar positivo ζ tal que $\zeta\lambda < 1$. Ao combinar (4.8), (4.20), (4.19) e usando que $x_0 \in \mathcal{B}_{r_*}[\bar{x}]$, temos

$$\begin{aligned} \|f'(\bar{x}) - B_k\| &\leq \|f'(\bar{x}) - B_{k-1}\| + \iota(\|x_{k-1} - \bar{x}\| + \|y_{k-1} - \bar{x}\|) \\ &\leq \|f'(\bar{x}) - B_0\| + \iota \sum_{j=1}^k (\|x_j - \bar{x}\| + \|y_j - \bar{x}\|) \\ &\leq \|f'(\bar{x}) - B_0\| + \iota \sum_{j=1}^k (\|x_j - \bar{x}\| + t\|x_j - \bar{x}\|) \\ &\leq \delta + (1+t)\iota \sum_{j=0}^k \|x_j - \bar{x}\| \\ &\leq \delta + 2\iota \sum_{j=0}^{\infty} \tilde{t}_0^j \|x_0 - \bar{x}\| \\ &\leq \delta + \frac{2\iota a}{1 - \tilde{t}_0}, \end{aligned}$$

ou seja, para todo $k \geq 1$

$$\|f'(\bar{x}) - B_k\| \leq \delta + \frac{2\iota a}{1 - \tilde{t}_0}. \quad (4.22)$$

A estimativa acima combinada com (4.11) implica em

$$\begin{aligned} \|-f(\bar{x}) + f(x_k) + B_k(x_k - \bar{x})\| &\leq \|-f(\bar{x}) + f(x_k) + f'(\bar{x})(\bar{x} - x_k)\| \\ &\quad + \|f'(\bar{x}) - B_k\| \|\bar{x} - x_k\| \\ &\leq \epsilon \|\bar{x} - x_k\| + \left(\delta + \frac{2\iota a}{1 - \tilde{t}_0} \right) \|\bar{x} - x_k\| \\ &\leq \left(\epsilon + \delta + \frac{2\iota a}{1 - \tilde{t}_0} \right) r_* \leq b. \end{aligned}$$

Assim, $-f(\bar{x}) + f(x_k) + B_k(x_k - \bar{x}) \in \mathcal{B}_b[0] \cap G_{x_k}(\bar{x})$. Em particular, $\bar{x} \in G_{x_k}^{-1}(-f(\bar{x}) + f(x_k) + B_k(x_k - \bar{x}))$. Então, usando (4.22), (4.12), (4.13) e a última desigualdade, podemos aplicar a Proposição 2.1 para obter

$$\begin{aligned} d(\bar{x}, G_{x_k}^{-1}(0)) &\leq \kappa \left(\epsilon + \delta + \frac{2\iota a}{1 - \tilde{t}_0} \right) \|\bar{x} - x_k\| \\ &\leq \frac{t(1 - \kappa\mu)}{2} \|\bar{x} - x_k\| = \frac{c(1 - \kappa\mu)}{2}. \end{aligned}$$

Com isso as condições (i) e (ii) do Teorema 4.1 são satisfeitas com $u = x_k$. Agora, sabendo que $x_k \in \mathcal{B}_{r_*}[\bar{x}]$, $c = t\|x_k - \bar{x}\|$ e por (4.15), obtemos $c \leq r_* \leq a$ e $c/\mu \leq b$, assim (4.10) implica na validade da condição (iii). Além disso, de (2.9) concluímos que a condição (iv) do Teorema 4.1 vale para a aplicação $\Gamma = G_{x_j}^{-1}$ com $j = k$. Assim, as condições do Teorema 4.1 são válidas com $c = t\|x_k - \bar{x}\|$. Por isso, segue que existem $y_k \in \mathcal{B}_c[\bar{x}]$ e $v_{k+1} \in \mathcal{B}_{c/\mu}[0]$ tais que $y_k \in G_{x_k}^{-1}(v_{k+1})$ e $v_{k+1} \in R_k(x_k, y_k)$, ou seja,

$$v_{k+1} \in G_{x_k}(y_k) \cap R_k(x_k, y_k) = (f(x_k) + B_k(y_k - x_k) + F(y_k)) \cap R_k(x_k, y_k). \quad (4.23)$$

Além do mais, o Teorema 4.1 implica que $\|y_k - \bar{x}\| \leq t\|x_k - \bar{x}\|$. A inclusão (4.23) implica que y_k satisfaz (4.4) para todo $k \geq 0$. Desde que $t < 1$, $y_k \in \mathcal{O} = \mathcal{B}_{r_*}[\bar{x}]$.

Novamente, se $y_k \in C \cap \mathcal{B}_c[\bar{x}]$ escolha $x_{k+1} = y_k$.

Caso contrário, tome $x_{k+1} \in P_{C \cap \mathcal{B}_c[\bar{x}]}(y_k, x_k, \theta_k)$. Dessa forma, usando o Lema 3.1 com $w = x_{k+1}$, $y = y_k$, $x = x_k$ e $\theta = \theta_k$, obtemos:

$$\|x_{k+1} - \bar{x}\| \leq [(1 + \sqrt{2\theta_k})t + \sqrt{2\theta_k}]\|\bar{x} - x_k\| < \tilde{t}\|x_k - \bar{x}\|. \quad (4.24)$$

Como $\tilde{t} < 1$ então $x_{k+1} \in \mathcal{O} = \mathcal{B}_{r_*}[\bar{x}]$ o que conclui a demonstração. \square

Da mesma forma como fizemos no capítulo anterior, pensamos imediatamente em quem poderia ser essa atualização B_k e de forma análoga, percebemos que a atualização de Broyden poderia ser adaptada de forma semelhante.

Então, de posse do Teorema 4.2 e da Proposição 3.1 podemos pensar em um teorema análogo ao Teorema 3.6, mas para o caso inexato. Assim temos:

Teorema 4.3. *Seja λ e μ duas constantes positivas tais que $\lambda\mu < 1$. Suponha que a aplicação $f + F$ é metricamente regular em \bar{x} para 0 com constante λ . Considere o método IQNM-InexP com θ_k uma sequência de números reais não crescente e assuma que*

$$\hat{\theta} := \sup \theta_k < \frac{1}{2}, \quad \|B_0 - f'(\bar{x})\| < 1/(2\lambda). \quad (4.25)$$

Suponha que para $k = 0, 1, \dots$, e para cada B_k e cada $x_k, y_k \in U$, $x_k \neq y_k$, satisfazendo (4.4), teremos a derivada Fréchet da função f lipschitz contínua com constante L em uma vizinhança U de um ponto x_ , ou seja,*

$$\|f'(u) - f'(v)\| \leq L\|u - v\| \quad \text{para todo } u, v \in U.$$

Dessa forma, podemos tomar

$$B_{k+1} := B_k + \frac{(z_k - B_k s_k) \langle s_k, \cdot \rangle}{\|s_k\|^2} \quad (4.26)$$

onde $z_k := f(y_k) - f(x_k)$ e $s_k := y_k - x_k$. Também, suponha que para cada $k = 0, 1, \dots$, a aplicação $(u, x) \mapsto R_k(u, x)$ é parcialmente Aubin Contínua com respeito a x em \bar{x} para 0

uniformemente em u em torno de \bar{x} com constante μ . E ainda, suponha que existe escalares positivos $\gamma < (1 - \lambda\mu)/2\mu$ e β tais que

$$d(0, R_k(u, \bar{x})) \leq \gamma \|u - \bar{x}\| \quad \text{para todo } u \in \mathcal{B}_\beta(\bar{x}) \quad \text{e todo } k = 0, 1, \dots. \quad (4.27)$$

Então existe uma vizinhança \mathcal{O} de \bar{x} tal que, para cada ponto inicial $x_0 \in \mathcal{O}$ existe uma sequência $\{x_k\}$ gerada pelo método IQNM-InexP, que está contida em \mathcal{O} e que converge q -linearmente para \bar{x} .

5 Aplicações e Testes Numéricos

Este capítulo será dividido em duas seções. Na primeira, falaremos sobre algumas aplicações do problema. Essas aplicações foram fundamentais para instigar a abordagem a esse problema restrito que teve uma grande abordagem em [1]. Vamos separar as aplicações analisadas nesse capítulo e tentar mostrar sua importância na medida do alcance deste trabalho.

Na segunda parte, faremos alguns testes em problemas de pequeno porte, seguindo [3], mas incluindo a restrição ao problema. A falta de referências numéricas na literatura a cerca desse problema dificulta ainda mais esse tipo de abordagem.

5.1 Aplicações

5.1.1 Desigualdade Variacional com Restrição

Em terapia de radiação de intensidade modulada (IMRT), feixes de radiação penetrante são direcionados à lesão tumoral de fontes externas, ver [38]. Um *multileaf colimador* (MLC) é usado para dividir cada feixe em muitos *beamlets* com controle individual intensidades. Existem dois aspectos principais da teleterapia por radiação que requerem modelagem computacional. O primeiro é o cálculo da dose de radiação absorvida no tecido irradiado com base em uma determinada distribuição de intensidades de feixe. Este cálculo de dose é um problema avançado. O segundo aspecto é o problema inverso do primeiro: encontrar uma distribuição de intensidades de radiação (mapa de intensidade de radiação) entregues por todos os feixes, o que resultaria em uma distribuição de dose clinicamente aceitável (ou seja, de modo que a dose para cada tecido estaria dentro dos limites superior e inferior desejados, que são prescritos com base no diagnóstico médico, conhecimento e experiência).

Para ter valor prático, no entanto, este mapa de intensidade de radiação deve ser implementável, de uma forma clinicamente aceitável, na máquina de tratamento disponível. Portanto, além dos parâmetros físicos e biológicos do objeto irradiado, as informações relevantes sobre as capacidades e especificações do tratamento disponível da máquina (ou seja, fonte de radiação) deve ser levado em consideração. O conceito de dose uniforme equivalente (EUD) foi introduzido em estudos recentes para descrever distribuições de dose com maior relevância clínica. EUD é definido para tumores como a dose equivalente biológica que, se administrada uniformemente, levará à mesma morte celular no volume tumoral como a distribuição real não uniforme da dose.

Este assunto foi amplamente estudado em [31], onde foi proposto um algoritmo

de otimização que minimiza uma função de proximidade ponderada que mede a soma dos quadrados das distâncias aos conjuntos de restrições, dentre outros problemas. Ainda nesse trabalho foram apresentados resultados computacionais que demonstram a validade do modelo e o poder do esquema algorítmico proposto. A junção desse trabalho a [39] mediante algumas adaptações do estudo do método dos extragradientes feito em [40], foram a base para o estudo dos CVIP feito em [29].

O Problema de Desigualdade Variacional com Restrições, que já mencionamos em (5.1), foi bem abordado em [29], onde foi denominado "Constrained Variational Inequality Problem" ou, simplesmente, "CVIP". O problema de desigualdade variacional restrito, com as devidas adaptações ao nosso caso, é definido como:

$$\text{Encontrar } x^* \in U \cap \Omega \text{ tal que } \langle f(x^*), x - x^* \rangle \geq 0, \forall x \in U. \quad (5.1)$$

Considerando $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função contínua e U e Ω são conjuntos não vazios, fechados e convexos em \mathbb{R}^n . Nas seções anteriores mostramos como abordar o problema via Equação Generalizada com Restrição e agora vamos expor a forma em que o problema foi abordado anteriormente. Para a garantir a convergência do Teorema, as condições seguintes são necessárias:

Condição 1 f é monótona em U .

Condição 2 f é Lipschitz Contínua em \mathbb{R}^n com constante $\kappa > 0$.

Condição 3 $\Omega \cap SOL(U, f) \neq \emptyset$.

Seja $\{\lambda_k\}_{k=0}^\infty \subset [a, b]$ para algum $a, b \in (0, \frac{1}{\kappa})$ e seja $\{\alpha_k\}_{k=0}^\infty \subset [c, d]$ para algum $c, d \in (0, 1)$. Então o algoritmo seguinte gera duas sequencias que convergem para um ponto $z \in \Omega \cap SOL(U, f)$, como garante o Teorema exposto solo em seguida.

CVIP Algorithm

Inicialização Escolha um ponto inicial arbitrário $x^0 \in \mathbb{R}^n$.

Passo Iterativo. Dado um iterado x^k , compute:

$$y^k = P_U(x^k - \lambda_k f(x^k)), \quad (5.2)$$

Construir o semi espaço T_k cujo hiperplano delimitador suporta U em y^k ,

$$T_k := \{w \in \mathbb{R}^n \mid \langle (x^k - \lambda_k f(x^k)) - y^k, w - y^k \rangle \leq 0\} \quad (5.3)$$

E então, calcule o próximo iterando por:

$$x^{k+1} = \alpha_k x^k + (1 - \alpha_k) P_{\Omega}(P_{T_k}(x^k - \lambda_k f(y^k))) \quad (5.4)$$

Observação 5.1. *O algoritmo acima, embora não use a teoria desenvolvida neste trabalho, assemelha-se bastante ao algoritmo proposto no capítulo anterior e por isso enunciamos aqui.*

Teorema 5.1. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, e sejam U e Ω conjuntos não vazios, fechados e convexos de \mathbb{R}^n . Assuma que as condições 1-3 mencionadas acima são válidas e sejam $\{x_k\}_{k=0}^{\infty}$ e $\{y_k\}_{k=0}^{\infty}$ duas sequências geradas pelo CVIP Algorith com $\{\lambda_k\}_{k=0}^{\infty} \subset [a, b]$ para algum $a, b \in (0, \frac{1}{\kappa})$ e $\{\alpha_k\}_{k=0}^{\infty} \subset [c, d]$ para algum $c, d \in (0, 1)$. Então $\{x_k\}_{k=0}^{\infty}$ e $\{y_k\}_{k=0}^{\infty}$ convergem para o mesmo ponto $z \in \Omega \cap SOL(U, f)$ e*

$$z = \lim_{k \rightarrow \infty} P_{\Omega \cap SOL(U, f)}(x^k). \quad (5.5)$$

A demonstração desse Teorema foge ao objetivo desse trabalho, mas está feita em detalhes em [29]. Esse problema, assim definido, foi relacionado, na época, a outros trabalhos com em [41, p.288] e [42] mas não entraremos em detalhes.

5.1.2 Desigualdade Variacional Split

A clássica Desigualdade Variacional, como está em [43] (1963) e [44] (1964), foi bastante usada para o estudo de programação matemática, economia de rede, pesquisa de transporte e ciências regionais. Nas últimas décadas, a teoria e os algoritmos numéricos para desigualdades variacionais foram estudados extensivamente.

Nesta seção trabalharemos com a recém-introduzida desigualdade variacional "split" em [29], que consiste em duas desigualdades variacionais com uma restrição linear de ligação. O "SVI", como chamaremos 'split variational inequality', pode ser visualizado como uma generalização de uma desigualdade variacional, que é principalmente motivada para tratar uma gama de problemas reais mundiais decorrentes da recuperação de sinal e processamento de imagem, como o bem estudado (conjunto múltiplo) problema de viabilidade 'split' [45] [31] [46] [47] [48] [49] [50] [51] [52] [53] [54], o problema de viabilidade convexa [55] [56] [57] [58] [59], as soluções comuns para desigualdades variacionais [60] [61] [62] [63], a minimização 'split' (separável) e problemas de 'split' zero [64] [65] [66] [67] [68] [69] e outras aplicações, como pode ser visto também em [30].

Técnicas "splits" foram empregadas em diversas áreas. Por exemplo, na temática tratamento da terapia de radiação modulada por intensidade (IMRT) [38] [70], um

“multileaf collimator” é usado para dividir cada feixe em muitos beamlets com controle individual de intensidades. A dose de radiação então precisa ser otimizada, dada a distribuição de intensidades de feixe de luz e, inversamente, a distribuição de intensidades de radiação entregues a todos os beamlets são necessários para serem otimizados dentro de uma determinada dose clinicamente aceitável de distribuição (a dose para cada tecido deve estar dentro do desejado superior e inferior, limites prescritos com base no diagnóstico médico, conhecimento e experiência). Um modelo matemático para o IMRT é a chamada viabilidade “split” de conjuntos múltiplos problema [47] que pode ser resolvido por uma abordagem de desigualdade variacional que é, em voltas, um caso especial de uma desigualdade variacional “split”. Abaixo, enunciaremos tal problemas com as devidas adaptações ao nosso caso.

Seja $U \subset \mathbb{R}^n$ e $\Omega \subset \mathbb{R}^m$ conjuntos não vazios, fechados e convexos, e $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ um operador linear. Dadas funções $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ e $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$, o SVIP³ é formulado como: Encontrar um ponto $x_* \in U$ tal que

$$\langle f(x_*), x - x_* \rangle \geq 0, \quad \forall x \in U,$$

e tal que o ponto $y_* = Ax_* \in \Omega$ satisfaz

$$\langle g(y_*), y - y_* \rangle \geq 0, \quad \forall y \in \Omega.$$

Observe agora, como feito em [1] e em [29] podemos tomar $\mathbb{R}^{nm} := \mathbb{R}^n \times \mathbb{R}^m$, $D := U \times \Omega$ e $V := \{w = (x, y) \in \mathbb{R}^{nm} : Ax = y\}$ o SVIP é equivalente ao CVIP:

$$\text{Encontrar } w_* \in D \cap V \text{ tal que } \langle h(w_*), w - w_* \rangle \geq 0, \quad \forall w \in D, \quad (5.6)$$

onde $w = (x, y)$ e $h : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{nm}$ são definidos por $h(x, y) := (f(x), g(y))$. E dessa forma o problema foi resolvido seguindo o algoritmo mencionado na seção anterior, diferentemente do que propomos no Exemplo (2.3). Uma abordagem bem diferente foi explorada em [30], mas foge muito ao objetivo do nosso trabalho.

5.2 Testes Numéricos

O Algoritmo foi implementado em Matlab© R2016b, usando precisão dupla, sob o sistema GNU/Linux Ubuntu 20.04. Por simplicidade, supomos $\Omega = \mathbb{R}^n$. A fim de escrever a inclusão $\bar{z} \in F(x)$, supomos que o conjunto $F(x)$ é definido por equações/inequações, digamos,

$$F(x) = \{z \in \mathbb{R}^m \mid h(x, z) = 0, g(x, z) \leq 0\}, \quad (5.7)$$

de modo que

$$\bar{z} \in F(x) \Leftrightarrow h(x, \bar{z}) = 0, g(x, \bar{z}) \leq 0.$$

³ SVIP do inglês Split Variational Inequality Problem.

Com isso, determinar uma solução y_k do subproblema $f(x_k) + B_k(y - x_k) + F(y) \ni 0$ equivale a encontrar $y_k \in \mathbb{R}^n$ e $z_k \in \mathbb{R}^m$ tais que

$$z_k = -[f(x_k) + B_k(y_k - x_k)], \quad h(y_k, z_k) = 0, \quad g(y_k, z_k) \leq 0.$$

As expressões acima serão satisfeitas se (y_k, z_k) for solução do problema

$$\min_{y,z} \frac{1}{2} \|z + [f(x_k) + B_k(y - x_k)]\|^2 \quad (5.8)$$

$$h(y, z) = 0, \quad (5.9)$$

$$g(y, z) \leq 0 \quad (5.10)$$

com valor funcional nulo. Este problema pode ser resolvido, em tese, por algoritmos usuais de programação não linear. Nos testes, foi utilizado o método de pontos interiores implementado na rotina `fmincon`, com os seguintes parâmetros: tolerância para otimalidade 10^{-6} e número máximo de iterações 1000; gradientes das h_i 's e g_j 's são fornecidos.

Por sua vez, verificar a validade da equação generalizada $f(x_k) + F(x_k) \ni 0$ (passo 1) equivale a testar se

$$h(x_k, -f(x_k)) = 0, \quad g(x_k, -f(x_k)) \leq 0.$$

Assim, utilizamos a seguinte medida aproximada como critério de parada do algoritmo:

$$e_k = \max\{\|h(x_k, -f(x_k))\|_\infty, \|g(x_k, -f(x_k))_+\|_\infty\} \leq 10^{-6}.$$

Note que $x_k \in C$ por construção.

O iterando x_{k+1} do passo de projeção (passo 2), isto é, aquele que satisfaz $x_{k+1} \in P(y_k, x_k, \theta_k)$, é obtido resolvendo o problema convexo relativo à projeção exata de y_k sobre C (equivalente à tomar $\theta_k = 0$)

$$\min_x \frac{1}{2} \|x - y_k\|^2 \quad x \in C. \quad (5.11)$$

Novamente, utilizamos o método de pontos interiores do Matlab presente da rotina `fmincon`. Na tentativa de refletir a inexatidão permitida na teoria, fixamos a precisão para otimalidade no método em θ_k^2 , onde $\theta_0 = 10^{-2}$ e $\theta_k = \max\{0, 9 \cdot \theta_{k-1}, 10^{-8}\}$, $k \geq 1$. Esta escolha é puramente empírica, não há relação aparente com a teoria; uma possível escolha com embasamento teórico pode ser investigada, especialmente tendo em vista o Lema 2.1.

No algoritmo, é requerido que $x_0 \in C$. Caso o ponto inicial x_0 fornecido não pertença a C , atualizamos $x_0 \leftarrow P_C(x_0)$.

5.2.1 Problema 1

Inspirados em [3], consideramos $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ e C dados por

$$\begin{aligned} f(x_1, x_2) &= (3x_1^3 - 2x_1^2, 0), \\ F(x_1, x_2) &= \begin{cases} \{-x_1, x_1\} \times \{0\}, & x_1 \geq 0 \\ \emptyset \times \{0\}, & x_1 < 0 \end{cases}, \\ C &= \mathbb{R} \times [1/2, 1]. \end{aligned}$$

As soluções são da equação generalizada $f(x) + F(x) \ni 0$ restrita à $x \in C$ são $(0, x_2)$ e $(1, x_2)$, $x_2 \in [1/2, 1]$.

Existem diferentes maneiras de escrever $F(x)$ como em (5.7). Cada uma delas afeta a resolução do problema de forma distinta. Neste caso, fazemos $F(x) = \{z \in \mathbb{R}^2 \mid z_1^2 - x_1 = 0\}$.

A Tabela 1 mostra a execução tomando $x_0 = (0, 7; 0)$. Rodamos o algoritmo para diferentes pontos iniciais, todos convergem para a solução $(0; \max\{0, 5; y_0\})$.

iter	e_k	e_k/e_{k-1}
1	6.975990e-01	
2	3.075904e-01	4.409272e-01
3	4.694257e-02	1.526139e-01
4	1.808056e-04	3.851635e-03
5	2.367476e-10	1.309404e-06

Tabela 1 – Resolução do Problema 1

Embora a análise da tabela acima nos sugira, ou pelo menos nos instigue a pensar numa convergência superlinear, o teorema de convergência mostra que pelo menos convergência linear temos.

5.2.2 Problema 2

Modificamos o problema anterior impondo uma maior relação entre as variáveis x_1 e x_2 :

$$\begin{aligned} f(x_1, x_2) &= (3x_1^3 - 2x_1^2, -x_1x_2), \\ F(x_1, x_2) &= \begin{cases} \{-x_1, x_1\} \times \{1 - x_1\}, & x_1 \geq 0 \\ \emptyset \times \{0\}, & x_1 < 0 \end{cases}, \\ C &= \mathbb{R} \times [0, 2]. \end{aligned}$$

As soluções são $(0, x_2)$, $x_2 \in [0, 2]$, e $(1, 0)$. Aqui, $F(x) = \{z \in \mathbb{R}^2 \mid z_1^2 - x_1 = 0, z_2 = 1 - x_1\}$. Curiosamente, o algoritmo é atraído para a solução $(0, 1)$ para diferentes pontos iniciais, inclusive partindo da solução $(0, 0)$ (evidentemente, isso depende da implementação). A Tabela 2 ilustra a execução para $x_0 = (0, 9; 0, 5)$.

iter	e_k	e_k/e_{k-1}
1	5.785110e-01	
2	5.344404e-01	9.238207e-01
3	2.963137e-01	5.544372e-01
4	1.049330e-01	3.541282e-01
5	3.895669e-02	3.712529e-01
6	6.432412e-04	1.651170e-02
7	6.269612e-06	9.746906e-03
8	2.101206e-09	3.351413e-04

Tabela 2 – Resolução do Problema 2

5.2.3 Problema 3

Consideramos o problema de dois níveis

$$\begin{aligned} \min_{x,y} \phi(x, y) \\ (x, y) \in D \\ y \in \arg \min_y q(x, y) \\ Yy \leq c - Xx, \end{aligned}$$

onde $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ e $c \in \mathbb{R}^p$. Supomos $D \neq \emptyset$ convexo e fechado, e $q(x, \cdot)$ uma função convexa e de classe C^1 para todo x .

Computar um ponto viável deste problema pode ser difícil, pois, além de y ser minimizador do problema inferior, estas variáveis estão presentes nas restrições $(x, y) \in D$ do problema superior. Uma das formas de lidar com problemas deste tipo é reescrever o problema do nível inferior sob suas condições KKT:

$$\begin{aligned} \nabla_y q(x, y) + Y^t \lambda &= 0 \\ \lambda &\geq 0 \\ (Xx + Yy - c)^t \lambda &= 0 \\ Yy &\leq c - Xx. \end{aligned}$$

Note que estas condições são suficientes para otimalidade pois estamos supondo que o problema da minimização de $q(x, \cdot)$ é convexo. Definindo

$$\begin{aligned} F(x, y, \lambda) &= \{0\}^2 \times N_{\mathbb{R}}(\lambda_1) \times \cdots \times N_{\mathbb{R}}(\lambda_p) \\ f(x, y, \lambda) &= (\nabla_y q(x, y) - Y^t \lambda, Xx + Yy - c) \\ C &= (D \cap \{(x, y) \mid Xx + Yy \leq c\}) \times \mathbb{R}_+^p, \end{aligned}$$

um ponto (x, y) é viável para o problema de dois níveis se, e somente se, existir λ tal que

$$0 \in f(x, y, \lambda) + F(x, y, \lambda), \quad (x, y, \lambda) \in C.$$

Observamos que a equação generalizada acima não envolve inequações; elas ficam a cargo do passo de projeção.

Para ilustrarmos o funcionamento do algoritmo, consideramos o problema particular com

$$q(x, y) = \frac{1}{2}(\|x\|^2 + \|y\|^2)$$

$$X = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad c = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

$$D = \{(x, y) \in \mathbb{R}^4 \mid x_1 + x_2 + y_1 + y_2 \leq 1\}$$

Em geral, o algoritmo apresenta comportamento bem diverso quando variamos o ponto de partida, convergindo a pontos diferentes. Observamos que, de fato, o problema de dois níveis admite vários pontos viáveis: podemos pensar o minimizador do nível inferior como uma função $y(x)$, que depende de x ; assim, para cada x teremos um par viável $(x, y(x))$ a depender, é claro, da existência de um multiplicador de Lagrange $\lambda(x)$. Partindo do ponto $(x_0, y_0, \lambda_0) = (0, 5; 0; 0; 1; 0, 1; 0)$, o algoritmo obtém

$$x^* = (-0,4231 ; -0.2694), \quad y^* = (0,0578 ; 0,0579), \quad \lambda^* = (0,0561 ; 0,0018)$$

(números arredondados na 4ª casa decimal), cujo resíduo KKT do problema do nível inferior é aproximadamente $2,82 \cdot 10^{-7}$. A Tabela 3 traz a execução do algoritmo.

iter	e_k	e_k/e_{k-1}
1	5.285721e-01	
2	7.818909e-02	1.479251e-01
3	2.928575e-02	3.745503e-01
4	8.903280e-03	3.040141e-01
5	8.555339e-03	9.609199e-01
6	9.162587e-03	1.070979e+00
7	9.458863e-03	1.032335e+00
8	9.841317e-03	1.040433e+00
9	7.568209e-03	7.690240e-01
10	7.309003e-03	9.657506e-01
11	7.506857e-03	1.027070e+00
12	4.989917e-03	6.647145e-01
13	2.197578e-04	4.404038e-02
14	1.499476e-05	6.823312e-02
15	1.945861e-03	1.297694e+02
16	9.971884e-05	5.124665e-02
17	1.766916e-05	1.771898e-01
18	1.619592e-03	9.166212e+01
19	1.854440e-03	1.145004e+00
20	1.321907e-03	7.128334e-01
21	1.387525e-03	1.049639e+00
22	1.607281e-03	1.158380e+00
23	1.274481e-03	7.929424e-01
24	1.325508e-03	1.040037e+00
25	6.816113e-04	5.142265e-01
26	7.024004e-04	1.030500e+00
27	2.823618e-07	4.019955e-04

Tabela 3 – Resolução do Problema 3

6 Considerações Finais

Neste trabalho buscamos contribuir um pouco no que se trata de Equações Generalizadas com Restrições, aqui além de acompanhar todo o esforço já praticado a respeito, fizemos uma abordagem Quasi-Newton para o problema. Provamos alguns teoremas que garantem a convergência das sequências geradas pelos algoritmos propostos, sob certas condições de regularidade da aplicação ponto-conjunto, bem como alguma propriedade sobre a atualização da matriz de aproximação da derivada. Percebemos, que assim como no problema irrestrito, a atualização de Broyden era de boa utilização no nosso contexto.

Fizemos, também, uma explicação de problemas já existentes na literatura que tinham outra proposta de solução mas que podiam ser repaginados na forma do nosso problema em questão.

Por fim, realizamos alguns testes numéricos, em problemas de pequeno porte, afim de observar como se comportaria computacionalmente a teoria desenvolvida embora sejam problemas difíceis de abordar.

Referências

- 1 OLIVEIRA, F. R. de; FERREIRA, O. P.; SILVA, G. N. Newton's method with feasible inexact projections for solving constrained generalized equations. *Comput Optim Appl*, v. 72, p. 159–177, 2019.
- 2 JOSEPHY, N. H. Newton's method for generalized equations and the pies energy model. *University of Wisconsin-Madison.*, 1979.
- 3 ARTACHO, F. J. A.; DONTCHEV, A. L.; BELYAKOV, A.; LOPEZ, M. Local convergence of quasi-Newton methods under metric regularity. *Comput Optim Appl*, v. 58, p. 225–247, 2014.
- 4 DONTCHEV, A. L.; ROCKAFELLAR, R. T. *Implicit Functions and Solution Mappings*. New York: Springer Series in Operations Research and Financial Engineering., 2014.
- 5 ROBINSOM, S. M. Extension of Newton's method to nonlinear functions with values in a cone. *Numer. Math.*, v. 19, p. 341–347, 1972.
- 6 ROBINSOM, S. M. Strongly regular generalized equations. *Math. Oper. Res.*, v. 5, p. 43–62, 1980.
- 7 ROBINSOM, S. M. Generalized equations. *Mathematical programming: the state of the art*, p. 346–367, 1983.
- 8 ARTACHO, F. J. A.; DONTCHEV, A.; GAYDU, M.; GEOFFROY, M. H.; VELIOV, V. M. Metric regularity of Newton's iteration. *SIAM J. Control Optim.*, v. 49, p. 339–362, 2011.
- 9 BONNANS, J. F. Local analysis of Newton-type methods for variational inequalities and nonlinear programming. *Appl. Math. Optim.*, v. 29, p. 161–186, 1994.
- 10 DONTCHEV, A. L. Local analysis of a Newton-type method based on partial linearization. *The mathematics of numerical analysis (Park City, UT, 1995), Lectures in Appl. Math.*, v. 32, p. 295–306, 1996.
- 11 DONTCHEV, A. L.; ROCKAFELLAR, R. T. Convergence of inexact Newton methods for generalized equations. *Mathematical Programming*, v. 139, p. 115–137, 2013.
- 12 FERREIRA, O. P. A robust semi-local convergence analysis of Newton's method for cone inclusion problems in Banach spaces under affine invariant majorant condition. *Comput. Appl. Math.*, v. 279, p. 318–335, 2015.
- 13 FERREIRA, O. P.; SILVA, G. N. Kantorovich's theorem on Newton's method for solving strongly regular generalized equation. *SIAM J. Optim.*, v. 27, p. 910–926, 2017.
- 14 FERREIRA, O. P.; SILVA, G. N. Local convergence analysis of Newton's method for solving strongly regular generalized equations. *J. Math. Anal. Appl.*, v. 458, p. 481–496, 2018.

- 15 DEMBO, R. S.; EISENSTAT, S. C.; STEIHAUG, T. Inexact Newton methods. *SIAM J. Numer. Anal.*, v. 19, p. 400–408, 1982.
- 16 KELLEY, C. T. *Solving nonlinear equations with Newton's method, Fundamentals of Algorithms*. Philadelphia, PA: SIAM, 2003.
- 17 CIBULKA, R.; DONTCHEV, A. L.; KRUGER, Y. A. Strong metric subregularity of mappings in variational analysis and optimization. *J. Math. Anal. Appl.*, v. 457, p. 1247–1282, 2018.
- 18 IZMAILOV, A. F.; SOLODOV, M. V. Inexact Josephy-Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization. *Comput. Optim. Appl.*, v. 46, p. 347–368, 2010.
- 19 RIBEIRO, A. A.; KARAS, E. W. *Otimização Contínua: aspectos teóricos e computacionais*. São Paulo: Cengage Learning, 2013.
- 20 IZMAILOV, A. F.; SOLODOV, M. V. *Otimização, Volume 2: Métodos Computacionais*. Rio de Janeiro: IMPA, 2012.
- 21 MARTÍNEZ, J. M.; SANTOS, S. A. *Métodos Computacionais de Otimização*. Rio de Janeiro: Sociedade Brasileira de Matemática, 1995.
- 22 FRIEDLANDER, A. *Elementos de Programação Não Linear*. Campinas(SP): Editora da Unicamp, 1994.
- 23 BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. M. *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley, 1993.
- 24 DENNIS, J. E.; SCHNABEL, R. B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Philadelphia: Englewood Cliffs, N. J., 1996.
- 25 GRIVA, I.; NASH, S. G.; SOFER, A. *Linear and Nonlinear Optimization, 2 ed.* Philadelphia: SIAM, 2009.
- 26 LUENBERGER, D.; YE, Y. *Linear and Nonlinear Programming, 3 ed.* New York: Springer, 2008.
- 27 NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. New York: Springer, 1999.
- 28 IZMAILOV, A. F.; SOLODOV, M. V. *Otimização, Volume 1: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. Rio de Janeiro: IMPA, 2005.
- 29 CENSOR, Y.; GIBALI, A.; REICH, S. Algorithms for the split variational inequality problem. *Numer. Algorithms*, v. 59(2), p. 301–323, 2012.
- 30 HE, H.; LING, C.; XU, H. K. A relaxed projection method for split variational inequalities. *J. Optim. Theory Appl.*, v. 166, p. 213–233, 2015.
- 31 CENSOR, Y.; BORTFELD, T.; MARTIN, B.; TROFIMOV, A. A unified approach for inversion problems in intensity-modulated radiation therapy. *Physics in Medicine and Biology*, v. 51, p. 2353–2365, 2006.
- 32 DONTCHEV, A. L.; ROCKAFELLAR, R. T. Convergence of inexact Newton methods for generalized equations. *Math. Program.*, v. 139, p. 115–137, 2013.

- 33 DONTCHEV, A. L.; LEWIS, A. S.; ROCKAFELLAR, R. T. The radius of metric regularity. *Transactions of the American Mathematical Society*, v. 355(2), p. 493–517, 2003.
- 34 KHAN, A. A.; TAMMER, C.; ZALINESCU, C. *Set-valued Optimization, An Introduction with Applications*. Verlag, Berlin: Springer, 2015.
- 35 BORWEIN, J. M.; ZHU, Q. J. *Techniques of Variational Analysis, An Introduction*. New York, NY: Springer, 2005.
- 36 BERTSEKAS, D. *Nonlinear Programming*. Belmont MA: Athena Scientific Optimization and Computation Series., 1999.
- 37 GONCALVES, M. L. N.; MELO, J. G. A Newton conditional gradient method for constrained nonlinear systems. *J. Comput. Appl. Math.*, v. 311, p. 473–483, 2017.
- 38 PALTA, J. R.; MACKIE, T. R. *Intensity-Modulated Radiation Therapy: The State of The Art*. Madison WI: Medical Physics Publishing, 2003.
- 39 CENSOR, Y.; CHEN, W.; COMBETTES, P. L.; DAVIDI, R.; HERMAN, G. T. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints. *Computational Optimization and Applications*, v. 51, p. 1065–1088, 2012.
- 40 KORPELEVICH, G. M. The extragradient method for finding saddle points and other problems. *Ekonomika i Matematicheskie Metody*, v. 12, p. 747–756, 1976.
- 41 BERTSEKAS, D. P.; TSITSIKLIS, J. N. *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall International, 1989.
- 42 DANG, Y.; GAO, Y. The strong convergence of a km–cq-like algorithm for a split feasibility problems. *Inverse Problems*, v. 27, p. 015007, 2011.
- 43 FICHERA, G. Sul problema elastostatico di signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.*, v. 34, p. 138–142, 1963.
- 44 STAMPACCHIA, G. Formes bilinéaires coercitives sur les ensembles convexes. *CR Acad. Sci. Paris*, v. 258, p. 4413–4416, 1964.
- 45 BYRNE, C. L. Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Probl.*, v. 18, p. 441–453, 2002.
- 46 CENSOR, Y.; ELFVING, T. A multiprojection algorithm using Bregman projections in product space. *Numer. Algorithm*, v. 8, p. 221–239, 1994.
- 47 CENSOR, Y.; ELFVING, T.; KOPF, N.; BORTFELD, T. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Probl.*, v. 21, p. 2071–2084, 2005.
- 48 LÓPEZ, G.; MARTIN-MARQUEZ, V.; WANG, F. H.; XU, H. K. Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Probl.*, v. 28, p. 085004, 2012.
- 49 MASAD, E.; REICH, S. A note on the multiple-set split convex feasibility problem in Hilbert space. *Nonlinear Convex Anal.*, v. 8, p. 367, 2007.

- 50 XU, H. K. A variable Krasnosel'skii–mann algorithm and the multiple-set split feasibility problem. *Inverse Probl.*, v. 22, p. 2021–2034, 2006.
- 51 XU, H. K. Iterative methods for the split feasibility problem in infinite dimensional Hilbert spaces. *Inverse Probl.*, v. 26, p. 105018, 2010.
- 52 YANG, Q. The relaxed cq algorithm solving the split feasibility problem. *Inverse Probl.*, v. 20, p. 1261–1266, 2004.
- 53 ZHANG, W. X.; HAN, D. R.; LI, Z. B. A self-adaptive projection method for solving the multiple-sets split feasibility problem. *Inverse Probl.*, v. 25, p. 115001, 2009.
- 54 ZHANG, W. X.; HAN, D. R.; YUAN, X. M. An efficient simultaneous method for the constrained multiplesets split feasibility problem. *Comput. Optim. Appl.*, v. 52, p. 825–843, 2012.
- 55 BAUSCHKE, H. H.; BORWEIN, J. M. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, v. 38, p. 367–426, 1996.
- 56 BYRNE, C. L. Iterative projection onto convex sets using multiple Bregman distances. *Inverse Probl.*, v. 15, p. 1295–1313, 1999.
- 57 BYRNE, C. L. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.*, v. 20, p. 103–120, 2004.
- 58 CENSOR, Y.; CHEN, W.; COMBETTES, P.; DAVIDI, R.; HERMAN, G. T. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints. *Comput. Optim. Appl.*, v. 51, p. 1065–1088, 2012.
- 59 CENSOR, Y.; DAVIDI, R.; HERMAN, G. T. Perturbation resilience and superiorization of iterative algorithms. *Inverse Probl.*, v. 26, p. 065008, 2010.
- 60 BYRNE, C.; CENSOR, Y.; GIBALI, A.; REICH, S. The split common null point problem. *Nonlinear Convex Anal.*, v. 13, p. 759–775, 2012.
- 61 CENSOR, Y.; GIBALI, A.; REICH, S. A Von Neumann alternating method for finding common solutions to variational inequalities. *Nonlinear Anal.*, v. 75, p. 4596–4603, 2012.
- 62 CENSOR, Y.; GIBALI, A.; REICH, S.; SABACH, S. Common solutions to variational inequalities. *Set-Valued Var. Anal.*, v. 20, p. 229–247, 2012.
- 63 COLAO, V.; MARINO, G.; XU, H. K. An iterative method for finding common solutions of equilibrium and fixed point problems. *Math. Anal. Appl.*, v. 344, p. 340–352, 2008.
- 64 CHEN, G.; TEBOULLE, M. A proximal-based decomposition method for convex minimization problems. *Math. Program.*, v. 64, p. 81–101, 1994.
- 65 FUKUSHIMA, M. Application of the alternating direction method of multipliers to separable convex programming problems. *Comput. Optim. Appl.*, v. 2, p. 93–111, 1992.
- 66 KAZMI, K. R.; RIZVI, S. H. An iterative method for split variational inclusion problem and fixed point problem for a nonexpansive mapping. *Optim. Lett.*, v. 8, p. 1113–1124, 2014.

-
- 67 KASSAY, G.; REICH, S.; SABACH, S. Iterative methods for solving systems of variational inequalities in reflexive Banach spaces. *SIAM J. Optim.*, v. 21, p. 1319–1344, 2011.
- 68 MOUDAFI, A. Split monotone variational inclusion. *J. Optim. Theory Appl.*, v. 150, p. 275–283, 2011.
- 69 TSENG, P. Alternating projection-proximal methods for convex programming and variational inequalities. *SIAM J. Optim.*, v. 7, p. 951–965, 1997.
- 70 WU, Q.; MOHAN, R.; NIEMIERKO, A.; SCHMIDT-ULLRICH, R. Optimization of intensity-modulated radiotherapy plan based on the equivalent uniform dose. *Int. J. Radiat. Oncol. Biol. Phys.*, v. 52, p. 224–235, 2003.