

UNICAMP

UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

LIZET SANTA CRUZ CALDERÓN

**Análise de Complexidade para Otimização
Multiobjetivo**

Campinas

2020

Lizet Santa Cruz Calderón

Análise de Complexidade para Otimização Multiobjetivo

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 Doutora em Matemática Aplicada.

Orientador: José Mario Martínez Pérez

Coorientadora: Maria Aparecida Diniz Ehrhardt

Este exemplar corresponde à versão final da Tese defendida pela aluna Lizet Santa Cruz Calderón e orientada pelo Prof. Dr. José Mario Martínez Pérez.

Campinas

2020

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

Santa Cruz Calderón, Lizet, 1989-
Sa59a Análise de complexidade para otimização multiobjetivo / Lizet Santa Cruz Calderón. – Campinas, SP : [s.n.], 2020.

Orientador: José Mario Martínez Pérez.
Coorientador: Maria Aparecida Diniz Ehrhardt.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Otimização multiobjetivo. 2. Complexidade (Matemática). 3. Hölder continuidade. I. Martínez Pérez, José Mario, 1948-. II. Ehrhardt, Maria Aparecida Diniz, 1956-. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Complexity analysis for multiobjective optimization

Palavras-chave em inglês:

Multiobjective optimization

Complexity (Mathematics)

Hölder continuity

Área de concentração: Matemática Aplicada

Titulação: Doutora em Matemática Aplicada

Banca examinadora:

José Mario Martínez Pérez [Orientador]

Sandra Augusta Santos

Roberto Andreani

Luís Felipe Cesar da Rocha Bueno

Leandro da Fonseca Prudente

Data de defesa: 03-04-2020

Programa de Pós-Graduação: Matemática Aplicada

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

- ORCID do autor: <https://orcid.org/0000-0001-6285-7761>

- Currículo Lattes do autor: <http://lattes.cnpq.br/6672949462618145>

**Tese de Doutorado defendida em 03 de abril de 2020 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). JOSÉ MARIO MARTÍNEZ PÉREZ

Prof(a). Dr(a). SANDRA AUGUSTA SANTOS

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

Prof(a). Dr(a). LUÍS FELIPE CESAR DA ROCHA BUENO

Prof(a). Dr(a). LEANDRO DA FONSECA PRUDENTE

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.

Agradecimentos

Ao meu orientador Mario Martínez e a minha co-orientadora Maria Aparecida Diniz pela excelente orientação, dedicação e paciência, pela oportunidade de me orientar neste trabalho, e pelos conhecimentos compartilhados, são um exemplo que levarei para minha vida acadêmica.

Agradeço primeiramente a Deus por me dar força de vontade o suficiente para chegar até aqui. Ao meu pai, Apolinar, minhas irmãs, Tania e Susana, e meu sobrinho Gianpiero, por todo o suporte e carinho dedicados a mim, sem os quais eu não chegaria a lugar algum. Agradeço ao meu namorado John pelo apoio, pela ajuda e paciência em todo o percurso do doutorado, e porque todo dia é um exemplo para mim de dedicação e esforço.

Agradeço aos Professores Roberto, Sandra e Lúcio pelas sugestões no Exame de Qualificação.

Agradeço aos meus amigos Daniel, Heraclio, Jackeline, Jamielli, Marina, Rui e Vitaliano que fizeram parte do percurso pelas conversas, apoio e estudos em grupo. Agradeço Fernanda Bia Peteam, pela paciência e ajuda que sempre deu para mim com relação ao MatLab.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. E com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) tendo número de processo 140781/2018-8.

*“ Talvez não tenha conseguido fazer o melhor,
mas lutei para que o melhor fosse feito.
Não sou o que deveria ser,
mas graças a Deus, não sou o que era antes.
(Martin Luther King)*

Resumo

Nesta tese é introduzido um método de regularização de ordem p para encontrar pontos estacionários fracos de problemas de otimização multiobjetivo com restrições. Sob condições de Hölder nas derivadas das funções objetivo, são obtidos resultados de complexidade que generalizam propriedades recentemente provadas para otimização escalar.

Palavras-chave: otimização multiobjetivo. complexidade. Hölder continuidade.

Abstract

In this thesis, a regularization method of order p is introduced to find weak stationary points of constrained multiobjective optimization problems. Under Hölder conditions on the derivatives of the objective functions, complexity results are obtained that generalize recently proved properties for scalar optimization.

Keywords: multiobjective optimization. complexity. Hölder continuity.

Lista de tabelas

Tabela 1 – Resultados do Algoritmo 4 quando o modelo é o polinômio de Taylor de ordem 2	68
Tabela 2 – Resultados do Algoritmo 4 quando o modelo é o polinômio de Taylor de ordem 3	74
Tabela 3 – Resultados do Algoritmo 4 para polinômios de ordem 2 e de ordem 3 .	79

Lista de abreviaturas e siglas

KKT	Karush-Kuhn-Tucker
OM	Otimização Multiobjetivo
NP	Número de problema

Lista de símbolos

$\ \cdot\ $	Denota a norma euclidiana.
$\min\{v, w\}$	Se $v, w \in \mathbb{R}^n$, $\min\{v, w\}$ denota o vetor com componentes, $\min\{v_1, w_1\}, \dots, \min\{v_n, w_n\}$.
v_+	Se $v \in \mathbb{R}^n$, v_+ é o vetor com componentes $\max\{v_1, 0\}, \dots, \max\{v_n, 0\}$.
$v \leq w$	Se $v, w \in \mathbb{R}^n$, denotamos $v \leq w$ se $v_i \leq w_i$, para todo $i = 1, \dots, n$. Analogamente, $v < w$ significa que $v_i < w_i$ para todo $i = 1, \dots, n$.
$M = O(g(x))$	Existe $C > 0$ tal que $M \leq Cg(x)$.
$[x]$	Denota o maior número inteiro menor ou igual a x , isto é, $\max\{m \in \mathbb{Z} m \leq x\}$.

Sumário

	Introdução	13
1	ANÁLISE DE COMPLEXIDADE PARA PROBLEMAS DE MINIMIZAÇÃO ESCALAR	17
1.1	Análise de complexidade para o caso $p = 1$	18
1.2	Análise de complexidade para o caso $p = 2$	24
2	ANÁLISE DE COMPLEXIDADE PARA PROBLEMAS MULTIOBJETIVO IRRESTRITOS	31
3	ANÁLISE DE COMPLEXIDADE PARA PROBLEMAS MULTIOBJETIVO COM RESTRIÇÕES	39
4	RESULTADOS NUMÉRICOS	49
5	CONSIDERAÇÕES FINAIS	54
	REFERÊNCIAS	55
	APÊNDICE A – PROBLEMAS TESTADOS	58
	APÊNDICE B – TABELAS	67

Introdução

Problemas de Otimização Multiobjetivo (OM) frequentemente aparecem nas áreas de matemática aplicada e engenharia [1], [2], [3], [4]. Dado um conjunto de funções objetivo $\{f_i\}_{i \in I}$ e um conjunto viável Ω , uma solução do problema multiobjetivo é um ponto viável x^* tal que não existe $x \in \Omega$ satisfazendo $f_i(x) \leq f_i(x^*), \forall i \in I$. Essa propriedade define "pontos de Pareto". Definições mais fortes de otimalidade foram dadas em diversos trabalhos (ver [5], [4], [6]).

Na otimização multiobjetivo procuramos resolver problemas de minimização de duas ou mais funções objetivo, considerando que quase nunca existe uma solução que minimize todas estas funções ao mesmo tempo, ou seja, enquanto na otimização escalar o valor ótimo é único, não podemos dizer o mesmo para problemas multiobjetivo porque uma solução que minimiza uma das funções geralmente não minimiza alguma ou todas as funções objetivo restantes (ver [7], [8]). A principal diferença entre a otimização escalar e a otimização multiobjetivo é a maneira de apresentar as soluções do problema. Como os objetivos geralmente são conflitantes, ou seja, a melhora de um acarreta na deterioração de outro(s), a resposta de um problema multiobjetivo não corresponde a uma única solução ótima, mas sim a um grupo de soluções que caracteriza o comprometimento entre os diversos objetivos. Este conjunto de soluções é chamado de fronteira Pareto, isto é, o conjunto de valores ótimos que não são dominados por nenhum outro ponto.

Uma das principais estratégias para a resolução de problemas de otimização multiobjetivo é a técnica de escalarização, que consiste em resolver um ou mais problemas de minimização escalar, de modo que os ótimos obtidos sejam pontos de Pareto para o problema original. Uma natural vantagem desta técnica é o fato de que existem vários métodos eficientes para resolver problemas escalares. Entre as diversas técnicas de escalarização, destaca-se o assim chamado método dos pesos, em que se minimiza uma combinação linear (não-negativa) dos objetivos. No entanto, os "pesos" adequados nem sempre são conhecidos a priori, de forma que os problemas escalares gerados pelo método podem ser ilimitados e, portanto, sequer possuir ótimos (ver [9]).

Algoritmos para resolver problemas de OM não lineares são iterativos. Dado um iterando x^k , é calculado um ponto x^{k+1} tal que $f_i(x^{k+1})$ seja suficientemente menor que $f_i(x^k)$ para todo $i \in I$. Isso é feito definindo um modelo para cada $f_i(x) - f_i(x^k)$ em torno de x^k e minimizando, aproximadamente, o máximo destes modelos. Fliege e Svaiter [10] parecem ser os primeiros a usar essa abordagem para resolver problemas multiobjetivo. Se o passo é reduzido suficientemente, teoremas de convergência adequados podem ser provados.

Quando as restrições do problema de OM definem um politopo, ou mesmo um conjunto convexo, o processo de encontrar minimizadores aproximados do modelo no conjunto viável pode ser simples e aproximações sucessivas para a solução do problema original podem ser obtidas usando busca linear. Na presença de restrições não-convexas a busca linear não preserva a viabilidade e, portanto, deve ser substituída por abordagens de região de confiança ou regularização. Felizmente, resolver subproblemas dentro de certas restrições não-convexas é, muitas vezes, acessível. É o caso de restrições que definem esferas, interseções de esferas com outros conjuntos simples, matrizes de densidade em cálculos eletrônicos [11], [12] e outras [13]. Além disso, se avaliações das funções objetivo são muito caras em comparação com as avaliações de modelos, métodos iterativos são aceitáveis para resolver subproblemas, desde que tenhamos um adequado critério para definir soluções aproximadas.

Neste trabalho vamos abordar problemas de OM com restrições, onde a única condição para as restrições é que estas sejam suficientemente simples, tornando possível a obtenção da solução aproximada dos subproblemas, e gerando soluções viáveis com satisfação moderada de otimalidade. Dado um iterando x^k , o método proposto minimizará aproximadamente um modelo do problema original mais um termo de regularização sujeito a restrições. Provaremos que, se o termo de regularização é suficientemente grande, ocorrerá um decréscimo em cada função objetivo. Neste caso, o ponto encontrado no processo de minimização é aceito como um novo iterando.

Se o modelo, em torno de x^k , do termo f_i da função objetivo concorda com $f_i(x)$ com erro $O(\|x - x^k\|^{p+1})$, devemos empregar termos de regularização da forma $\sigma\|x - x^k\|^{p+1}$. Portanto, mesmo no caso em que o erro é $O(\|x - x^k\|^{p+\beta})$ com $\beta < 1$, desde que $p + \beta > 1$, o emprego de uma regularização de ordem $p + 1$ é admissível. Por exemplo, erros da forma $O(\|x - x^k\|^{p+\beta})$ ocorrem quando o modelo é a p -ésima aproximação de Taylor de $f_i(x)$ e as p -ésimas derivadas não são Lipschitz contínuas, mas são Hölder contínuas com parâmetro β . Será mostrado que o algoritmo introduzido neste trabalho gera soluções aproximadas com precisão ϵ depois de $O(\epsilon^{-\frac{p+\beta}{p+\beta-1}})$ iterações e avaliações funcionais.

Uma maneira de obter soluções aproximadas é modificar o problema de otimização adicionando ao objetivo um termo de regularização (ou função de regularização), cujas propriedades tendem a favorecer a seleção de vetores desconhecidos com a estrutura desejada (ver [14]). Assim, obtemos problemas de otimização regularizada com a seguinte forma:

$$\text{Minimizar } f(x) + \sigma z(x), \quad (1)$$

onde f é a função objetivo, z é o termo de regularização, e σ é o parâmetro positivo de regularização.

Resultados de complexidade para problemas de otimização escalar irrestrita e restrita usando regularização foram comprovados em muitos trabalhos (ver [15], [16], [17], [18], [19], [20], [21]). Algoritmos que utilizam subproblemas com regularização quadrática para problemas de minimização escalar sem restrições, com complexidade de ordem $O(\epsilon^{-3/2})$, são apresentados em [16], [19], [18], [17], [22], [23], [24], [25]. A regularização cúbica foi introduzida em [26] e, em [27] é aplicada na resolução de problemas de engenharia. Em todos estes trabalhos foi assumida a Lipschitz continuidade da segunda derivada. Em 2016, Grapiglia e Nesterov [28] consideraram problemas sem restrições, em que, em vez da Lipschitz continuidade da segunda derivada, foi assumida a Hölder continuidade com parâmetro $\beta \in (0, 1]$. Em 2017, uma generalização da regularização cúbica, a p -regularização em problemas irrestritos, foi abordada em [29]. Em trabalhos recentes, foram propostos métodos de regularização baseados em modelos polinomiais de Taylor para minimização escalar, que apenas dependem das condições de Hölder nas derivadas de ordem superior.

Em 2017, Martínez [30] apresenta modelos gerais (não necessariamente de Taylor) de ordem arbitrária e emprega um critério de descida, trabalhando com otimização escalar e restrições não lineares que podem ser não convexas. Resultados de complexidade são mostrados.

Este trabalho de Martínez foi a base para o desenvolvimento deste tese. O algoritmo aqui proposto para problemas multiobjetivo calcula, em cada iteração, um ponto KKT aproximado para o subproblema, sem exigir o cumprimento de alguma condição de qualificação. Em cada iteração se a norma do gradiente do Lagrangiano associado ao subproblema é maior do que ϵ , o decréscimo requerido de cada função f_i é $\frac{\alpha \epsilon^{\frac{p+1}{p}}}{(2p+4)^{\frac{p+1}{p}} \sigma^{\frac{1}{p}}}$, onde σ é o parâmetro de regularização e $\alpha \in (0, 1)$. A quantidade de iterações e avaliações da função objetivo e suas derivadas necessárias para alcançar condições aproximadas de KKT com precisão ϵ é $O(\epsilon^{\frac{-p+\beta}{p+\beta-1}})$. Para o caso escalar, assumindo que a p -ésima derivada da função é Hölder contínua, Martínez apresenta resultados de complexidade compatíveis com os resultados apresentados em [17]. Com base nessa teoria, surgiu o interesse de estudar a complexidade para problemas multiobjetivo, baseados também em modelos gerais. A teoria de minimização multiobjetivo é estudada por diversos autores ([31], [32], [33]) mas, a diferença desses trabalhos com relação ao nosso, é que os resultados de complexidade serão obtidos usando método de regularização de ordem superior, o qual generalizará assim os resultados apresentados em [30] para o caso escalar. Para problemas de OM irrestrita e OM com restrições convexas, resultados de complexidade para o caso $p = 1$ foram apresentados em [34], [35].

Faremos também uma análise do desempenho do nosso algoritmo aplicado a testes numéricos baseados nos problemas de Moré, Garbow e Hilstrom [36], com dois valores diferentes de p ($p = 1, p = 2$).

Em resumo, as principais contribuições desta tese são:

- Proposta de um método para resolver problemas de OM com restrições, usando o método de regularização de ordem p , com $p > 1$, supondo a Hölder continuidade da função objetivo.
- Implementação computacional do método proposto.
- Testes numéricos dos problemas de Moré, Garbow e Hilstrom, reescritos na forma multiobjetivo com caixas.

Vale destacar que, a partir do trabalho desenvolvido nesta tese, escrevemos um artigo, já publicado online [37].

Esta tese está organizada da seguinte maneira: no Capítulo 1, fazemos uma revisão das definições básicas e dos resultados obtidos em [30] adaptados problemas irrestritos. Abordamos particularmente os casos em que $\beta = 1$, com $p = 2$ e $p = 1$ (isto é, para os casos em que as derivadas da função são Lipschitz contínuas). No Capítulo 2, apresentamos um algoritmo para problemas multiobjetivo irrestritos, supondo que as primeiras derivadas são Lipschitz contínuas (isto é $\beta = 1$) e também estudamos o caso em que as primeiras derivadas são Hölder contínuas. No Capítulo 3, introduzimos um algoritmo para problemas multiobjetivo com restrições e mostramos alguns resultados de complexidade. No Capítulo 4, apresentamos alguns resultados numéricos do algoritmo proposto aplicado aos problemas de Moré, Garbow e Hilstrom. E, concluindo, colocamos nossas considerações finais.

1 Análise de Complexidade para Problemas de Minimização Escalar

Neste capítulo será feita uma revisão dos resultados apresentados por Martínez em [30], adaptados para problemas de otimização escalar irrestrita.

Consideremos o problema

$$\text{Minimizar } f(x), \quad (1.1)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$. A Hölder continuidade com expoente β , da p -ésima derivada da função objetivo, é assumida, onde a p -ésima derivada de f em x é definida como o tensor de ordem p , dada por

$$\nabla^p f(x) = \left[\frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}} \right]_{i_j \in \{1, \dots, n\}, j=1, \dots, p}.$$

Definição 1.1. Dizemos que a p -ésima derivada de f é Hölder contínua com expoente β , se $\forall x, \bar{x} \in \mathbb{R}^n$, tem-se

$$\|\nabla^p f(x) - \nabla^p f(\bar{x})\|_{[p]} \leq L \|x - \bar{x}\|^\beta, \quad (1.2)$$

$p \in \{1, 2, 3, \dots\}$, $L > 0$ e $\beta \in [0, 1]$, com $p + \beta > 1$.

Em (1.2), $\|\cdot\|_{[p]}$ é a norma tensorial induzida pela norma euclidiana $\|\cdot\|$ no espaço de tensores de ordem p , que é dada por

$$\|T\|_{[p]} := \max_{\|v_1\| = \dots = \|v_p\|} |T[v_1, \dots, v_p]|,$$

onde $T[v_1, \dots, v_p]$ representa o tensor de ordem p resultante da aplicação do tensor T de ordem p nos vetores v_1, \dots, v_p (ver [29]).

Note que para o caso $\beta = 1$, a Definição 1.1 se reduz à definição de Lipschitz continuidade da p -ésima derivada da função.

Definimos um modelo suave de $f(x)$ em torno de \bar{x} , $M_{\bar{x}} : \mathbb{R}^n \rightarrow \mathbb{R}$. Dizemos que x e \bar{x} satisfazem as hipóteses (1.3)–(1.5) quando

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L \|x - \bar{x}\|^{p+\beta-1}, \quad (1.3)$$

$$M_{\bar{x}}(\bar{x}) = f(\bar{x}), \quad (1.4)$$

$$f(x) \leq M_{\bar{x}}(x) + L \|x - \bar{x}\|^{p+\beta}. \quad (1.5)$$

A interpretação de (1.3) e (1.5) é que o gradiente de f no ponto x é aproximado pelo gradiente de $M_{\bar{x}}$ no ponto x com um erro proporcional à potência $p + \beta - 1$ da distância entre x e \bar{x} e, mais ainda, $M_{\bar{x}}(x) + L\|x - \bar{x}\|^{p+\beta}$ superestima $f(x)$. Quando $M_{\bar{x}}$ é o polinômio de Taylor de ordem p em torno de \bar{x} e a p -ésima derivada satisfaz a condição de Hölder com expoente β e constante L , as condições mencionadas acima são satisfeitas para $\bar{x}, x \in \mathbb{R}^n$.

A seguir vamos apresentar um algoritmo para a resolução do problema (1.1), considerando que o gradiente da função objetivo (convexa ou não convexa) é Lipschitz contínuo, obtendo que o número máximo de iterações e avaliações da função objetivo e suas derivadas para garantir que, em uma certa iteração k , $\|\nabla f(x^k)\| \leq \epsilon$, ou $f(x^k) \leq f_{\text{target}}$, é $O(\epsilon^{-2})$, onde f_{target} é considerado como um limitante inferior para a função objetivo no seu domínio.

Analogamente será provado para o caso em que a função é duas vezes diferenciável e cuja Hessiana satisfaz a condição de Lipschitz continuidade (isto é equivalente a dizer que a Hessiana é Hölder contínua, com $\beta = 1$). Será provado que o número máximo de iterações e avaliações da função objetivo e suas derivadas para garantir que $\|\nabla f(x^k)\| \leq \epsilon$, ou $f(x^k) \leq f_{\text{target}}$, é $O(\epsilon^{-\frac{3}{2}})$.

1.1 Análise de complexidade para o caso $p = 1$.

No problema (1.1), para o caso $p = 1$, em que exigimos somente que a primeira derivada de f seja Lipschitz contínua, um modelo linear com uma regularização quadrática, vai ser minimizado.

Dado o problema

$$\min \quad M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^2, \quad (1.6)$$

onde

$$M_{\bar{x}}(x) := f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}), \quad (1.7)$$

as seguintes condições são satisfeitas:

$$M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^2 \leq f(\bar{x}) \quad (1.8)$$

e

$$\|\nabla[M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^2]\| \leq \theta\|x - \bar{x}\|, \quad (1.9)$$

onde σ é o parâmetro de regularização, e θ é um parâmetro positivo ($\theta > 0$). Além disso, \bar{x} e x satisfazem

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L\|x - \bar{x}\|, \quad (1.10)$$

$$M_{\bar{x}}(\bar{x}) = f(\bar{x}), \quad (1.11)$$

$$f(x) \leq M_{\bar{x}}(x) + L\|x - \bar{x}\|^2. \quad (1.12)$$

Para obter (1.12), note que, de (1.7), temos

$$\nabla M_{\bar{x}}(x) = \nabla f(\bar{x}).$$

Consideramos a aproximação de Taylor de ordem 1 com resto de Lagrange da função f

$$f(x) = f(\bar{x}) + \nabla f(\bar{x} + \xi(x - \bar{x})), \quad \xi \in [0, 1].$$

Note que,

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + [\nabla f(\bar{x} + \xi(x - \bar{x})) - \nabla f(\bar{x})]^T(x - \bar{x}),$$

logo, pela desigualdade de Cauchy-Schwarz,

$$f(x) \leq M_{\bar{x}}(x) + \|\nabla f(\bar{x} + \xi(x - \bar{x})) - \nabla f(\bar{x})\| \|x - \bar{x}\|.$$

Sabendo que a derivada é Lipschitz contínua, temos

$$f(x) \leq M_{\bar{x}}(x) + L\|x - \bar{x}\|^2.$$

Além disso $M_{\bar{x}}(\bar{x}) = f(\bar{x})$. Para obter (1.10), vemos que,

$$\nabla M_{\bar{x}}(x) = \nabla f(\bar{x}),$$

então

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| = \|\nabla f(x) - \nabla f(\bar{x})\| \leq L\|x - \bar{x}\|,$$

portanto

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L\|x - \bar{x}\|.$$

A seguir, apresentamos o Algoritmo 1, baseado no Algoritmo 2.1 de [30], voltado para problemas irrestritos e com $p = 1$.

Algoritmo 1. Assuma que $x^0 \in \mathbb{R}^n$, $\alpha \in (0, 1)$, $\epsilon \in (0, 1)$, $f_{target} \in \mathbb{R}$, $\theta > 0$ e $\sigma_{min} > 0$.

Inicialize $k \leftarrow 0$ e $\sigma_0 = \sigma_{min}$.

Passo 1: Defina $\sigma \leftarrow \sigma_k$.

Passo 2: Defina $\bar{x} \leftarrow x^k$.

Passo 3: Construa o modelo $M_{\bar{x}}(x)$ a partir de (1.7).

Passo 3.1 Encontre $x \in \mathbb{R}^n$ tal que as condições (1.8)-(1.9) sejam satisfeitas.

Passo 4: Se $\|\nabla f(x)\| \leq \epsilon$ ou $f(x) \leq f_{target}$, pare.

Passo 5: Se a condição de decréscimo suficiente

$$f(x) \leq f(\bar{x}) - \frac{\alpha \epsilon^2}{36 \sigma} \quad (1.13)$$

não se cumpre, faça $\sigma \leftarrow 2\sigma$ e volte para o Passo 2. Caso contrário vá para o Passo 6.

Passo 6: Defina $x^{k+1} = x$, $k \leftarrow k + 1$, $\sigma_k = \sigma$, e vá para o Passo 1.

Os resultados apresentados a seguir, são adaptações dos resultados apresentados em [30], voltados para problemas irrestritos.

No Lema seguinte, provaremos que, no caso em que o ponto x não satisfaz o critério de estacionariedade, então deve ser esperado um decréscimo na função objetivo. Este decréscimo apoia os resultados de complexidade, que indicam o número de iterações e avaliações necessárias para atingir o critério de estacionariedade.

Lema 1.1. *Assuma que $\alpha \in (0, 1)$ e que (1.8), (1.9), (1.12), (1.10) são satisfeitas para \bar{x} e x . Então.*

$$\|\nabla f(x)\| \leq (\theta + 2\sigma)\|x - \bar{x}\| + L\|x - \bar{x}\|. \quad (1.14)$$

Além disso, se

$$\|\nabla f(x)\| \geq \epsilon \quad (1.15)$$

e

$$\sigma \geq \max \left\{ \theta, \max \left\{ \frac{L}{3}, \frac{L}{1 - \alpha} \right\} \right\}, \quad (1.16)$$

temos que

$$\sigma \|x - \bar{x}\| \geq \frac{\epsilon}{6}, \quad (1.17)$$

$$f(x) \leq f(\bar{x}) - \alpha\sigma \|x - \bar{x}\|^2 \quad (1.18)$$

e

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{36} \frac{\epsilon^2}{\sigma}. \quad (1.19)$$

Demonstração: Por (1.9) e (1.10) temos

$$\begin{aligned} \|\nabla f(x)\| &\leq \|\nabla f(x) - \nabla M_{\bar{x}}(x)\| + \|\nabla M_{\bar{x}}(x)\| \\ &\leq L\|x - \bar{x}\| + \|\nabla[M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^2]\| + \|\nabla[\sigma\|x - \bar{x}\|^2]\| \\ &\leq L\|x - \bar{x}\| + \theta\|x - \bar{x}\| + 2\sigma\|x - \bar{x}\| \end{aligned}$$

Portanto

$$\|\nabla f(x)\| \leq (2\sigma + \theta)\|x - \bar{x}\| + L\|x - \bar{x}\|,$$

isto é (1.14) está provado. Por (1.15)-(1.16) temos

$$\begin{aligned} \epsilon &\leq \|\nabla f(x)\| \leq (2\sigma + \theta)\|x - \bar{x}\| + L\|x - \bar{x}\| \\ &\leq 3\sigma\|x - \bar{x}\| + L\|x - \bar{x}\| \end{aligned}$$

De (1.16) temos que $L \leq 3\sigma$, isto é, $6\sigma\|x - \bar{x}\| \geq \epsilon$, logo $\sigma\|x - \bar{x}\| \geq \frac{\epsilon}{6}$ e (1.13) está provado.

De (1.8) e (1.12) segue que

$$\begin{aligned} f(x) &\leq M_{\bar{x}}(x) + L\|x - \bar{x}\|^2 \\ &= M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^2 - \sigma\|x - \bar{x}\|^2 + L\|x - \bar{x}\|^2 \\ &\leq f(\bar{x}) + L\|x - \bar{x}\|^2 - \sigma\|x - \bar{x}\|^2. \end{aligned}$$

Temos que provar que

$$f(x) \leq f(\bar{x}) - \alpha\sigma\|x - \bar{x}\|^2.$$

De (1.16) temos

$$\sigma \geq \frac{L}{1 - \alpha},$$

logo

$$(1 - \alpha)\sigma \geq L$$

e

$$(1 - \alpha)\sigma\|x - \bar{x}\|^2 \geq L\|x - \bar{x}\|^2 \geq f(x) - f(\bar{x}) + \sigma\|x - \bar{x}\|^2,$$

o que satisfaz (1.18).

De (1.17) temos que

$$\sigma\|x - \bar{x}\| \geq \frac{\epsilon}{6}.$$

Logo

$$\begin{aligned} \sigma^2\|x - \bar{x}\|^2 &\geq \frac{\epsilon^2}{36}, \\ \|x - \bar{x}\|^2 &\geq \frac{\epsilon^2}{36\sigma^2}, \\ \alpha\sigma\|x - \bar{x}\|^2 &\geq \frac{\alpha\epsilon^2}{36\sigma}. \end{aligned}$$

Então

$$f(x) \leq f(\bar{x}) - \alpha\sigma\|x - \bar{x}\|^2 \leq f(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma}$$

Isto é,

$$f(x) \leq f(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma},$$

que é o que queríamos provar. ■

A partir do Lema 1.1, será visto no seguinte teorema que, se o Algoritmo 1 não parar no passo 4, será obtido um decréscimo funcional de $O(\epsilon^{-2})$.

Teorema 1.1. *Suponha que x^{k+1} é calculado pelo Algoritmo 1 e que as hipóteses (1.8)-(1.12) são satisfeitas para $\bar{x} = x^k$ em todos os pontos x obtidos no Passo 3 da iteração k . Definamos*

$$c_1 := \min \left\{ \frac{1}{72\theta}, \frac{1}{72 \max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\}} \right\}. \quad (1.20)$$

Então

$$f(x^{k+1}) \leq f(x^k) - \alpha c_1 \epsilon^2. \quad (1.21)$$

Demonstração: Seja

$$\sigma_{max} = 2 \max \left\{ \theta, \max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\} \right\}.$$

Pelo Lema 1.1 temos que $\sigma_k \leq \sigma_{max}$ e

$$f(x) \leq f(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma_k}.$$

Como $\frac{1}{\sigma_{max}} \leq \frac{1}{\sigma_k}$, então

$$\begin{aligned} f(x) &\leq f(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma_k} \\ &\leq f(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma_{max}}. \end{aligned}$$

Pela definição de σ_{max} temos que

$$f(x) \leq f(\bar{x}) - \frac{\alpha\epsilon^2}{72\theta}$$

e

$$f(x) \leq f(\bar{x}) - \frac{\alpha\epsilon^2}{72 \max\left\{\frac{L}{3}, \frac{L}{1-\alpha}\right\}}.$$

De (1.20) temos

$$c_1 \leq \frac{1}{72\theta}$$

e

$$c_1 \leq \frac{1}{72 \max\left\{\frac{L}{3}, \frac{L}{1-\alpha}\right\}}.$$

Logo

$$-\frac{\alpha\epsilon^2}{72\theta} \leq -\alpha c_1 \epsilon^2$$

ou

$$-\frac{\alpha\epsilon^2}{72 \max\left\{\frac{L}{3}, \frac{L}{1-\alpha}\right\}} \leq -\alpha c_1 \epsilon^2.$$

Portanto $f(x) \leq f(\bar{x}) - \alpha c_1 \epsilon^2$. ■

No teorema a seguir, será provado que o número de iterações realizadas pelo Algoritmo 1, é no máximo um múltiplo de $\epsilon^{-\frac{p+\beta}{p+\beta-1}}$.

Teorema 1.2. *Assuma que (1.8)-(1.10) são satisfeitas para $\bar{x} = x^k$ e em todos os pontos x calculados a cada iteração k dado pelo Algoritmo 1. Seja c_1 definido por (1.20). Então, depois de, no máximo,*

$$\left\lceil \frac{(f(x^0) - f_{target}) \epsilon^{-2}}{\alpha c_1} \right\rceil \tag{1.22}$$

iteraões, o Algoritmo 1 calcula $x \in \mathbb{R}^n$, verificando $f(x) \leq f_{target}$ ou

$$\|\nabla f(x)\| \leq \epsilon. \tag{1.23}$$

Além disso, excluindo as avaliações do gradiente para testar o critério de parada no Passo 4, o mesmo limitante superior (1.22) é válido para o número de avaliações das derivadas da função objetivo.

Demonstração: Do Teorema 1.1, temos $f(x^{k+1}) \leq f(x^k) - \alpha c_1 \epsilon^2$, do critério de parada do Algoritmo 1, temos que o número máximo de iterações é dado por (1.22). Além disso o Algoritmo 1 exige avaliar as derivadas de f apenas nos iterandos x^k . Portanto o Teorema fica provado. ■

Teorema 1.3. *Suponha que as hipóteses do Teorema 1.2 são satisfeitas. Então o número de avaliações de f e suas derivadas calculadas pelo Algoritmo 1 é limitado superiormente por*

$$(f(x^0) - f_{target}) \frac{\epsilon^{-2}}{\alpha c_1} + \max\{\log_2(\theta), c_l\} - \log_2(\sigma_{min}) + 1, \quad (1.24)$$

onde

$$c_l := \log_2 \left(\max \left\{ \frac{1}{3}L, \frac{L}{1-\alpha} \right\} \right). \quad (1.25)$$

Demonstração: Sabemos, pelo Teorema 1.2, que o número de iterações de f não pode ultrapassar

$$\left\lfloor \frac{(f(x^0) - f_{target})\epsilon^{-2}}{\alpha c_1} \right\rfloor.$$

De acordo com o Passo 5 do Algoritmo 2, sabemos que, se (1.13) não ocorre, então o valor do parâmetro σ é duplicado. No entanto, se $\sigma \geq \sigma_{max}$, (1.13) necessariamente é satisfeito. Daí temos que o número máximo de vezes que o parâmetro σ é duplicado é $\log_2 \left(\frac{\sigma_{max}}{\sigma_{min}} \right)$.

Como

$$\sigma_{max} = 2 \max \left\{ \theta, \max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\} \right\},$$

temos

$$\begin{aligned} \log_2 \left(\frac{\sigma_{max}}{\sigma_{min}} \right) &= \log_2(\sigma_{max}) - \log_2(\sigma_{min}) \\ &= \log_2 \left(2 \max \left\{ \theta, \max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\} \right\} \right) - \log_2(\sigma_{min}) \\ &= \log_2 2 + \max \left\{ \log_2(\theta), \log_2 \left(\max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\} \right) \right\} - \log_2(\sigma_{min}) \\ &= 1 + \max \left\{ \log_2(\theta), \log_2 \left(\max \left\{ \frac{L}{3}, \frac{L}{1-\alpha} \right\} \right) \right\} - \log_2(\sigma_{min}) \\ &= 1 + \max\{\log_2(\theta), c_l\} - \log_2(\sigma_{min}). \end{aligned}$$

Assim concluímos que o número de avaliações de f e suas derivadas é limitado por

$$(f(x^0) - f_{target}) \frac{\epsilon^{-2}}{\alpha c_1} + \max\{\log_2(\theta), c_l\} - \log_2(\sigma_{min}) + 1.$$

com c_l definido em (1.25). ■

Concluímos que, para $p = 1$, o Algoritmo 1 vai encontrar um ponto estacionário aproximado em no máximo $O(\epsilon^{-2})$ iterações e avaliações de f , isto é, o Algoritmo 1 calcula

$x \in \mathbb{R}^n$ verificando $f(x) \leq f_{target}$ ou $\|\nabla f(x)\| \leq \epsilon$ e, além disso, o número de avaliações de f e suas derivadas calculadas pelo Algoritmo 1 é limitado superiormente por (1.24), levando à complexidade $O(\epsilon^{-2})$.

1.2 Análise de complexidade para o caso $p = 2$.

Consideremos o problema (1.1) onde f é tal que sua segunda derivada é Lipschitz contínua. Para todo $\bar{x} \in \mathbb{R}^n$ definimos $M_{\bar{x}}(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, como o polinômio de Taylor de ordem 2

$$M_{\bar{x}}(x) := f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}). \quad (1.26)$$

Neste caso, as desigualdades (1.3)–(1.5) passam a ser escritas como:

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L\|x - \bar{x}\|^2, \quad (1.27)$$

$$M_{\bar{x}}(\bar{x}) = f(\bar{x}) \quad (1.28)$$

e

$$f(x) \leq M_{\bar{x}}(x) + L\|x - \bar{x}\|^3. \quad (1.29)$$

Para obter a desigualdade (1.29) vemos que:

$$\nabla M_{\bar{x}}(x) = \nabla f(\bar{x}) + \nabla^2 f(\bar{x})(x - \bar{x}),$$

que é obtido da definição (1.26). Consideremos a aproximação de Taylor de ordem 2 com resto de Lagrange da função f

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x} + \xi(x - \bar{x}))(x - \bar{x}), \quad \xi \in [0, 1].$$

Note que

$$\begin{aligned} f(x) &= f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}) \\ &\quad + \frac{1}{2}(x - \bar{x})^T [\nabla^2 f(\bar{x} + \xi(x - \bar{x})) - \nabla^2 f(\bar{x})](x - \bar{x}). \end{aligned}$$

Logo, pela desigualdade de Cauchy-Schwarz,

$$f(x) \leq M_{\bar{x}}(x) + \|x - \bar{x}\| \cdot \|\nabla^2 f(\bar{x} + \xi(x - \bar{x})) - \nabla^2 f(\bar{x})\| \cdot \|x - \bar{x}\|$$

Sabendo que a segunda derivada é Lipschitz contínua temos

$$f(x) \leq M_{\bar{x}}(x) + L\|x - \bar{x}\|^3.$$

Além disso,

$$M_{\bar{x}}(\bar{x}) = f(\bar{x}).$$

E, para obter a desigualdade (1.27), vemos que:

$$\begin{aligned}\nabla M_{\bar{x}}(x) &= \nabla f(\bar{x}) + \nabla^2 f(\bar{x})(x - \bar{x}), \\ f(x) &= f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x} + \xi(x - \bar{x}))(x - \bar{x}),\end{aligned}$$

então

$$\begin{aligned}\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| &= \|\nabla f(x) - \nabla f(\bar{x}) - \nabla^2 f(\bar{x})(x - \bar{x})\| \\ &= \|\nabla^2 f(\bar{x} + \xi(x - \bar{x}))(x - \bar{x}) - \nabla^2 f(\bar{x})(x - \bar{x})\| \\ &= \|[\nabla^2 f(\bar{x} + \xi(x - \bar{x})) - \nabla^2 f(\bar{x})](x - \bar{x})\| \\ &\leq L\|x - \bar{x}\|\|x - \bar{x}\| \\ &= L\|x - \bar{x}\|^2,\end{aligned}$$

portanto

$$\|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L\|x - \bar{x}\|^2.$$

Se $M_{\bar{x}}$ é o polinômio de Taylor de ordem 2 em torno de \bar{x} , e a segunda derivada satisfaz a condição de Lipschitz com constante L , as condições (1.29) e (1.27) são satisfeitas. Além disso, para todo \bar{x} , $x \in \mathbb{R}^n$, $M_{\bar{x}}(x)$ satisfaz as seguintes condições

$$M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^3 \leq f(\bar{x}) \quad (1.30)$$

e

$$\|\nabla[M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^3]\| \leq \theta\|x - \bar{x}\|^2, \quad (1.31)$$

No caso de $\theta = 0$, temos que (1.31) são as condições de estacionariedade para o problema de minimização irrestrito

$$\min \quad M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^3 \quad (1.32)$$

(ver [38]). Isto é, resolvendo o problema (1.32), podemos obter as desigualdades (1.30)-(1.31).

A partir das definições acima, apresentamos o Algoritmo 2 a seguir, voltado para a resolução do problema (1.1). Este algoritmo é uma adaptação, para problemas irrestritos, do Algoritmo 1.2 de [30].

Algoritmo 2. Assuma que $x^0 \in \mathbb{R}^n$, $\alpha \in (0, 1)$, $\epsilon \in (0, 1)$, $f_{target} \in \mathbb{R}$, $\theta > 0$, e $\sigma_{min} > 0$. Inicialize $k \leftarrow 0$ e $\sigma_0 = \sigma_{min}$.

Passo 1: Defina $\sigma \leftarrow \sigma_k$.

Passo 2: Defina $\bar{x} \leftarrow x^k$.

Passo 3: Construa o modelo $M_{\bar{x}}(x)$ a partir de (1.26).

Passo 3.1: Encontre $x \in \mathbb{R}^n$ tal que as condições (1.30)-(1.31) se cumpram.

Passo 4: Se $\|\nabla f(x)\| \leq \epsilon$ ou $f(x) \leq f_{target}$, pare.

Passo 5: Se a condição de decréscimo suficiente

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\sigma^{\frac{1}{2}}}, \quad (1.33)$$

não se cumpre, faça $\sigma \leftarrow 2\sigma$ e volte para o Passo 2. Caso contrário vá para o Passo 6.

Passo 6: Defina $x^{k+1} = x$, $k \leftarrow k + 1$, $\sigma_k = \sigma$ e vá para o Passo 1.

Analogamente ao caso $p = 1$, adaptamos os resultados apresentado em [30], voltados para problemas irrestritos, considerando $p = 2$. No seguinte lema asseguramos que, no caso em que o ponto \bar{x} não satisfaz o critério de estacionariedade, então deve ser esperado um decréscimo na função objetivo. Isto mostrará a boa definição do Algoritmo 2.

Lema 1.2. Assuma que $\alpha \in (0, 1)$ e que (1.29), (1.27), (1.30), (1.31) são satisfeitas para \bar{x} e x . Então

$$\|\nabla f(x)\| \leq (\theta + 3\sigma) \|x - \bar{x}\|^2 + L \|x - \bar{x}\|^2. \quad (1.34)$$

Além disso, se

$$\|\nabla f(x)\| \geq \epsilon \quad (1.35)$$

e

$$\sigma \geq \max \left\{ \theta, \max \left\{ \frac{L}{4}, \frac{L}{1 - \alpha} \right\} \right\}, \quad (1.36)$$

temos que

$$\sigma \|x - \bar{x}\|^2 \geq \frac{\epsilon}{8}, \quad (1.37)$$

$$f(x) \leq f(\bar{x}) - \alpha\sigma \|x - \bar{x}\|^3 \quad (1.38)$$

e

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\sigma^{\frac{1}{2}}}. \quad (1.39)$$

Demonstração: Por (1.27) temos

$$\|\nabla f(x)\| - \|\nabla M_{\bar{x}}(x)\| \leq \|\nabla f(x) - \nabla M_{\bar{x}}(x)\| \leq L \|x - \bar{x}\|^2,$$

$$\|\nabla f(x)\| \leq \|\nabla M_{\bar{x}}(x)\| + L \|x - \bar{x}\|^2.$$

Logo, de (1.31), temos

$$\begin{aligned} \|\nabla f(x)\| &\leq \|\nabla [M_{\bar{x}}(x) + \sigma \|x - \bar{x}\|^3] + \nabla [\sigma \|x - \bar{x}\|^3]\| + L \|x - \bar{x}\|^2 \\ &\leq \theta \|x - \bar{x}\|^2 + 3\sigma \|x - \bar{x}\|^2 + L \|x - \bar{x}\|^2. \end{aligned}$$

Portanto

$$\|\nabla f(x)\| \leq (\theta + 3\sigma) \|x - \bar{x}\|^2 + L \|x - \bar{x}\|^2,$$

isto é, (1.34) está provado.

Por (1.35), e como $\sigma \geq \theta$, então

$$\begin{aligned} \epsilon &\leq \|\nabla f(x)\| \leq (\theta + 3\sigma)\|x - \bar{x}\|^2 + L\|x - \bar{x}\|^2 \\ &\leq 4\sigma\|x - \bar{x}\|^2 + L\|x - \bar{x}\|^2. \end{aligned}$$

De (1.36) temos que $L \leq 4\sigma$, isto é, $8\sigma\|x - \bar{x}\|^2 \geq \epsilon$, logo $\sigma\|x - \bar{x}\|^2 \geq \frac{\epsilon}{8}$ e (1.37) ficou provado.

De (1.29) e (1.30), temos

$$f(x) \leq M_{\bar{x}}(x) + \sigma\|x - \bar{x}\|^3 + L\|x - \bar{x}\|^3 - \sigma\|x - \bar{x}\|^3 \leq f(\bar{x}) + L\|x - \bar{x}\|^3 - \sigma\|x - \bar{x}\|^3.$$

Temos que provar que

$$f(x) \leq f(\bar{x}) - \alpha\sigma\|x - \bar{x}\|^3.$$

De (1.36) temos que

$$\sigma \geq \frac{L}{1 - \alpha},$$

logo

$$(1 - \alpha)\sigma \geq L,$$

e

$$(1 - \alpha)\sigma\|x - \bar{x}\|^3 \geq L\|x - \bar{x}\|^3 \geq f(x) - f(\bar{x}) + \sigma\|x - \bar{x}\|^3,$$

o que satisfaz (1.38).

De (1.37) temos que

$$\|x - \bar{x}\| \geq \frac{\epsilon^{\frac{1}{2}}}{8^{\frac{1}{2}}\sigma^{\frac{1}{2}}}.$$

Logo

$$\begin{aligned} \|x - \bar{x}\|^3 &\geq \frac{\epsilon^{\frac{3}{2}}}{8^{\frac{3}{2}}\sigma^{\frac{3}{2}}}, \\ \sigma\|x - \bar{x}\|^3 &\geq \frac{\epsilon^{3/2}}{8^{\frac{3}{2}}\sigma^{\frac{1}{2}}}, \\ \alpha\sigma\|x - \bar{x}\|^3 &\geq \frac{\alpha\epsilon^{\frac{3}{2}}}{8^{\frac{3}{2}}\sigma^{\frac{1}{2}}}. \end{aligned}$$

Então

$$f(x) \leq f(\bar{x}) - \alpha\sigma\|x - \bar{x}\|^3 \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}}\frac{\epsilon^{\frac{3}{2}}}{\sigma^{\frac{1}{2}}},$$

isto é,

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}}\frac{\epsilon^{\frac{3}{2}}}{\sigma^{\frac{1}{2}}},$$

que é o que queríamos provar. ■

Teorema 1.4. *Suponha que x^{k+1} é calculado pelo Algoritmo 2 e que as hipóteses (1.29)-(1.31) são satisfeitas para $\bar{x} = x^k$ e em todos os pontos x calculados no Passo 3 da iteração k .*

Definamos

$$c_2 := \min \left\{ \frac{1}{8^{\frac{3}{2}}(2\theta)^{\frac{1}{2}}}, \frac{1}{8^{\frac{3}{2}} \left\{ 2 \max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right\}^{\frac{1}{2}}} \right\}. \quad (1.40)$$

Então

$$f(x^{k+1}) \leq f(x^k) - \alpha c_2 \epsilon^{\frac{3}{2}}. \quad (1.41)$$

Demonstração: Seja

$$\sigma_{max} := 2 \max \left\{ \theta, \max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right\}.$$

Pelo Lema 1.2 temos que $\sigma_k \leq \sigma_{max}$ para todo k . Por (1.39) temos

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\sigma_k^{\frac{1}{2}}}.$$

Como $\frac{1}{\sigma_{max}} \leq \frac{1}{\sigma_k}$, daí $-\frac{1}{\sigma_k} \leq -\frac{1}{\sigma_{max}}$, logo

$$\begin{aligned} f(x) &\leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\sigma_k^{\frac{1}{2}}} \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\sigma_{max}^{\frac{1}{2}}} \\ &\leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{(2\theta)^{\frac{1}{2}}}, \end{aligned}$$

e

$$f(x) \leq f(\bar{x}) - \frac{\alpha}{8^{\frac{3}{2}}} \frac{\epsilon^{\frac{3}{2}}}{\left(2 \left(\max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right) \right)^{\frac{1}{2}}}.$$

De (1.40) temos

$$\begin{aligned} c_2 &\leq \frac{1}{8^{\frac{3}{2}}} \frac{1}{(2\theta)^{\frac{1}{2}}}, \\ -\frac{\alpha \epsilon^{\frac{3}{2}}}{8^{\frac{3}{2}}(2\theta)^{\frac{1}{2}}} &\leq -\alpha c_2 \epsilon^{\frac{3}{2}} \end{aligned}$$

e

$$\begin{aligned} c_2 &\leq \frac{1}{8^{\frac{3}{2}} \left\{ 2 \max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right\}^{\frac{1}{2}}}, \\ -\frac{\alpha \epsilon^{\frac{3}{2}}}{8^{\frac{3}{2}} \left\{ 2 \max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right\}^{\frac{1}{2}}} &\leq -\alpha c_2 \epsilon^{\frac{3}{2}}. \end{aligned}$$

Logo

$$f(x) \leq f(\bar{x}) - \alpha c_2 \epsilon^{\frac{3}{2}}.$$

■

No próximo teorema será provado que o número máximo de iterações e avaliações da função objetivo e suas derivadas para garantir que, em um certo iterando k , $\|\nabla f(x)\| \leq \epsilon$ ou $f(x) < f_{\text{target}}$, é $O(\epsilon^{-\frac{3}{2}})$.

Teorema 1.5. *Assuma que (1.29)-(1.31) ocorrem para $\bar{x} = x^k$ e em todos os pontos x calculados em cada iteração k gerada pelo Algoritmo 2. Seja c_2 definido por (1.40). Então, depois de, no máximo,*

$$\lfloor (f(x^0) - f_{\text{target}}) \frac{\epsilon^{-\frac{3}{2}}}{\alpha c_2} \rfloor \quad (1.42)$$

iterações, o Algoritmo 2 calcula $x \in \mathbb{R}^n$, verificando $f(x) \leq f_{\text{target}}$ ou

$$\|\nabla f(x)\| \leq \epsilon. \quad (1.43)$$

Além disso, excluindo as avaliações do gradiente para testar o critério de parada no Passo 4 do Algoritmo 2, o mesmo limitante superior (1.42) é válido para o número de avaliações das derivadas da função objetivo.

Demonstração: Do Teorema 1.4 temos que $f(x^{k+1}) \leq f(x^k) - \alpha c_2 \epsilon^{\frac{3}{2}}$ e, do critério de parada do Passo 3 do Algoritmo 2, segue o resultado desejado no número de iterações. Além disso o Algoritmo 2 exige a avaliação das derivadas de f apenas nos iterandos x^k . Portanto o Teorema fica provado. ■

Teorema 1.6. *Suponha que as hipóteses do Teorema 1.5 são satisfeitas. Então o número de avaliações de f e suas derivadas calculadas pelo Algoritmo 2 é limitado superiormente por*

$$(f(x^0) - f_{\text{target}}) \frac{\epsilon^{-\frac{3}{2}}}{\alpha c_2} + \max\{\log_2(\theta), c_l\} - \log_2(\sigma_{\min}) + 1, \quad (1.44)$$

onde

$$c_l := \log_2 \left(\max \left\{ \frac{1}{4}L, \frac{L}{(1-\alpha)} \right\} \right). \quad (1.45)$$

Demonstração: Sabemos, pelo Teorema 1.5, que o número de avaliações de f não pode ultrapassar

$$\left\lfloor \frac{(f(x^0) - f_{\text{target}}) \epsilon^{-\frac{3}{2}}}{\alpha c_2} \right\rfloor.$$

De acordo com o Passo 5 do Algoritmo 2 sabemos que, se (1.33) não ocorre, então o valor do parâmetro σ é duplicado. No entanto, se $\sigma \geq \sigma_{\max}$, (1.33) necessariamente é satisfeito. Daí temos que o número máximo de vezes que o parâmetro σ é duplicado é $\log_2\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)$.

Como

$$\sigma_{\max} = 2 \max \left\{ \theta, \max \left\{ \frac{L}{4}, \frac{L}{1-\alpha} \right\} \right\},$$

temos

$$\begin{aligned} \log_2\left(\frac{\sigma_{max}}{\sigma_{min}}\right) &= \log_2(\sigma_{max}) - \log_2(\sigma_{min}) \\ &= \log_2\left(2 \max\left\{\theta, \max\left\{\frac{L}{4}, \frac{L}{1-\alpha}\right\}\right\}\right) - \log_2(\sigma_{min}) \\ &= 1 + \log_2\left(\max\left\{\theta, \max\left\{\frac{L}{4}, \frac{L}{1-\alpha}\right\}\right\}\right) - \log_2(\sigma_{min}). \end{aligned}$$

Como o logaritmo é uma função crescente temos

$$\log_2\left(\frac{\sigma_{max}}{\sigma_{min}}\right) = 1 + \max\left\{\log_2(\theta), \log_2\left(\max\left\{\frac{L}{4}, \frac{L}{1-\alpha}\right\}\right)\right\} - \log_2(\sigma_{min}).$$

Logo concluímos que o número de avaliações de f e suas derivadas é limitado por

$$(f(x^0) - f_{target}) \frac{\epsilon^{-\frac{3}{2}}}{\alpha c_2} + \max\{\log_2(\theta), c_l\} - \log_2(\sigma_{min}) + 1,$$

com c_l definido em (1.45). ■

Sob as hipóteses de diferenciabilidade assumidas nesta seção vemos que o Algoritmo 2 leva a um ponto estacionário aproximado de (1.1) em, no máximo, $O(\epsilon^{-\frac{3}{2}})$ iterações e avaliações da função objetivo e suas derivadas. Este é um caso particular do resultado obtido em [30], em que Martínez mostra que, para o caso de p arbitrário ($p \geq 1$) e $\beta \in (0, 1]$, o número máximo de iterações é igual a $O(\epsilon^{-\frac{p+\beta}{p+\beta-1}})$.

Com os resultados obtidos para os casos $p = 1$ e $p = 2$, conclui-se que em termos de complexidade, o melhor é trabalhar com $p = 2$, pois o número máximo de iterações necessárias para calcular o ponto estacionário aproximado é menor. Note que, embora com derivadas de ordem superior o número de iterações seja menor, as iterações se tornam mais caras computacionalmente, pois o fato de usar modelos de ordem superior, requer calcular derivadas de ordem maior.

Em [30] (Lema 2.1), o caso definido por $p = 1$ e $\beta \approx 0$ é o pior possível em termos de redução do valor da função objetivo em cada iteração, sendo a pior redução esperada $\alpha c_p \epsilon^{1+\frac{1}{\beta}}$ pois, para $\epsilon < 1$, a redução vai para zero quando $\beta \rightarrow 0$, implicando assim em um grande número de iterações para atingir o critério de parada.

2 Análise de Complexidade para Problemas Multiobjetivo Irrestritos

Em otimização multiobjetivo, dois ou mais objetivos devem ser minimizados simultaneamente, considerando que quase nunca existe um minimizador comum a todas as funções a serem minimizadas. Por este motivo, a noção clássica de otimalidade não é conveniente neste contexto. Uma das principais estratégias para a resolução de problemas de otimização multiobjetivo é a técnica de escalarização ([39], [40], [41], [42]), que consiste em resolver um ou mais problemas de minimização escalar, de modo que os ótimos obtidos sejam pontos de Pareto para o problema original. Resolver um problema de otimização multiobjetivo consiste em achar todos seus ótimos de Pareto, um conjunto cuja imagem é chamada fronteira Pareto. Achar a fronteira Pareto é um problema difícil. Geralmente conseguem-se aproximações dela.

Neste capítulo apresentamos um algoritmo para problemas multiobjetivo irrestritos, supondo que as primeiras derivadas de cada termo da função objetivo são Hölder contínuas, e analisaremos sua complexidade, adaptando os resultado de [30]. O conteúdo deste capítulo será uma preparação para o Capítulo 3 onde apresentaremos o problema multiobjetivo com restrições, o principal tema deste tese.

Vamos considerar o seguinte problema:

$$\text{Minimizar } F(x), \quad (2.1)$$

onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ pode ser escrita como $F(x) = [f_1(x), \dots, f_m(x)]$, com $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $\forall i = 1, \dots, m$.

Definição 2.1. (Ver [33]). Um ponto $\bar{x} \in \mathbb{R}^n$ é um ponto de Pareto do problema (2.1) se $\nexists x \in \mathbb{R}^n$, tal que

$$F(x) \leq F(\bar{x}) \text{ e } F(x) \neq F(\bar{x}),$$

onde o sinal de desigualdade \leq entre vetores deve ser entendido no sentido componente a componente.

Suponha, ao longo deste capítulo, que as primeiras derivadas de f_i , $\forall i = 1, \dots, m$, são Hölder contínuas. Seja $M_{\bar{x}}^i(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função definida como:

$$M_{\bar{x}}^i(x) := f_i(\bar{x}) + \nabla f_i(\bar{x})^T(x - \bar{x}), \forall i = 1, \dots, m, \quad (2.2)$$

e considere $\beta \in (0, 1]$. Então

$$f_i(x) \leq M_{\bar{x}}^i(x) + L\|x - \bar{x}\|^{\beta+1}, \quad (2.3)$$

$$M_{\bar{x}}^i(x) = f_i(\bar{x}), \quad (2.4)$$

e

$$\|\nabla f_i(x) - \nabla M_{\bar{x}}^i(x)\| \leq L\|x - \bar{x}\|^\beta. \quad (2.5)$$

De fato, a aproximação de Taylor de ordem 1, com resto de Lagrange, para f_i , com $i = 1, \dots, m$, em torno de \bar{x} , é dada por

$$f_i(x) = f_i(\bar{x}) + \nabla f_i(\bar{x} + \xi(x - \bar{x}))^T(x - \bar{x}),$$

Logo,

$$\begin{aligned} f_i(x) &= f_i(\bar{x}) + \nabla f_i(\bar{x})^T(x - \bar{x}) + [\nabla f_i(\bar{x} + \xi(x - \bar{x})) - \nabla f_i(\bar{x})]^T(x - \bar{x}) \\ &= M_{\bar{x}}^i(x) + [\nabla f_i(\bar{x} + \xi(x - \bar{x})) - \nabla f_i(\bar{x})]^T(x - \bar{x}), \forall i = 1, \dots, m. \end{aligned}$$

Pela desigualdade de Cauchy-Schwarz e, sabendo que a primeira derivada de cada f_i é Hölder contínua, temos

$$f_i(x) \leq M_{\bar{x}}^i(x) + L\|x - \bar{x}\|^{\beta+1}.$$

Para provar (2.5), lembramos que

$$\nabla M_{\bar{x}}^i(\bar{x}) = \nabla f_i(\bar{x}), \quad \forall i = 1, \dots, m,$$

logo

$$\begin{aligned} \|\nabla f_i(x) - \nabla M_{\bar{x}}^i(x)\| &= \|\nabla f_i(x) - \nabla f_i(\bar{x})\| \\ &\leq L\|x - \bar{x}\|^\beta, \forall i = 1, \dots, m. \end{aligned}$$

Portanto (2.5) é satisfeita.

A seguir, propomos o Algoritmo 3, voltado para o problema (2.1).

Algoritmo 3. Dados: $x^0 \in \mathbb{R}^n$ um ponto viável, $\alpha \in (0, 1)$, $\epsilon \in (0, 1)$, $\delta > 0$, $\theta > 0$ e $\sigma_{min} > 0$.

Inicialize $k \leftarrow 0$, $\sigma_0 = \sigma_{min}$.

Passo 1: Defina $\sigma \leftarrow \sigma_k$.

Passo 2: Defina $\bar{x} \leftarrow x^k$.

Passo 3: Construa o modelo $M_{\bar{x}}(x)$ a partir de (2.3)-(2.5).

Passo 3.1: Encontre $x \in \mathbb{R}^n$, $\gamma \in \mathbb{R}^m$, tais que, se $\bar{x} = x^k$

$$M_{\bar{x}}^i(x) + \sigma\|x - \bar{x}\|^2 \leq f_i(\bar{x}), \quad \text{para todo } i = 1, \dots, m \quad (2.6)$$

$$\gamma \geq 0, \quad \sum_{i=1}^m \gamma_i = 1 \quad (2.7)$$

e

$$\left\| \sum_{i=1}^m \gamma_i \nabla [M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^2] \right\| \leq \theta \|x - \bar{x}\|, \quad (2.8)$$

Passo 4: Se $\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| \leq \epsilon$, pare.

Passo 5: Se a condição de decréscimo suficiente

$$f_i(x) \leq f_i(x^k) - \frac{\alpha \epsilon^2}{36\sigma} \quad (2.9)$$

não é satisfeita para algum i , faça $\sigma \leftarrow 2\sigma$, e volte para o Passo 2. Caso contrário vá para o Passo 6.

Passo 6: Defina $x^{k+1} = x$, $k \leftarrow k + 1$, $\sigma_k = \sigma$, e vá ao Passo 1.

Note que o Passo 3.1 do Algoritmo 3 é equivalente a resolver o subproblema

$$\min_{i=1, \dots, m} \max \langle \nabla f_i(x^k), x - x^k \rangle + \sigma \|x - x^k\|^2, \quad (2.10)$$

o qual está baseado no modelo de Fliege e Svaiter [10].

Para o caso $m = 1$, a direção de máxima descida da função objetivo ocorre quando resolvemos o problema

$$\min \langle \nabla f_1(x^k), x - x^k \rangle + \frac{1}{2} \|x - x^k\|^2, \quad (2.11)$$

(ver [43]). Para otimização multiobjetivo, Fliege e Svaiter [10] propuseram tomar a direção de máxima descida como a solução do problema

$$\min_{i=1, \dots, m} \max \langle \nabla f_i(\bar{x}), x - \bar{x} \rangle + \sigma \|x - \bar{x}\|^2. \quad (2.12)$$

Note que (2.11) é um caso particular de (2.12) quando $m = 1$ e $\sigma = \frac{1}{2}$.

O seguinte teorema mostra que, se as funções f_i avaliadas no ponto x^k são uma combinação convexa então o ponto x^k é um ponto estacionário, candidato a ser um ponto Pareto.

Teorema 2.1. Se o problema

$$\min_{i=1, \dots, m} \max \langle \nabla f_i(x^k), x - x^k \rangle + \sigma \|x - x^k\|^2, \quad (2.13)$$

tem solução x^{trial} , então existe $\gamma \in \mathbb{R}_+^m$ tal que $x^{trial} - x^k = -\frac{1}{2\sigma} \sum_{i=1}^m \gamma_i \nabla f_i(x^k)$, $\sum_{i=1}^m \gamma_i = 1$

Demonstração: Resolver o problema (2.13) é equivalente a resolver

$$\begin{aligned} \min \quad & z + \sigma \|x - x^k\|^2 \\ \text{sujeita a} \quad & \langle \nabla f_i(x^k), x - x^k \rangle \leq z, \forall i = 1, \dots, m. \end{aligned} \quad (2.14)$$

Das condições de Karush, Kuhn e Tucker - KKT [44], tem-se que $\exists \mu \in \mathbb{R}_+^m$ tal que:

$$\begin{aligned} & \begin{pmatrix} 1 \\ 2\sigma(x^{trial} - x^k) \end{pmatrix} + \sum_{i=1}^m \mu_i \begin{pmatrix} -1 \\ \nabla f_i(x^k) \end{pmatrix} = 0, \\ & \mu_i (\langle \nabla f_i(x^k), x^{trial} - x^k \rangle - z) = 0, i = 1, \dots, m. \end{aligned}$$

Daí temos

$$x^{trial} - x^k = -\frac{1}{2\sigma} \sum_{i=1}^m \mu_i \nabla f_i(x^k), \quad \sum_{i=1}^m \mu_i = 1. \quad (2.15)$$

Considerando $\gamma = \mu$, a prova segue. ■

O seguinte lema afirma que, sempre que o ponto x não satisfaça o critério de parada do Algoritmo 3, deve ser esperado um decréscimo significativo no valor de cada função objetivo f_i , $\forall i = 1, \dots, m$. Provaremos que as equações (2.3)-(2.8), (2.17)-(2.18), implicam (2.16), (2.19)-(2.21).

Lema 2.1. *Suponha que $\alpha \in (0, 1)$ e (2.3), (2.5), (2.6), (2.7), (2.8) são satisfeitas em $\bar{x} = x^k, x$, então para todo $\gamma \in \mathbb{R}_+^m$, $\sum_{i=1}^m \gamma_i = 1$, tem-se*

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| \leq (\theta + 2\sigma) \|x - \bar{x}\| + L \|x - \bar{x}\|^\beta. \quad (2.16)$$

Se

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| \geq \epsilon, \text{ para todo } \gamma \in \mathbb{R}_+^m, \sum_{i=1}^m \gamma_i = 1 \quad (2.17)$$

e

$$\sigma \geq \max \left\{ \theta, \epsilon^{\frac{\beta-1}{\beta}} \max \left\{ \frac{2^{\frac{-\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}}} \cdot \frac{1}{6^{\frac{\beta-1}{\beta}}} \right\} \right\}, \quad (2.18)$$

então tem-se

$$\sigma \|x - \bar{x}\| \geq \frac{\epsilon}{6}, \quad (2.19)$$

$$f_i(x) \leq f_i(\bar{x}) - \alpha \sigma \|x - \bar{x}\|^2, \quad \forall i = 1, \dots, m, \quad (2.20)$$

e

$$f_i(x) \leq f_i(\bar{x}) - \frac{\alpha}{36} \frac{\epsilon^2}{\sigma}. \quad (2.21)$$

Demonstração: De (2.5), temos

$$\|\nabla f_i(x) - \nabla M_{\bar{x}}^i(x)\| \leq L \|x - \bar{x}\|^\beta, \quad \forall i = 1, \dots, m.$$

Para todo $\gamma \in \mathbb{R}_+^m$, $\sum_{i=1}^m \gamma_i = 1$, tem-se

$$\|\gamma_i \nabla f_i(x) - \gamma_i \nabla M_{\bar{x}}^i(x)\| \leq \gamma_i L \|x - \bar{x}\|^\beta, \quad \forall i = 1, \dots, m.$$

Do fato de que $\sum_{i=1}^m \gamma_i = 1$, e de (2.8) tem-se

$$\begin{aligned}
 \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) - \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) \right\| &\leq L \|x - \bar{x}\|^\beta \\
 \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| &\leq \left\| \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) \right\| + L \|x - \bar{x}\|^\beta \\
 &\leq \left\| \sum_{i=1}^m \gamma_i \nabla [M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^2] \right\| + \|\nabla[\sigma \|x - \bar{x}\|^2]\| \\
 &\quad + L \|x - \bar{x}\|^\beta \\
 &\leq \theta \|x - \bar{x}\| + 2\sigma \|x - \bar{x}\| + L \|x - \bar{x}\|^\beta \\
 &= (\theta + 2\sigma) \|x - \bar{x}\| + L \|x - \bar{x}\|^\beta.
 \end{aligned}$$

Portanto (2.16) é satisfeita.

De (2.16), (2.17), (2.18), como $\sigma \geq \theta$, temos

$$\begin{aligned}
 \epsilon &\leq (\sigma + 2\sigma) \|x - \bar{x}\| + L \|x - \bar{x}\|^\beta \\
 &\leq 3\sigma \|x - \bar{x}\| + L \|x - \bar{x}\|^\beta.
 \end{aligned}$$

Definamos $t = \sigma \|x - \bar{x}\|$, $c_1 = \frac{L}{3\sigma^\beta}$, então $t + c_1 t^\beta \geq \frac{\epsilon}{3}$.

Se $t \geq c_1 t^\beta$, e do fato de que $t > \frac{\epsilon}{3} - c_1 t^\beta$ tem-se

$$\begin{aligned}
 t &\geq \frac{\epsilon}{3} - t, \\
 2t &\geq \frac{\epsilon}{3}, \\
 t &\geq \frac{\epsilon}{6}, \\
 \sigma \|x - \bar{x}\| &\geq \frac{\epsilon}{6}.
 \end{aligned}$$

Se $c_1 t^\beta > t$, e do fato de que $c_1 t^\beta \geq \frac{\epsilon}{3} - t$ tem-se

$$\begin{aligned}
 c_1 t^\beta &\geq \frac{\epsilon}{3} - c_1 t^\beta, \\
 2c_1 t^\beta &\geq \frac{\epsilon}{3}, \\
 t &\geq \left(\frac{\epsilon}{6c_1} \right)^{\frac{1}{\beta}}.
 \end{aligned}$$

Provemos que $\frac{\epsilon}{6c_1} \geq \frac{\epsilon^\beta}{6^\beta}$. Isto é equivalente a provar que

$$\begin{aligned}
 \epsilon &\geq \frac{\epsilon^\beta 6L}{6^\beta 3\sigma^\beta} \\
 \sigma^\beta &\geq \frac{\epsilon^{\beta-1} 6^{1-\beta} L}{3} \\
 \sigma &\geq \frac{\epsilon^{\frac{\beta-1}{\beta}} 2^{\frac{1-\beta}{\beta}} L^{\frac{1}{\beta}}}{3}
 \end{aligned}$$

Portanto de (2.18), tem-se que $\frac{\epsilon}{6c_1} \geq \frac{\epsilon^\beta}{6^\beta}$, logo

$$\begin{aligned} t &\geq \left(\frac{\epsilon^\beta}{6^\beta}\right)^{\frac{1}{\beta}}, \\ t &\geq \frac{\epsilon}{6}. \end{aligned}$$

Assim (2.19) é satisfeita.

Por (2.3) e (2.6) tem-se:

$$\begin{aligned} f_i(x) &\leq M_{\bar{x}}^i(x) + L\|x - \bar{x}\|^{\beta+1} \\ &\leq M_{\bar{x}}^i(x) + \sigma\|x - \bar{x}\|^2 - \sigma\|x - \bar{x}\|^2 + L\|x - \bar{x}\|^{\beta+1} \\ &\leq f_i(\bar{x}) - \sigma\|x - \bar{x}\|^2 + L\|x - \bar{x}\|^{\beta+1}. \end{aligned}$$

Provemos que

$$\sigma\|x - \bar{x}\|^2 - L\|x - \bar{x}\|^{\beta+1} \geq \alpha\sigma\|x - \bar{x}\|^2,$$

o que é equivalente a provar que

$$\sigma(1 - \alpha) \geq L\|x - \bar{x}\|^{\beta-1}. \quad (2.22)$$

De (2.19) tem-se

$$\begin{aligned} \sigma\|x - \bar{x}\| &\geq \frac{\epsilon}{6}, \\ \|x - \bar{x}\|^{\beta-1} &\leq \frac{\epsilon^{\beta-1}}{(6\sigma)^{\beta-1}}. \end{aligned}$$

Para que (2.22) se cumpra é suficiente provar que

$$\sigma(1 - \alpha) \geq \frac{L\epsilon^{\beta-1}}{(6\sigma)^{\beta-1}},$$

isto é equivalente a provar que

$$\sigma \geq \frac{L^{\frac{1}{\beta}} \epsilon^{\frac{\beta-1}{\beta}}}{(1 - \alpha)^{\frac{1}{\beta}} 6^{\frac{\beta-1}{\beta}}}.$$

Logo, por (2.18), tem-se que (2.22) é satisfeita. Agora provemos (2.21). Por (2.19) temos

$$\begin{aligned} \|x - \bar{x}\| &\geq \frac{\epsilon}{6\sigma}, \\ \alpha\sigma\|x - \bar{x}\|^2 &\geq \frac{\alpha\epsilon^2}{36\sigma}, \end{aligned}$$

portanto $f_i(x) \leq f_i(\bar{x}) - \frac{\alpha\epsilon^2}{36\sigma}$. ■

No seguinte teorema, será provado que, se o Algoritmo 3 não parar no Passo 4, um decréscimo na função objetivo é obtido.

Teorema 2.2. Assuma que x^{k+1} é calculado pelo Algoritmo 3, e satisfaz (2.3), (2.5), (2.6), (2.7) (2.8) para x^k , e para todos os pontos x calculados no Passo 2 da iteração k . Definamos

$$c_1 := \min \left\{ \frac{1}{36} \frac{1}{2\theta}, \frac{1}{36 \left\{ 2 \max \left\{ \frac{2^{\frac{1-\beta}{\beta}} L^{\frac{1}{\beta}}}{3}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}} 6^{\frac{\beta-1}{\beta}}} \right\} \right\}} \right\}. \quad (2.23)$$

Então

$$f_i(x^{k+1}) \leq f_i(x^k) - \alpha c_1 \epsilon^{\frac{\beta+1}{\beta}}, \forall i = 1, \dots, m. \quad (2.24)$$

Demonstração: Considerando

$$\sigma_{\max} = 2 \max \left\{ \theta, \epsilon^{\frac{\beta-1}{\beta}} \max \left\{ \frac{2^{\frac{-\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}} \epsilon^{\frac{\beta-1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}} 6^{\frac{\beta-1}{\beta}}} \right\} \right\}, \quad (2.25)$$

pelo Lema 2.1 temos que $\sigma_k \leq \sigma_{\max}$ para todo k . Portanto, por (2.21), em cada iteração temos

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^2}{36 \sigma_k} \leq f_i(x^k) - \frac{\alpha \epsilon^2}{36 \sigma_{\max}}.$$

Da definição de σ_{\max} temos

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^2}{36(2\theta)} \quad (2.26)$$

e

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^2}{36 \left\{ 2 \epsilon^{\frac{\beta-1}{\beta}} \max \left\{ \frac{2^{\frac{-\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}} 6^{\frac{\beta-1}{\beta}}} \right\} \right\}}. \quad (2.27)$$

No caso de (2.26), (2.24) é satisfeito pela definição de c_1 e $\epsilon \leq 1$. Para o caso (2.27), temos que

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{\beta+1}{\beta}}}{36 \left\{ 2 \max \left\{ \frac{2^{\frac{1-\beta}{\beta}} L^{\frac{1}{\beta}}}{3}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}} 6^{\frac{\beta-1}{\beta}}} \right\} \right\}}.$$

Pela definição de c_1 , temos que (2.24) se cumpre. ■

No teorema a seguir, provaremos que o número de iterações realizadas pelo Algoritmo 3, é no máximo um múltiplo de $\epsilon^{-\frac{1+\beta}{\beta}}$

Teorema 2.3. Suponha que (2.3), (2.5), (2.6), (2.7), (2.8) são satisfeitas para x^k em todos os pontos x calculados no Passo 2 de cada iteração k executada pelo Algoritmo 3. Seja c_1 definido por (2.23). Então, dado $f_{\text{target}} \in \mathbb{R}$ arbitrário e $i \in \{1, \dots, m\}$, depois de, no máximo,

$$\lceil (f_i(x^0) - f_{\text{target}}) \frac{\epsilon^{-\frac{1+\beta}{\beta}}}{\alpha c_1} \rceil \quad (2.28)$$

iterações, o Algoritmo 3 calcula $x \in \mathbb{R}^n$, tal que existe $\gamma \in \mathbb{R}_+^m$, $\sum_{i=1}^m \gamma_i = 1$, verificando, $f_i(x) \leq f_{target}$ ou

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| \leq \epsilon.$$

Demonstração: O resultado segue da definição de c_1 e do Teorema 2.2. ■

Para completar a análise de complexidade em termos de avaliações funcionais, devemos adicionar o número de vezes em que σ é aumentado no Passo 5. Pelo Lema 2.1, não é necessário aumentar o σ , se $\sigma \geq \sigma_{\max}$. Portanto, o número máximo de vezes que o σ será limitado por um múltiplo de $-\log_2(\epsilon)$.

Teorema 2.4. *Suponha que as hipóteses do Teorema 2.3 são satisfeitas. Então o número de avaliações de F realizadas pelo Algoritmo 3, antes de obter $f_i(x^k) \leq f_{target}$ ou $\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) \right\| \leq \epsilon$, é limitado superiormente por*

$$(f(x^0) - f_{target}) \frac{\epsilon^{-\frac{1+\beta}{\beta}}}{\alpha c_1} + 1 + \max \left\{ \log_2(\theta), \frac{\beta-1}{\beta} \log_2(\epsilon) + C_l \right\} - \log_2(\sigma_{\min}), \quad (2.29)$$

onde

$$C_l := \log_2 \left(\max \left\{ \frac{2^{-\frac{\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}}} \frac{\epsilon^{\frac{\beta-1}{\beta}}}{6^{\frac{\beta-1}{\beta}}} \right\} \right).$$

Demonstração: Pelo Teorema 2.3, a avaliação de f_i e de suas derivadas em cada iteração x^k é dada pelo primeiro termo de (2.29). Cada vez que a condição de descida do Algoritmo 3 falha, o parâmetro de regularização σ é duplicado mas, pelo Lema 2.1, se $\sigma \geq \sigma_{\max}$, a condição de descida necessariamente ocorre, isto é, o número máximo em que σ é duplicado é limitado superiormente por $\log_2\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)$. Portanto, por (2.25), temos que

$$\begin{aligned} \log_2 \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right) &= \log_2(\sigma_{\max}) - \log_2(\sigma_{\min}) \\ &= \log_2 \left(2 \max \left\{ \theta, \epsilon^{\frac{\beta-1}{\beta}} \max \left\{ \frac{2^{-\frac{\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}}} \frac{\epsilon^{\frac{\beta-1}{\beta}}}{6^{\frac{\beta-1}{\beta}}} \right\} \right\} \right) - \log_2(\sigma_{\min}) \\ &= 1 + \max \left\{ \log_2(\theta), \frac{\beta-1}{\beta} \log_2(\epsilon) + \right. \\ &\quad \left. \log_2 \left(\max \left\{ \frac{2^{-\frac{\beta+1}{\beta}}}{3} L^{\frac{1}{\beta}}, \frac{L^{\frac{1}{\beta}}}{(1-\alpha)^{\frac{1}{\beta}}} \frac{\epsilon^{\frac{\beta-1}{\beta}}}{6^{\frac{\beta-1}{\beta}}} \right\} \right) \right\} - \log_2(\sigma_{\min}). \end{aligned}$$

Isto completa a prova. ■

3 Análise de Complexidade para Problemas Multiobjetivo com Restrições

Neste capítulo introduziremos um algoritmo para problemas multiobjetivo com restrições. Será mostrado que o algoritmo gera, para o problema proposto, soluções aproximadas com precisão ϵ depois de realizar $O(\epsilon^{-\frac{p+\beta}{p+\beta-1}})$ iterações.

Vamos considerar o seguinte problema:

$$\begin{aligned} \text{minimizar} \quad & F(x) \\ \text{sujeita a} \quad & h_E(x) = 0, \\ & h_I(x) \leq 0, \end{aligned} \tag{3.1}$$

onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h_E : \mathbb{R}^n \rightarrow \mathbb{R}^q$, $h_I : \mathbb{R}^n \rightarrow \mathbb{R}^s$. Defina $F(x) = [f_1(x), \dots, f_m(x)]$, com $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ funções cujas p -ésimas derivadas são Hölder contínuas $\forall i = 1, \dots, m$. Assumiremos que $p \in \{1, 2, 3, \dots\}$, $L > 0$, $\delta > 0$ e $\beta \in (0, 1]$, com $p + \beta > 1$. Também vamos supor que h_E e h_I são funções diferenciáveis e que $h'_E(x)$ e $h'_I(x)$ são seus Jacobianos respetivamente. Seja Ω o conjunto viável do problema (3.1), definido por

$$\Omega = \{x \in \mathbb{R}^n : h_E(x) = 0, h_I(x) \leq 0\}.$$

Dizemos que o ponto \bar{x} é viável quando $\bar{x} \in \Omega$.

Para todo $\bar{x} \in \mathbb{R}^n$, $i \in \{1, \dots, m\}$, definamos $M_{\bar{x}}^i : \mathbb{R}^n \rightarrow \mathbb{R}$ (um "modelo" de $f_i(x)$ em torno de \bar{x}). Assuma que as derivadas até ordem p de $M_{\bar{x}}^i$ existem para todo $x \in \mathbb{R}^n$. Dizemos que \bar{x} e x satisfazem as condições (3.2) e (3.4) quando

$$\|\nabla f_i(x) - \nabla M_{\bar{x}}^i(x)\| \leq L\|x - \bar{x}\|^{p+\beta-1}, \tag{3.2}$$

$$M_{\bar{x}}^i(\bar{x}) = f_i(\bar{x}) \tag{3.3}$$

e

$$f_i(x) \leq M_{\bar{x}}^i(x) + L\|x - \bar{x}\|^{p+\beta}, \tag{3.4}$$

para todo $i = 1, \dots, m$.

Se $M_{\bar{x}}(x)$ é o polinômio de Taylor de ordem p em torno de \bar{x} e a p -ésima derivada satisfaz a condição de Hölder definida para β e L , as condições (3.2) e (3.4) são satisfeitas para todo $x, \bar{x} \in \mathbb{R}^n$.

A seguir propomos o Algoritmo 4 voltado para o problema (3.1), o qual implementa um método que minimizará aproximadamente um modelo do problema original

mais um termo de regularização sujeito a restrições. Provaremos que, se o termo de regularização é suficientemente grande, ocorrerá um decréscimo em cada termo da função objetivo. Neste caso, o correspondente ponto “ x ” é aceito como uma nova iteração.

Algoritmo 4. Dados: $x^0 \in \mathbb{R}^n$ um ponto viável, $\alpha \in (0, 1)$, $\epsilon \in (0, 1)$, $\delta > 0$, $\theta > 0$ e $\sigma_{min} > 0$.

Inicialize $k \leftarrow 0$, $\sigma_0 = \sigma_{min}$.

Passo 1: Defina $\sigma \leftarrow \sigma_k$.

Passo 2: Defina $\bar{x} \leftarrow x^k$.

Passo 3: Construa o modelo $M_{\bar{x}}(x)$ a partir de (3.2)-(3.4).

Passo 3.1: Encontre $x \in \mathbb{R}^n$, $\gamma \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^s$, $\mu \in \mathbb{R}_+^q$, tais que, se $\bar{x} = x^k$

$$M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^{p+1} \leq f_i(\bar{x}), \text{ para todo } i = 1, \dots, m \quad (3.5)$$

$$\gamma \geq 0, \sum_{i=1}^m \gamma_i = 1 \quad (3.6)$$

e

$$\left\| \sum_{i=1}^m \gamma_i \nabla [M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^{p+1}] + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \leq \theta \|x - \bar{x}\|^p, \quad (3.7)$$

onde

$$\| \min\{\mu, -h_I(x)\} \| \leq \delta, \quad (3.8)$$

$$\|h_E(x)\| \leq \delta \text{ e } \|h_I(x)_+\| \leq \delta. \quad (3.9)$$

Passo 4: Se $\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \leq \epsilon$, pare.

Passo 5: Se a condição de decréscimo suficiente

$$f_i(x) \leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{2^{\frac{p+1}{p}} (p+2)^{\frac{p+1}{p}} \sigma^{\frac{1}{p}}} \quad (3.10)$$

não é satisfeita para algum i , faça $\sigma \leftarrow 2\sigma$, e volte para o Passo 2. Caso contrário vá para o Passo 6.

Passo 6: Defina $x^{k+1} = x$, $k \leftarrow k + 1$, $\sigma_k = \sigma$, e vá ao Passo 1.

Observação 3.1. Estamos supondo que as restrições de (3.1) são suficientemente fáceis para que seja possível encontrarmos uma solução aproximada para os subproblemas.

Se \bar{x} não é solução de (3.1), existe um ponto $x \in \mathbb{R}^n$ viável, tal que, $f_i(x) < f_i(\bar{x})$ para todo $i = 1, \dots, m$. Como $M_{\bar{x}}^i(x)$ é um modelo para $f_i(x)$, é razoável pedir um x viável tal que as componentes de $M_{\bar{x}}(x)$ diminuam tanto quanto seja possível.

Isso leva a minimizar $\max\{M_{\bar{x}}^1(x) - M_{\bar{x}}^1(\bar{x}), \dots, M_{\bar{x}}^m(x) - M_{\bar{x}}^m(\bar{x})\}$ sujeito à viabilidade. Equivalentemente, queremos resolver, aproximadamente,

$$\min \quad z \quad (3.11)$$

$$s.a \quad M_{\bar{x}}^i(x) - M_{\bar{x}}^i(\bar{x}) \leq z, i = 1, \dots, m, \quad (3.12)$$

$$h_E(x) = 0, \quad (3.13)$$

$$h_I(x) \leq 0. \quad (3.14)$$

Se as restrições (3.12)-(3.14) satisfazem uma condição de qualificação, a solução de (3.11)-(3.14) satisfaz as condições de KKT (ver [44]). Mas KKT para (3.11)-(3.14) implicam as condições (3.6)-(3.9) com $\sigma = 0$, $\theta = 0$ e $\delta = 0$. Portanto (3.6)-(3.9) satisfazem as condições de KKT aproximadas (ver [45]) para o problema

$$\min \quad \max\{M_{\bar{x}}^1(x) - M_{\bar{x}}^1(\bar{x}), \dots, M_{\bar{x}}^m(x) - M_{\bar{x}}^m(\bar{x})\} + \sigma \|x - \bar{x}\|^{p+1} \quad (3.15)$$

$$s.a \quad h_E(x) = 0, \quad (3.16)$$

$$h_I(x) \leq 0. \quad (3.17)$$

Como $M_{\bar{x}}$ é um modelo para $F(x)$, pode-se esperar que, restringindo o passo por meio de regularização, a condição (3.5) também será satisfeita.

Teorema 3.1. *Se o problema (3.15)-(3.17) tem solução x^* , então existem $\gamma \in \mathbb{R}_+^m$, $\mu \in \mathbb{R}_+^s$, $\lambda \in \mathbb{R}^q$, $\sum_{i=1}^m \gamma_i = 1$, tais que $x^* - x^k = -\frac{1}{(p+1)\sigma} \left(\sum_{i=1}^m \gamma_i \nabla M_{x^k}^i(x^*) + h'_I(x^*)^T \mu + h'_E(x^*)^T \lambda \right)$.*

Demonstração: O problema (3.15)-(3.17) é equivalente ao problema

$$\min \quad z + \sigma \|x - x^k\|^{p+1} \quad (3.18)$$

$$s.a \quad M_{x^k}^i(x) - M_{x^k}^i(x^k) \leq z, \forall i = 1, \dots, m, \quad (3.19)$$

$$h_E(x) = 0, \quad (3.20)$$

$$h_I(x) \leq 0. \quad (3.21)$$

Das condições KKT temos

$$\begin{aligned} & \begin{pmatrix} 1 \\ (p+1)\sigma(x^* - x^k) \end{pmatrix} + \sum_{i=1}^m \gamma_i \begin{pmatrix} -1 \\ \nabla M_{x^k}^i(x^*) \end{pmatrix} + \begin{pmatrix} 0 \\ h'_I(x^*)^T \mu \end{pmatrix} + \begin{pmatrix} 0 \\ \nabla h'_E(x^*)^T \lambda \end{pmatrix} \\ & = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned}$$

$$h_E(x^*) = 0,$$

$$h_I(x^*) \leq 0,$$

$$\gamma_i (M_{x^k}^i(x^*) - M_{x^k}^i(x^k) - z) = 0, \forall i = 1, \dots, m,$$

$$\mu^T h_I(x^*) = 0,$$

$$\gamma \in \mathbb{R}^m, \mu \in \mathbb{R}_+^s.$$

Logo

$$(p+1)\sigma(x^* - x^k) + \sum_{i=1}^m \gamma_i \nabla M_{x^k}^i(x^*) + h'_I(x^*)^T \mu + h'_E(x^*)^T \lambda = 0, \quad (3.22)$$

$$\sum_{i=1}^m \gamma_i = 1. \quad (3.23)$$

Assim

$$x^* - x^k = -\frac{1}{(p+1)\sigma} \left(\sum_{i=1}^m \gamma_i \nabla M_{x^k}^i(x^*) + h'_I(x^*)^T \mu + h'_E(x^*)^T \lambda \right).$$

■

No teorema a seguir, provaremos que se x não satisfaz o critério de estacionariedade, então deve ser esperado um decréscimo em cada uma das funções objetivo. Isto mostrará a boa definição do algoritmo proposto neste capítulo.

Teorema 3.2. *Suponha que x^k é um iterando calculado pelo Algoritmo 4, $\alpha \in (0, 1)$ e (3.2)-(3.9) são satisfeitas para $\bar{x} = x^k$, x , γ , λ e μ . Então*

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \leq (\theta + (p+1)\sigma) \|x - \bar{x}\|^p + L \|x - \bar{x}\|^{p+\beta-1}. \quad (3.24)$$

Além disso, se

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \geq \epsilon \quad (3.25)$$

e

$$\sigma \geq \max \left\{ \theta, \epsilon^{\frac{\beta-1}{p+\beta-1}} \max \left\{ \frac{2^{\frac{-\beta+1}{p+\beta-1}} L^{\frac{p}{p+\beta-1}}}{p+2}, \frac{L^{\frac{p}{p+\beta-1}}}{(1-\alpha)^{\frac{p}{p+\beta-1}} (2p+4)^{\frac{\beta-1}{p+\beta-1}}} \right\} \right\}, \quad (3.26)$$

tem-se

$$\sigma \|x - \bar{x}\|^p \geq \frac{\epsilon}{2p+4}, \quad (3.27)$$

$$f_i(x) \leq f_i(\bar{x}) - \alpha \sigma \|x - \bar{x}\|^{p+1}, \forall i = 1, \dots, m, \quad (3.28)$$

$$f_i(x) \leq f_i(\bar{x}) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{(2p+4)^{\frac{p+1}{p}} \sigma^{\frac{1}{p}}}. \quad (3.29)$$

Demonstração: Para todo $i = 1, \dots, m$, sabemos que

$$\|\nabla f_i(x) - M_{\bar{x}}^i(x)\| \leq L \|x - \bar{x}\|^{p+\beta-1}.$$

Assim

$$\|\gamma_i \nabla f_i(x) - \gamma_i \nabla M_{\bar{x}}^i(x)\| \leq \gamma_i L \|x - \bar{x}\|^{p+\beta-1},$$

logo, por (3.23),

$$\sum_{i=1}^m \|\gamma_i \nabla f_i(x) - \gamma_i \nabla M_{\bar{x}}^i(x)\| \leq L \|x - \bar{x}\|^{p+\beta-1},$$

então

$$\left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) - \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) \right\| \leq L \|x - \bar{x}\|^{p+\beta-1}.$$

Daí temos

$$\begin{aligned} \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| &= \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) - \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + \right. \\ &\quad \left. \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \\ &\leq \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) - \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) \right\| + \\ &\quad \left\| \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \\ &\leq \left\| \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \\ &\quad + L \|x - \bar{x}\|^{p+\beta-1}, \\ &\leq \left\| \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^{p+1} \right\| + h'_E(x)^T \lambda + \\ &\quad h'_I(x)^T \mu + \|\nabla[\sigma \|x - \bar{x}\|^{p+1}]\| + \\ &\quad L \|x - \bar{x}\|^{p+\beta-1}, \\ &\leq \left\| \sum_{i=1}^m \gamma_i \nabla M_{\bar{x}}^i(x) + \sigma \|x - \bar{x}\|^{p+1} \right\| + \\ &\quad h'_E(x)^T \lambda + h'_I(x)^T \mu + (p+1)\sigma \|x - \bar{x}\|^p + \\ &\quad L \|x - \bar{x}\|^{p+\beta-1}, \\ &\leq (\theta \|x - \bar{x}\|^p + \sigma(p+1) \|x - \bar{x}\|^p + L) \|x - \bar{x}\|^{p+\beta-1}, \\ &\leq (\theta + \sigma(p+1)) \|x - \bar{x}\|^p + L \|x - \bar{x}\|^{p+\beta-1}, \end{aligned}$$

o que conclui a prova de (3.24).

De (3.24) e (3.25) temos que

$$\epsilon \leq (\theta + \sigma(p+1)) \|x - \bar{x}\|^p + L \|x - \bar{x}\|^{p+\beta-1}.$$

De (3.26) segue que

$$\begin{aligned} \epsilon &\leq (\sigma + (p+1)\sigma) \|x - \bar{x}\|^p + L \|x - \bar{x}\|^{p+\beta-1}, \\ \epsilon &\leq (p+2)\sigma \|x - \bar{x}\|^p + L \|x - \bar{x}\|^{p+\beta-1}. \end{aligned} \tag{3.30}$$

Denotemos $t = \sigma \|x - \bar{x}\|^p$, $v = \frac{p + \beta - 1}{p}$, e $c_1 = \frac{L}{(p + 2)\sigma^v}$. Isto é, a desigualdade (3.30) pode ser expressa como

$$\frac{\epsilon}{(p + 2)} \leq t + c_1 t^v.$$

Se $t \geq c_1 t^v$, então

$$\frac{\epsilon}{(p + 2)} \leq 2t,$$

logo

$$\frac{\epsilon}{2(p + 2)} \leq \sigma \|x - \bar{x}\|^p.$$

Se $c_1 t^v \geq t$, então

$$\frac{\epsilon}{p + 2} \leq 2c_1 t^v,$$

logo

$$t^v \geq \frac{\epsilon}{2c_1(p + 2)},$$

então

$$t \geq \left(\frac{\epsilon}{2c_1(p + 2)} \right)^{\frac{1}{v}}.$$

Provemos que

$$\frac{\epsilon}{2(p + 2)c_1} \geq \left[\frac{\epsilon}{2(p + 2)} \right]^v.$$

Equivalentemente, basta provar que

$$c_1 \leq \epsilon^{1-v} [2(p + 2)]^{v-1}.$$

Assim, precisamos apenas provar a condição

$$\frac{L}{(p + 2)\sigma^v} \leq \epsilon^{1-v} [2(p + 2)]^{v-1}.$$

Isolando σ na desigualdade acima, tem-se

$$\begin{aligned} \sigma^v &\geq \frac{L}{(p + 2)\epsilon^{1-v} [2(p + 2)]^{v-1}}, \\ \sigma &\geq \frac{L^{\frac{1}{v}}}{(p + 2)^{\frac{1}{v}} \epsilon^{\frac{1}{v}-1} 2^{1-\frac{1}{v}} (p + 2)^{1-\frac{1}{v}}}, \end{aligned}$$

ou seja,

$$\sigma \geq L^{\frac{1}{v}} (p + 2)^{-1} \epsilon^{1-\frac{1}{v}} 2^{\frac{1}{v}-1}.$$

Usando a definição de v , essa desigualdade é equivalente a

$$\sigma \geq L^{\frac{p}{p+\beta-1}} (p + 2)^{-1} \epsilon^{\frac{\beta-1}{p+\beta-1}} 2^{\frac{1-\beta}{p+\beta-1}}.$$

Por (3.26), a última desigualdade ocorre. Portanto (3.27) é satisfeita.

Agora provemos (3.28). Por (3.4) e (3.5), para todo $i = 1, \dots, m$,

$$\begin{aligned} f_i(x) &\leq M_{\bar{x}}^i(x) + L\|x - \bar{x}\|^{p+\beta} \\ &\leq M_{\bar{x}}^i(x) + \sigma\|x - \bar{x}\|^{p+1} - \sigma\|x - \bar{x}\|^{p+1} + L\|x - \bar{x}\|^{p+\beta} \\ &\leq f_i(\bar{x}) - \sigma\|x - \bar{x}\|^{p+1} + L\|x - \bar{x}\|^{p+\beta}. \end{aligned}$$

Assim, para mostrar (3.28), basta provar que

$$\sigma\|x - \bar{x}\|^{p+1} - L\|x - \bar{x}\|^{p+\beta} \geq \alpha\sigma\|x - \bar{x}\|^{p+1}. \quad (3.31)$$

Vemos que (3.31) é equivalente a

$$L\|x - \bar{x}\|^{\beta-1} \leq \sigma(1 - \alpha). \quad (3.32)$$

Mas, por (3.27), temos que

$$\sigma^{1/p}\|x - \bar{x}\| \geq \frac{\epsilon^{1/p}}{(2p+4)^{1/p}},$$

logo,

$$\|x - \bar{x}\|^{\beta-1} \leq \frac{\epsilon^{(\beta-1)/p}}{((2p+4)\sigma)^{(\beta-1)/p}}.$$

Assim, uma condição suficiente para o cumprimento de (3.32) é:

$$\sigma(1 - \alpha) \geq \frac{L\epsilon^{(\beta-1)/p}}{((2p+4)\sigma)^{(\beta-1)/p}},$$

que é equivalente a

$$\begin{aligned} \sigma \cdot \sigma^{\frac{\beta-1}{p}} &\geq \frac{L\epsilon^{\frac{\beta-1}{p}}}{(2p+4)^{\frac{\beta-1}{p}}(1-\alpha)} \\ \sigma^{\frac{p+\beta-1}{p}} &\geq \frac{L\epsilon^{\frac{\beta-1}{p}}}{(2p+4)^{\frac{\beta-1}{p}}(1-\alpha)} \\ \sigma &\geq \frac{L^{\frac{p}{p+\beta-1}}\epsilon^{\frac{\beta-1}{p+\beta-1}}}{(2p+4)^{\frac{\beta-1}{p+\beta-1}}(1-\alpha)^{\frac{p}{p+\beta-1}}}. \end{aligned}$$

Pela definição de σ em (3.26), a prova de (3.28) está completa. Agora provemos (3.29).

Por (3.27), temos que

$$\|x - \bar{x}\| \geq \frac{\epsilon^{1/p}}{(2p+4)^{1/p}\sigma^{1/p}}.$$

Portanto,

$$\alpha\sigma\|x - \bar{x}\|^{p+1} \geq \frac{\alpha}{(2p+4)^{\frac{p+1}{p}}}\frac{\epsilon^{(p+1)/p}}{\sigma^{1/p}}.$$

Assim, temos que

$$f_i(x) \leq f_i(\bar{x}) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{2^{\frac{p+1}{p}} (p+2)^{\frac{p+1}{p}} \sigma^{\frac{1}{p}}}.$$

■

No teorema seguinte, provaremos que, se o Algoritmo 4 não parar no Passo 4, é esperado um decréscimo em cada função f_i .

Teorema 3.3. *Suponha que as hipóteses do Teorema 3.2 são satisfeitas. Então o iterando x^{k+1} está bem definido e, tomando*

$$c_p = \min \left\{ \frac{1}{[2(p+2)]^{\frac{p+1}{p}}} \frac{1}{(2\theta)^{\frac{1}{p}}}, \frac{1}{[2(p+2)]^{\frac{p+1}{p}} \left\{ 2 \max \left\{ \frac{2^{\frac{-\beta+1}{p+\beta-1}} L^{\frac{p}{p+\beta-1}}}{p+2}, \frac{L^{\frac{p}{p+\beta-1}}}{(1-\alpha)^{\frac{p}{p+\beta-1}} (2p+4)^{\frac{\beta-1}{p+\beta-1}}} \right\} \right\}^{\frac{1}{p}} \right\}, \quad (3.33)$$

temos que, para todo $i = 1, \dots, m$,

$$f_i(x^{k+1}) \leq f_i(x^k) - \alpha c_p \epsilon^{\frac{p+\beta}{p+\beta-1}}.$$

Demonstração: Seja

$$\sigma_{\max} = 2 \max \left\{ \theta, \epsilon^{\frac{\beta-1}{p+\beta-1}} \max \left\{ \frac{2^{\frac{-\beta+1}{p+\beta-1}} L^{\frac{p}{p+\beta-1}}}{p+2}, \frac{L^{\frac{p}{p+\beta-1}}}{(1-\alpha)^{\frac{p}{p+\beta-1}} [2(p+2)]^{\frac{\beta-1}{p+\beta-1}}} \right\} \right\},$$

Pelo Teorema 3.2 tem-se que $\sigma_k \leq \sigma_{\max}$ para todo k . Portanto

$$\begin{aligned} f_i(x^{k+1}) &\leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{2^{\frac{p+1}{p}} (p+2)^{\frac{p+1}{p}} \sigma_k^{\frac{1}{p}}} \\ &\leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{2^{\frac{p+1}{p}} (p+2)^{\frac{p+1}{p}} \sigma_{\max}^{\frac{1}{p}}}. \end{aligned}$$

Então

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{2^{\frac{p+1}{p}} (p+2)^{\frac{p+1}{p}} \sigma_{\max}^{\frac{1}{p}}}.$$

Pela definição de σ_{\max} tem-se

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{[2(p+2)]^{\frac{p+1}{p}} (2\theta)^{\frac{1}{p}}},$$

e

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{\alpha \epsilon^{\frac{p+1}{p}}}{[2(p+2)]^{\frac{p+1}{p}} 2 \epsilon^{\frac{\beta-1}{p+\beta-1}} \max \left\{ \frac{2^{\frac{\beta+1}{p+\beta-1}} L^{\frac{p}{p+\beta-1}}}{(p+2)}, \frac{L^{\frac{p}{p+\beta-1}}}{(1-\alpha)^{\frac{p}{p+\beta-1}} [2(p+2)]^{\frac{\beta-1}{p+\beta-1}}} \right\}}.$$

Pela definição de c_p , e com $\epsilon \leq 1$, o resultado segue. ■

Teorema 3.4. *Suponha que as hipóteses do Teorema 3.2 são satisfeitas para todo x^k calculado pelo Algoritmo 4. Seja c_p definida como no Teorema 3.3. Então, dado $f_{target} \in \mathbb{R}$ arbitrário e $i \in \{1, \dots, m\}$, depois de, no máximo*

$$\left[(f_i(x^0) - f_{target}) \frac{\epsilon^{-\frac{p+\beta}{p+\beta-1}}}{\alpha c_p} \right] \quad (3.34)$$

iterações, o Algoritmo 4 calcula $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}_+^m$, $\gamma \in \mathbb{R}_+^m$, $\sum_{i=1}^m \gamma_i = 1$, que verificam $f_i(x) \leq f_{target}$ ou

$$\begin{aligned} & \left\| \sum_{i=1}^m \gamma_i \nabla f_i(x) + h'_E(x)^T \lambda + h'_I(x)^T \mu \right\| \leq \epsilon, \\ & \|h_E(x)\| \leq \delta, \|h_I(x)_+\| \leq \delta \text{ e } \|\min\{\mu, -h_I(x)\}\| \leq \delta. \end{aligned}$$

Demonstração: O resultado desejado sobre o número de iterações segue de (3.8), (3.9), Teorema 3.3 e do critério de parada no Passo 4 do Algoritmo 4. Além disso, note que, excluindo as avaliações de gradiente para testar o critério de parada no Passo 4, o Algoritmo 4 exige a avaliação das derivadas de f_i apenas nos iterandos x^k (e não nos pontos x rejeitados). Portanto a tese está provada. ■

O Corolário a seguir afirma que, se pelo menos uma das funções f_i, \dots, f_m é limitada inferiormente, o Algoritmo 4 para após de $O(\epsilon^{-\frac{p+\beta}{p+\beta-1}})$ iterações e avaliações do gradiente.

Corolário 3.1. *Suponha que as hipóteses do Teorema 3.2 são satisfeitas para todo x^k calculado pelo Algoritmo 4. Se $i \in \{1, \dots, m\}$ e $f_{target} \in \mathbb{R}$ são tais que $f_i(x) > f_{target}$ para todo $x \in \mathbb{R}^n$ tal que $\|h_E(x)\| \leq \delta$ e $\|h_I(x)_+\| \leq \delta$, o Algoritmo 4 para realizando, no máximo, o número de iterações dado em (3.34).*

Demonstração: Os resultados seguem diretamente do Teorema 3.4. ■

O Teorema 3.4 e o Corolário 3.1 dizem respeito ao número de iterações realizadas pelo Algoritmo 4. O custo computacional associado a cada iteração envolve a avaliação dos gradientes em x^k (para testar o critério de parada no Passo 4), e das derivadas de ordem superior. Independentemente de tal esforço, a cada iteração precisamos avaliar f_1, \dots, f_m para diferentes pontos x . Isto significa que, além do trabalho relacionado a calcular o modelo $M_{x^k}^i(x)$ para cada i , devemos considerar as avaliações da função nos pontos rejeitados. Isso corresponde ao número de vezes em que o parâmetro de regularização σ é duplicado. Pelo Teorema 3.2, se $\sigma \geq \sigma_{max}$ não é necessário aumentar σ . Como consequência, um limitante superior para o número de avaliações funcionais são dados no seguinte teorema.

Teorema 3.5. *Suponha que as hipóteses do Teorema 3.2 são satisfeitas. Então o número de avaliações de F realizadas pelo Algoritmo 4 antes de obter $f_i(x^k) < f_{target}$ é limitado*

superiormente por

$$(f_i(x^0) - f_{target}) \frac{\epsilon^{-\frac{p+\beta}{p+\beta-1}}}{\alpha c_p} + \max \left\{ \log_2(\theta), \frac{1-\beta}{p+\beta-1} \log_2(\epsilon^{-1}) + c_l \right\} - \log_2(\sigma_{\min}) + 1,$$

onde

$$c_l = \log_2 \left(\max \left\{ \frac{2^{-\frac{\beta+1}{\beta}} L^{\frac{\beta}{p+\beta-1}}}{p+2}, \frac{L^{\frac{\beta}{p+\beta-1}}}{(1-\alpha)^{\frac{p}{p+\beta-1}} [2(p+2)]^{\frac{\beta-1}{p+\beta-1}}} \right\} \right).$$

Demonstração: Pelo Teorema 3.4 tem-se que o número de avaliações de f_i e suas derivadas em cada iteração x^k é dado por $(f_i(x^0) - f_{target}) \frac{\epsilon^{-\frac{p+\beta}{p+\beta-1}}}{\alpha c_p}$, para cada $i = 1, \dots, m$. Toda vez que a condição de descida do Algoritmo 4 falha para algum i , o parâmetro de regularização σ é duplicado. Se $\sigma \geq \sigma_{\max}$, a condição de descida necessariamente ocorre para todo $i = 1, \dots, m$, isto é, o número máximo em que σ é duplicado é limitado por $\log_2 \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)$. Portanto pela definição de σ_{\max} tem-se

$$\begin{aligned} \log_2 \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right) &= \log_2(\sigma_{\max}) - \log_2(\sigma_{\min}) \\ &= \left[\max \left\{ \log_2(\theta), \left(\frac{\beta-1}{p+\beta-1} \log_2(\epsilon) + c_l \right) \right\} \right] - \log_2(\sigma_{\min}) + 1. \end{aligned}$$

Isto completa a prova. ■

4 Resultados Numéricos

Com o objetivo de analisar o desempenho de nossa abordagem proposta com dois valores diferentes de p , ou seja, $p = 2$ e $p = 3$, realizamos a implementação do Algoritmo 4 no MatLab, versão R2019a. Para resolver, em cada iteração k , o subproblema (3.18)–(3.21), escolhemos a rotina *fmincon* do MatLab, com a opção que implementa um algoritmo de ponto interior, e com os parâmetros usuais da rotina. Com base nos problemas da coleção Moré, Garbow e Hillstrom [36], geramos nosso conjunto de testes em que os termos f_1, \dots, f_m da nossa função objetivo F são as funções que definem os problemas de quadrados mínimos não lineares descritos em [36]. Serão considerados os 35 problemas aí propostos e introduzimos limitantes para as variáveis. A formulação dos problemas testados será apresentada no Apêndice A. A numeração dos problemas é a mesma de [36].

Usamos $\alpha = 0.1$, $\epsilon = 1.e - 05$, $\delta = 0.1$, $\theta = 100$ e $\sigma_{min} = 1$. Foram escolhidas caixas arbitrariamente, nas quais a única condição sobre elas é conter a solução dos problemas dados para o caso de quadrados mínimos. Para facilitar a diagramação, as tabelas com os resultados obtidos encontram-se no Apêndice B. As Tabelas 1 e 2 mostram o desempenho do nosso algoritmo, para os 35 problemas testados em caixas, com relação ao valor da função objetivo F na solução obtida, e o número de iterações externas e internas (as realizadas por *fmincon*) usadas pelo método, para três diferentes pontos iniciais viáveis para cada problema. Somente omitimos os pontos iniciais e os correspondentes valores da função objetivo para o problema 19, pois este problema possui 11 variáveis e 65 termos em F . Para os 3 pontos iniciais escolhidos o algoritmo parou pelo critério de estacionariedade.

Observando o número de iterações externas e internas nessas tabelas, verificamos que a versão do Algoritmo 4 com $p = 3$ realiza, para cada ponto inicial escolhido, menos iterações externas e internas que a versão com $p = 2$, na maioria dos casos. No entanto, esta observação é feita sem levar em conta que, ainda que tenham sido usados os mesmos pontos iniciais para todos os testes com os modelos baseados nos polinômios de Taylor de ordem 2 e de ordem 3, as soluções obtidas são, em geral, diferentes. Assim, para uma análise um pouco mais precisa, resumimos, na Tabela 3, os casos em que ambas as versões (com $p = 2$ e $p = 3$) levam à mesma, ou aproximadamente à mesma, solução, a partir de um mesmo ponto inicial dado. As duas últimas colunas desta tabela exibem o número de iterações externas realizadas pelo algoritmo, com os diferentes polinômios de Taylor. Enquanto a versão com os polinômios de Taylor de ordem 2 realiza um total de 37 iterações externas nestes testes, a versão usando os polinômios de Taylor de ordem 3 performa um total de 22 iterações externas para os mesmos problemas, representando uma redução de aproximadamente 40% no que se refere às iterações externas, refletindo que, trabalhar com

modelos de maior ordem pode tornar o algoritmo mais eficiente em termos de número de iterações. Porém, devemos destacar que o uso de modelos de ordem $p = 3$ (ou de ordem superior) torna as iterações mais caras. Seria necessário o registro do tempo de execução gasto pelo algoritmo para uma análise de eficiência mais precisa.

Na Figura 1, exibimos os iterados calculados pelo Algoritmo 4 no Problema 2, iniciando de diferentes pontos iniciais. Os pontos circulados indicam os pontos que atenderam aos critérios de parada do Algoritmo 4, o qual os torna candidatos a pontos de Pareto. Mas note que, em três casos, as soluções obtidas apesar de satisfazerem o critério de parada, não são soluções multiobjetivo, pois são pontos dominados por outros. Esses pontos não fazem parte da Fronteira de Pareto do problema. Os asteriscos mostrados na Figura 1 representam os valores de (f_1, f_2) validados também nos pontos rejeitados, ou seja, as vezes que σ foi dobrado (Passo 5 do Algoritmo 4).

Finalmente, nas Figuras 2, 3 e 4, considerando F como a função de Rosenbrock (ver Apêndice A, Problema 1), exibimos os iterados calculados pelo Algoritmo 4 para três caixas escolhidas de diferentes maneiras. No primeiro caso, a caixa foi escolhida considerando a solução de quadrados mínimos como ponto exterior da caixa, no segundo caso a solução de quadrados mínimos foi considerada como ponto interior da caixa e no último caso, foi considerada na fronteira da caixa. Observamos que a aproximação da fronteira Pareto do problema varia de acordo com caixa escolhida. Em nossos testes, como mencionado anteriormente, optamos por definir as caixas considerando a solução de quadrados mínimos como ponto interior. Outra observação que deve ser feita é que, nas Figuras 2 e 4, há pontos que satisfizeram um critério de parada mas não são soluções multiobjetivos.

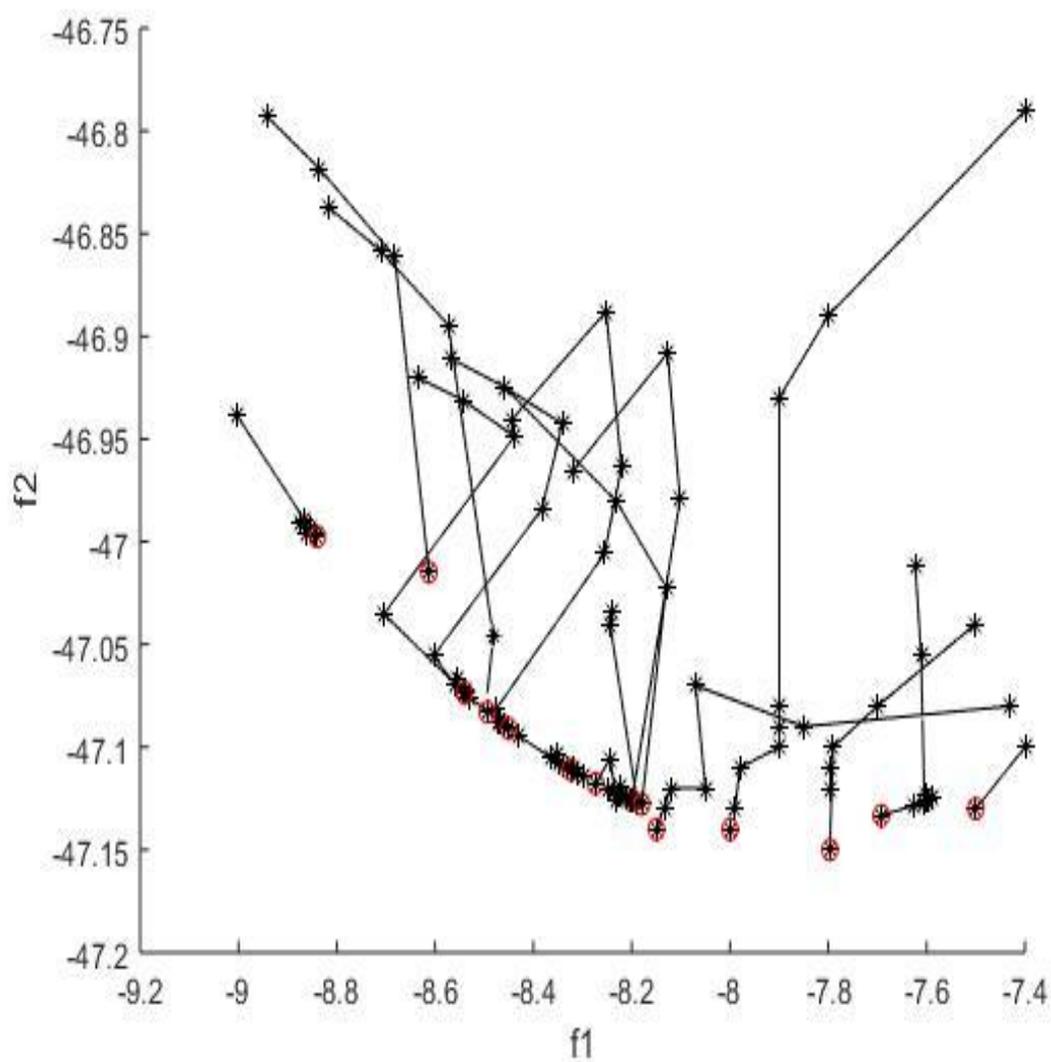


Figura 1 – Avaliações da função F nos iterados, ao resolver o Problema 2 a partir de 15 pontos iniciais arbitrários. (Note que em 3 casos, os pontos que satisfazem o critério de parada não são soluções multiobjetivos.)

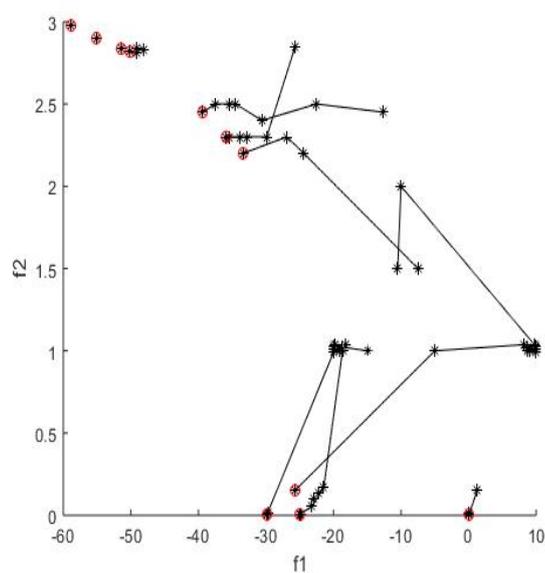


Figura 2 – Aproximação da Fronteira Pareto da função de Rosenbrock com $-2 \leq x_1 \leq 0$, $-2 \leq x_2 \leq 0$.

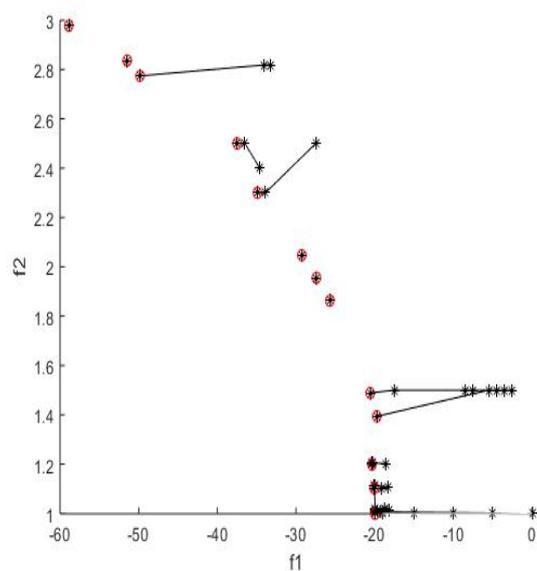


Figura 3 – Aproximação da Fronteira Pareto da função de Rosenbrock com $-2 \leq x_1 \leq 1.5$, $-2 \leq x_2 \leq 1.5$.

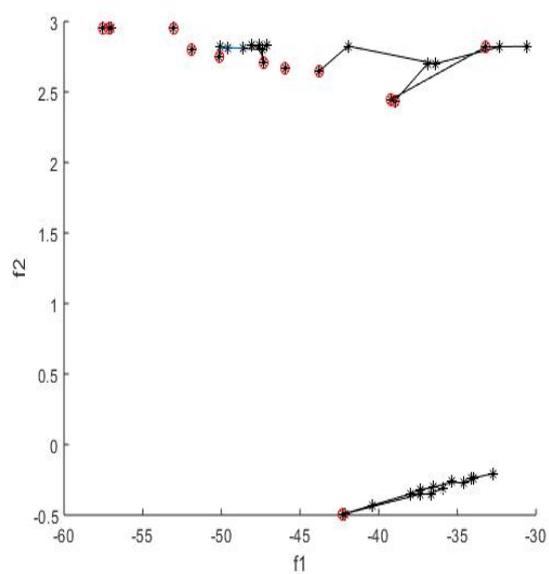


Figura 4 – Aproximação da Fronteira Pareto da função de Rosenbrock com $-2 \leq x_1 \leq 1$, $-2 \leq x_2 \leq 1$.

5 Considerações Finais

Nesta tese, estudamos a complexidade de algoritmos para problemas de otimização multiobjetivo. No entanto, iniciamos nosso trabalho com o estudo da complexidade para problemas de minimização escalar. Neste contexto, propusemos quatro algoritmos: os dois primeiros, para minimização escalar, são baseados em um caso particular do algoritmo proposto em [30], no qual assumimos a Lipchitz continuidade da primeira e segunda derivadas da função objetivo e obtivemos resultados de complexidade da ordem de $O(\epsilon^{-2})$ e $O(\epsilon^{-3/2})$ respectivamente.

Para o caso multiobjetivo, inicialmente irrestrito, apresentamos o Algoritmo 3, assumindo Hölder continuidade para a primeira derivada da função objetivo. Por fim, propusemos o Algoritmo 4 para o caso de problemas multiobjetivos com restrições, onde foi assumida a hipótese de Hölder continuidade da p -ésima derivada da função objetivo. Obtivemos resultados de complexidade para ambos os casos.

Apresentamos também a aplicabilidade do Algoritmo 4, realizando sua implementação computacional em MatLab. O objetivo principal desta implementação foi testar os modelos que usam derivadas terceiras da função objetivo, ou seja, o caso em que $p = 3$, e derivadas segundas, ou seja, o caso em que $p = 2$. Os resultados dos experimentos indicaram que a versão do Algoritmo 4 com $p = 3$ se mostra mais eficiente do que aquela com $p = 2$.

Concluindo, podemos destacar que, com esta tese, obtivemos contribuições relevantes para o estudo de complexidade de algoritmos para programação multiobjetivo.

Referências

- 1 HASLINGER, J.; MAKINEN, R. A. E. Introduction to shape optimization. *SIAM Publications, Philadelphia*, 2003.
- 2 BELTON, V.; STEWART, T. J. *Multiple Criteria Decision Analysis: An Integrated Approach*. Boston: Kluwer Academic Publishers, 2002.
- 3 LUC, D. T. *Notes in Economy and Mathematical*. Berlin: Springer, 1989.
- 4 MIETTINEN, K. M. Nonlinear multiobjective optimization. *Kluwer Academic Publishers, Boston*, 1999.
- 5 DRUMMOND, L. M. G.; IUSEM, A. N. A projected gradient method for vector optimization problems. *Computational Optimization and Applications*, n. 28, p. 5–29, 2004.
- 6 BRANKE, J.; DEB, K.; MIETTINEN, K.; SLOWINSKI, R. *Multiobjective Optimization*. Berlin: in Lecture Notes in Comput. Sci. 5252, Springer Verlag, 2008.
- 7 ESCHENAUER, H.; KOSKI, J.; OSYCZKA, A. *Multicriteria Design Optimization*. Berlin: Springer, 1990.
- 8 SAWARAGI, Y.; NAKAYAMA, H.; TANINO, T. *Theory of Multiobjective Optimization*. Orlando, San Diego, New York, London: of Mathematics in Science and Engineering, Academic Press, 1985. v. 176.
- 9 DRUMMOND, L. M. G.; SVAITER, B. F. *Métodos de Descida em Otimização Multiobjetivo*. IMPA: 30º Colóquio Brasileiro de Matemática, 2015.
- 10 FLIEGE, J.; SVAITER, B. F. Steepest descent methods for multicriteria optimization. *Math. Methods Oper*, n. 51, p. 479–494, 2000.
- 11 FRANCISCO, J. B.; MARTÍNEZ, J. M.; MARTÍNEZ, L. Globally convergent trust-region methods for self-consistent field electronic structure calculations. *Journal of Chemical Physics*, n. 121, p. 10863–10878, 2004.
- 12 FRANCISCO, J. B.; MARTÍNEZ, J. M.; MARTÍNEZ, L. Density-based globally convergent trust-region method for self-consistent field electronic structure calculations. *Journal of Mathematical Chemistry*, n. 40, p. 349–377, 2006.
- 13 FRANCISCO, J. B.; BAZÁN, F. S. V. Nonmonotone algorithm for minimization on closed sets with application to minimization on stiefel manifolds. *Journal of Computational and Applied Mathematics*, n. 236, p. 2717–2727, 2012.
- 14 SRA, S.; NOWOZIN, S.; WRIGHT, S. *Optimization for Machine Learning*. Cambridge, MA: MIT Press, 2011.
- 15 BIRGIN, E. G.; GARDENGHI, J. L.; MARTÍNEZ, J. M.; SANTOS, S. A.; TOINT, L. Evaluation complexity for nonlinear constrained optimization using unscaled kkt conditions and high-order models. *SIAM Journal on Optimization*, n. 26, p. 951–967, 2016.

- 16 BIRGIN, E. G.; MARTÍNEZ, J. M. The use of quadratic regularization with cubic descent for unconstrained optimization. *SIAM Journal on Optimization*, n. 27, p. 1049–1074, 2017.
- 17 CARTIS, C.; GOULD, N. I. M.; TOINT, L. On the complexity of steepest descent, Newton's and regularized Newton's methods for nonconvex unconstrained optimization. *SIAM Journal on Optimization*, n. 20, p. 2833–2852, 2010.
- 18 CARTIS, C.; GOULD, N. I. M.; TOINT, L. Adaptive cubic regularization methods for unconstrained optimization. PartII: worst-case function and derivate complexity. *Mathematical Programming*, n. 130, p. 295–319, 2011.
- 19 CARTIS, C.; GOULD, N. I. M.; TOINT, L. Adaptive cubic regularization methods for unconstrained optimization. PartI: motivation, convergence and numerical results. *Mathematical Programming*, n. 127, p. 245–295, 2011.
- 20 CARTIS, C.; GOULD, N. I. M.; TOINT, L. Universal regularization methods - varying the power, the smoothness and the accuracy. *SIAM Journal on Optimization*, n. 1, p. 595–615, 2016.
- 21 GRAPIGLIA, G. N.; NESTEROV, Y. Regularized newton methods for minimizing functions with hölder continuous Hessians. *SIAM Journal on Optimization*, n. 27, p. 478–506, 2017.
- 22 CURTIS, F. E.; ROBINSON, D. P.; SADAMI, M. A trust-region algorithm with a worst-case iteration complexity of $O(\epsilon^{-3/2})$. *Mathematical Programming*, n. 162, p. 1–32, 2017.
- 23 DUSSAULT, J. P. Simple unified convergence proofs for the trust-region and a new ARC variant. *Technical report, University of Sherbrooke, Sherbrooke, Canada*, 2015.
- 24 MARTÍNEZ, J. M.; RAYDAN, M. Cubic-regularization counterpart of a variable-norm trust-region method for unconstrained minimization. *Journal of Global Optimization*, to appear, n. 68, p. 367–385, 2017.
- 25 NESTEROV, Y.; POLYAK, B. T. Cubic regularization of Newton's method and its global performance. *Mathematical Programming*, n. 108, p. 177–205, 2006.
- 26 GRIEWANK, A. The modification of Newton's method for unconstrained optimization by bounding cubic terms. *Technical Report NA/12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge*, 1981.
- 27 WEISER, M.; DEUFLHARD, P.; ERDMANN, B. Affine conjugate adaptive newton methods for nonlinear elastomechanics. *Optimization Methods and Software*, n. 22, p. 413–431, 2007.
- 28 GRAPIGLIA, G. N.; NESTEROV, Y. Globally convergent second-order schemes for minimizing twice differentiable functions. *CORE Discussion Paper 2016/28, Université Catholique de Louvain, Louvain, Belgium*, 2016.
- 29 BIRGIN, E. G.; GARDENGHI, J. L.; MARTÍNEZ, J. M.; SANTOS, S. A.; TOINT, L. Worst-case evaluation complexity for unconstrained nonlinear optimization using high-order regularized models. *Mathematical Programming*, n. 163, p. 259–368, 2017.

- 30 MARTÍNEZ, J. M. On high-order model regularization for constrained optimization. *SIAM J. optim.*, n. 27, p. 2447–2458, 2017.
- 31 BONNEL, H.; IUSEM, A. N.; SVAITER, B. F. Proximal methods in vector optimization. *SIAM J. Optim.*, n. 15, p. 953–970, 2005.
- 32 BONNEL, H.; KAYA, C. Y. Optimization over efficient set of multi-objective convex optimal control problems. *J. Optim. Theory Appl.*, n. 147, p. 93–112, 2010.
- 33 FLIEGE, J.; DRUMMOND, L. M. G.; SVAITER, B. F. Newton’s method for multiobjective optimization. *SIAM J. Optim.*, n. 20, p. 602–626, 2009.
- 34 FLIEGE, J.; VAZ, A. I. F.; VICENTE, L. N. Complexity of gradient descent for multiobjective optimization. *Optimization Methods and Software*.
- 35 GRAPIGLIA, G. N. *Três contribuições em Otimização Não Linear e Não Convexa*. Tese (Doutorado) — Universidade Federal do Paraná, 2014.
- 36 MORÉ, J. J.; GARBOW, B.; HILLSTROM, K. E. Testing unconstrained optimization software. *ACM Transactions on Mathematical Software*, n. 7, p. 17–41, 1981.
- 37 CALDERÓN, L.; DINIZ-EHRHARDT, M. A.; MARTÍNEZ, J. M. On high-order model regularization for multiobjective optimization. *Optimization Methods and Software*.
- 38 BERTSEKAS, D. *Nonlinear Programming*. 2. ed. Massachusetts, USA: Athena Scientific, Belmont, 1999.
- 39 DUTTA, J.; KAYA, C. Y. A new scalarization and numerical method for constructing weak pareto front of multi-objective optimization problems. *Optim.*, n. 60, p. 1091–1104, 2011.
- 40 GEOFFRION, A. M. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, n. 22, p. 618–630, 1968.
- 41 PASCOLETTI, A.; SERAFINI, P. Scalarizing vector optimization problems. *J. Optim. Theory Appl.*, n. 42, p. 499–524, 1984.
- 42 BURACHIK, R. S.; KAYA, C. Y.; RIZVI, M. M. A new scalarization technique to approximate pareto fronts on problems with disconnected feasible sets. *J. Optim. Theory Appl.*, n. 162, p. 428–446, 2014.
- 43 DRUMMOND, L. M. G.; SVAITER, B. F. A steepest descent method for vector optimization. *Journal of Computational and Applied Mathematics*, n. 175, p. 395–414, 2005.
- 44 MARTÍNEZ, J. M.; SANTOS, S. A. *Métodos computacionais de otimização*. IMPA: 20º Colóquio Brasileiro de Matemática, 1995.
- 45 BIRGIN, E. G.; MARTÍNEZ, J. M. *Practical Augmented Lagrangian Methods for Constrained*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2014. v. 10 of Fundamentals of Algorithms.

APÊNDICE A – Problemas Testados

Problema 1

$$\begin{aligned} \min & \quad [10(x_2 - x_1^2), 1 - x_1] \\ \text{s.a} & \quad -2 \leq x_i \leq 2, \quad i=1,2. \end{aligned}$$

Problema 2

$$\begin{aligned} \min & \quad [-13 + x_1 + ((5 - x_2)x_2 - 2)x_2, -29 + x_1 + ((x_2 + 1)x_2 - 14)x_2] \\ \text{s. a} & \quad -2 \leq x_1 \leq 2 \\ & \quad -4 \leq x_2 \leq 4 \end{aligned}$$

Problema 3

$$\begin{aligned} \min & \quad [10^4 x_1 x_2 - 1, e^{-x_1} + e^{-x_2} - 1.0001] \\ \text{s. a} & \quad -1 \leq x_i \leq 1, \quad i=1,2. \end{aligned}$$

Problema 4

$$\begin{aligned} \min & \quad [x_1 - 10^6, x_2 - 2 \cdot 10^6, x_1 x_2 - 2] \\ \text{s. a} & \quad -1 \leq x_i \leq 1, \quad i=1,2. \end{aligned}$$

Problema 5

$$\begin{aligned} \min & \quad [1.5 - x_1(1 - x_2), 2.25 - x_1(1 - x_2^2), 2.625 - x_1(1 - x_2^3)] \\ \text{s. a} & \quad -3 \leq x_i \leq 3, \quad i=1,2. \end{aligned}$$

Problema 6

$$\begin{aligned} \min & \quad [4 - (e^{x_1} + e^{x_2}), 6 - (e^{2x_1} + e^{2x_2})] \\ \text{s. a} & \quad -5 \leq x_i \leq 5, \quad i=1,2. \end{aligned}$$

Problema 7

$$\begin{aligned} \text{minimizar} & \quad [f_1(x), 10(x_1^2 + x_2^2)^{\frac{1}{2}} - 1, x_3] \\ \text{s. a} & \quad -3 \leq x_i \leq 3, \quad i=1,2,3, \end{aligned}$$

onde

$$f_1(x) = \begin{cases} 10(x_3 - \frac{5}{\pi} \arctan(\frac{x_1}{x_2})), & x_1 > 0, \\ 10(x_3 - \frac{5}{\pi} \arctan(\frac{x_1}{x_2}) - 0.5), & x_1 < 0. \end{cases}$$

Problema 8

$$\begin{aligned} & \text{minimizar} && [f_1(x), \dots, f_{15}(x)] \\ & \text{s. a} && -2 \leq x_i \leq 3, \quad i=1,2,3, \end{aligned}$$

onde

$$f_i(x) = y_i - \left(x_1 + \frac{u_i}{v_i x_2 + w_i x_3} \right),$$

$$u_i = i,$$

$$v_i = 16 - i,$$

$$w_i = \min(u_i, v_i),$$

i	y	i	y	i	y
1	0.14	5	0.29	11	0.73
2	0.18	6	0.32	12	0.96
3	0.22	7	0.35	12	1.34
4	0.25	8	0.39	14	2.10
		9	0.37	15	4.39
		10	0.58		

Problema 9

$$\begin{aligned} & \text{minimizar} && [f_1(x), \dots, f_{15}(x)] \\ & \text{s. a} && -3 \leq x_i \leq 3, \quad i=1,2,3, \end{aligned}$$

onde

$$f_i(x) = x_1 \exp \left[\frac{-x_2(t_i - x_3)^2}{2} \right] - y_i,$$

$$t_i = (8 - i)/2,$$

i	y_i
1,15	0.0009
2,14	0.0044
3,13	0.0175
4,12	0.0540
5,11	0.1295
6,10	0.2420
7,9	0.3521
8	0.3989

Problema 10

$$\begin{aligned} & \text{minimizar} && [f_1(x), \dots, f_{16}(x)] \\ & \text{s. a} && -5 \leq x_i \leq 1, \quad i=1,2,3, \end{aligned}$$

onde $f_i(x) = x_1 \exp \left[\frac{x_2}{(t_i + x_3)} \right] - y_i$,
 $t_i = 45 + 5i$,

i	y_i	i	y_i
1	34780	9	8261
2	28610	10	7030
3	23650	11	6005
4	19630	12	5147
5	16370	13	4427
6	13720	14	3820
7	11540	15	3307
8	9744	16	2872

Problema 11

$$\begin{aligned} & \min && [f_1(x), f_2(x), f_3(x)] \\ & \text{s. a} && -4 \leq x_i \leq 4, \quad i=1,2,3, \end{aligned}$$

onde

$$\begin{aligned} f_i(x) &= e^{\left[-\frac{|3iy_i x_2|^{x_3}}{x_1} \right]} - t_i, \\ y_i &= 25 + (-50 \ln(t_i))^{2/3}, \text{ com } t_i = \frac{i}{100}, \quad i=1,2,3. \end{aligned}$$

Problema 12

$$\begin{aligned} & \min && [f_1(x), f_2(x), f_3(x)] \\ & \text{s. a} && -4 \leq x_i \leq 4, \quad i=1,2,3, \end{aligned}$$

onde

$$f_i(x) = e^{-t_i x_1} - e^{-t_i x_2} - x_3(e^{-t_i} - e^{-10t_i}), \quad i=1,2,3, \text{ com } t_i = (0.1)i.$$

Problema 13

$$\begin{aligned} & \min && [f_1(x), f_2(x), f_3(x), f_4(x)] \\ & \text{s. a} && -4 \leq x_i \leq 2, \quad i=1,2,3,4, \end{aligned}$$

onde

$$\begin{aligned} f_1(x) &= x_1 + 10x_2, \\ f_2(x) &= 5^{1/2}(x_3 - x_4), \\ f_3(x) &= (x_2 - 2x_3)^2, \\ f_4(x) &= 10^{1/2}(x_1 - x_4)^2. \end{aligned}$$

Problema 14

$$\begin{aligned} \min \quad & [f_1(x), f_2(x), f_3(x), f_4(x), f_5(x), f_6(x)] \\ \text{s. a} \quad & -3 \leq x_i \leq 3, \quad i=1,2,3,4, \end{aligned}$$

onde

$$\begin{aligned} f_1(x) &= 10(x_2 - x_1^2), \\ f_2(x) &= 1 - x_1, \\ f_3(x) &= (90)^{1/2}(x_4 - x_3^2), \\ f_4(x) &= 1 - x_3, \\ f_5(x) &= (10)^{1/2}(x_2 + x_4 - 2), \\ f_6(x) &= (10)^{1/2}(x_2 - x_4). \end{aligned}$$

Problema 15

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_{11}(x)] \\ \text{s. a} \quad & -5 \leq x_i \leq 1, \quad i=1,2,3,4, \end{aligned}$$

onde

$$f_i(x) = y_i - \frac{x_1(u_i^2 + u_i x_2)}{(u_i^2 + u_i x_3 + x_4)},$$

i	y_i	u_i	i	y_i	u_i
1	0.1957	4.0000	7	0.0456	0.1250
2	0.1947	2.0000	8	0.0342	0.1000
3	0.1735	1.0000	9	0.0323	0.0833
4	0.1600	0.5000	10	0.0235	0.0714
5	0.0844	0.2500	11	0.0246	0.0625
6	0.0627	0.1670			

Problema 16

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_4(x)] \\ \text{s. a} \quad & -3 \leq x_i \leq 3 \quad i=1,2,3,4, \end{aligned}$$

onde

$$f_i(x) = (x_1 + t_i x_2 - \exp[t_i])^2 + (x_3 + x_4 \text{sen}(t_i) - \cos(t_i))^2, \text{ com } t_i = i/5$$

Problema 17

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_{33}(x)] \\ \text{s. a} \quad & -4 \leq x_i \leq 4, \quad i=1,2,3,4,5, \end{aligned}$$

onde

$$f_i(x) = y_i(x_1 + x_2[-t_i x_4] + x_3 \exp[-t_i x_5]),$$

i	y_i	i	y_i	i	y_i	i	y_i
1	0.844	10	0.784	19	0.538	28	0.431
2	0.908	11	0.751	20	0.522	29	0.424
3	0.932	12	0.718	21	0.506	30	0.420
4	0.936	13	0.685	22	0.490	31	0.414
5	0.925	14	0.658	23	0.478	32	0.411
6	0.908	15	0.628	24	0.467	33	0.406
7	0.881	16	0.603	25	0.457		
8	0.850	17	0.580	26	0.448		
9	0.818	18	0.558	27	0.438		

Problema 18

$$\begin{aligned} \min & \quad [f_1(x), \dots, f_6(x)] \\ \text{s. a} & \quad -3 \leq x_i \leq 1, \quad i=1,2,3,4,5,6. \end{aligned}$$

onde

$$f_i(x) = x_3 \exp[-t_i x_1] - x_4 \exp[-t_i x_2] + x_6 \exp[-t_i x_5] - y_i,$$

$$t_i = 0.1i,$$

$$y_i = \exp[-t_i x_1] - 5 \exp[-10t_i] + 3 \exp[-4t_i].$$

Problema 19

$$\begin{aligned} \min & \quad [f_1(x), \dots, f_{65}(x)] \\ \text{s. a} & \quad -5 \leq x_i \leq 3, \quad i=1, \dots, 11. \end{aligned}$$

onde

$$f_i(x) = y_i - (x_1 \exp[-t_i x_5] + x_2 \exp[-(t_i - x_9)^2 x_6] + x_3 \exp[-(t_i - x_{10})^2 x_7] + x_4 \exp[-(t_i - x_{11})^2 x_8]),$$

$$t_i = (i - 1)/10,$$

i	y_i	i	y_i	i	y_i
1	1.366	23	0.694	45	0.672
2	1.191	24	0.644	46	0.708
3	1.112	25	0.624	47	0.633
4	0.1.013	26	0.661	48	0.668
5	0.991	27	0.612	49	0.645
6	0.885	28	0.558	50	0.632
7	0.831	29	0.533	51	0.591
8	0.847	30	0.495	52	0.559
9	0.786	31	0.500	53	0.597
10	0.725	32	0.423	54	0.625
11	0.746	33	0.395	55	0.739
12	0.679	34	0.375	56	0.710
13	0.608	35	0.372	57	0.729
14	0.655	36	0.391	58	0.720
15	0.616	37	0.396	59	0.636
16	0.606	38	0.405	60	0.581
17	0.602	39	0.428	61	0.428
18	0.626	40	0.429	62	0.292
19	0.651	41	0.523	63	0.162
20	0.724	42	0.562	64	0.098
21	0.649	43	0.607	65	0.098
22	0.649	44	0.653		

Problema 20

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_{31}(x)] \\ \text{s. a} \quad & -4 \leq x_i \leq 3, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = \sum_{j=2}^4 (j-1)x_j t_i^{j-2} - \left(\sum_{j=1}^4 x_j t_i^{j-1} \right)^2 - 1, \text{ com } t_i = i/29, 1 \leq i \leq 29,$$

$$f_{30}(x) = x_1,$$

$$f_{31}(x) = x_2 - x_1^2 - 1.$$

Problema 21

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_4(x)] \\ \text{s. a} \quad & -5 \leq x_i \leq 4, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$f_{2i-1}(x) = 10(x_{2i} - x_{i-1}^2),$$

$$f_{2i}(x) = 1 - x_{2i-1}.$$

Problema 22

$$\begin{aligned} \min \quad & [f_1(x), \dots, f_4(x)] \\ \text{s. a} \quad & -4 \leq x_i \leq 1, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$\begin{aligned} f_{4i-3}(x) &= x_{4i-3} + 10x_{4i-2}, \\ f_{4i-2}(x) &= 5^{1/2}(x_{4i-1} - x_{4i}), \\ f_{4i-1}(x) &= (x_{4i-2} - 2x_{4i-1})^2, \\ f_{4i} &= 10^{1/2}(x_{4i-3} - x_{4i}). \end{aligned}$$

Problema 23

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -2 \leq x_i \leq 2, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$\begin{aligned} f_i(x) &= a^{1/2}(x_i - 1), \quad 1 \leq i \leq 4, \\ f_5(x) &= \left(\sum_{j=1}^4 x_j^2 \right) - \frac{1}{4}, \quad a = 10^{-5}. \end{aligned}$$

Problema 24

$$\begin{aligned} \min & [f_1(x), \dots, f_8(x)] \\ \text{s. a} & -3 \leq x_i \leq 1, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$\begin{aligned} f_1(x) &= x_1 - 0.2, \\ f_i(x) &= a^{1/2} \left(\exp \left[\frac{x_i}{10} \right] + \exp \left[\frac{x_{i-1}}{10} \right] - y_i \right), \quad 2 \leq i \leq 4, \\ f_i(x) &= a^{1/2} \left(\exp \left[\frac{x_{i-3}}{10} \right] - \exp \left[\frac{-1}{10} \right] \right), \quad 4 < i < 8, \\ f_8(x) &= \left(\sum_{j=1}^4 (5-j)x_j^2 \right) - 1, \\ a &= 10^{-5}, \quad y_i = \exp \left[\frac{i}{10} \right] + \exp \left[\frac{i-1}{10} \right]. \end{aligned}$$

Problema 25

$$\begin{aligned} \min & [f_1(x), \dots, f_6(x)] \\ \text{s. a} & -2 \leq x_i \leq 5, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$\begin{aligned} f_i(x) &= x_i - 1, \quad i = 1, \dots, 4, \\ f_5(x) &= \sum_{j=1}^4 j(x_j - 1), \\ f_6(x) &= \left(\sum_{j=1}^4 j(x_j - 1) \right)^2. \end{aligned}$$

Problema 26

$$\begin{aligned} \min & [f_1(x), \dots, f_4(x)] \\ \text{s. a} & -3 \leq x_i \leq 3, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = 4 - \sum_{j=1}^4 \cos x_j + i(1 - \cos x_i) - \operatorname{sen} x_i, 1 \leq i \leq 4.$$

Problema 27

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -3 \leq x_i \leq 5, i=1, \dots, 5, \end{aligned}$$

onde

$$f_i(x) = x_i + \sum_{j=1}^5 x_j - 5, 1 \leq i < 5,$$

$$f_5(x) = \left(\prod_{j=1}^5 x_j \right) - 1.$$

Problema 28

$$\begin{aligned} \min & [f_1(x), \dots, f_4(x)] \\ \text{s. a} & -2 \leq x_i \leq 2, i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = 2x_i - x_{i-1} - x_{i+1} + h^2(x_i + t_i + 1)^3/2,$$

$$h = 1/5, t_i = ih.$$

Problema 29

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -3 \leq x_i \leq 5, i=1, \dots, 5, \end{aligned}$$

onde

$$f_i(x) = x_i + h \left[(1 - t_i) \sum_{j=1}^i t_j (x_j + t_j + 1)^3 + t_i \sum_{j=i+1}^5 (1 - t_j) (x_j + t_j + 1)^3 \right] / 2,$$

$$h = 1/6, t_i = ih.$$

Problema 30

$$\begin{aligned} \min & [f_1(x), \dots, f_4(x)] \\ \text{s. a} & -2 \leq x_i \leq 3, i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = (3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1,$$

onde $x(0) = x(5) = 0$.

Problema 31

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -2 \leq x_i \leq 3, i=1, \dots, 5, \end{aligned}$$

onde

$$f_i(x) = x_i(2 + 5x_i^2) + 1 - \sum_{i \in J_i} x_j(1 + x_j),$$

$$J_i = \{j : j \neq i, \max(1, i - 5) \leq j \leq \min(n, i + 1)\}.$$

Problema 32

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -2 \leq x_i \leq 4, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = x_i - \frac{2}{m} \left(\sum_{j=1}^4 x_j \right) - 1, \quad 1 \leq i \leq 4,$$

$$f_i(x) = -\frac{2}{m} \left(\sum_{j=1}^4 x_j \right) - 1, \quad i = 5.$$

Problema 33

$$\begin{aligned} \min & [f_1(x), \dots, f_5(x)] \\ \text{s. a} & -3 \leq x_i \leq 3, \quad i=1, \dots, 4, \end{aligned}$$

onde

$$f_i(x) = i \left(\sum_{j=1}^4 jx_j \right) - 1.$$

Problema 34

$$\begin{aligned} \min & [f_1(x), \dots, f_6(x)] \\ \text{s. a} & -4 \leq x_i \leq 3, \quad i=1, \dots, 5, \end{aligned}$$

onde

$$\begin{aligned} f_1(x) &= 1, \quad f_6(x) = -1, \\ f_i(x) &= (i - 1) \left(\sum_{j=2}^4 jx_j \right) - 1, \quad 2 \leq i \leq 5. \end{aligned}$$

Problema 35

$$\begin{aligned} \min & [f_1(x), f_2(x), f_3(x)] \\ \text{s. a} & -5 \leq x_i \leq 1, \quad i=1, 2, 3, \end{aligned}$$

onde

$$f_i(x) = \frac{1}{3} \sum_{j=1}^3 T_i(x_j) - \int_0^1 T_i(x) dx,$$

onde T_i é o i -ésimo polinômio de Chebyshev no intervalo $[0, 1]$, definido como

$$\int_0^1 T_i(x) dx = 0, \quad \text{para } i \text{ ímpar.}$$

$$\int_0^1 T_i(x) dx = \frac{-1}{i^2 - 1}, \quad \text{para } i \text{ par.}$$

APÊNDICE B – Tabelas

Tabela 1 – Resultados do Algoritmo 4 quando o modelo é o polinômio de Taylor de ordem 2.

NP	x^0	F	Outer	Inner
1	$x_I^0 = (1, -2)$	$[-29.391, 0.0284]$	4	38
	$x_{II}^0 = (-1.5, -2)$	$[-42.3502, 2.4996]$	2	22
	$x_{III}^0 = (-2, -0.8)$	$[-46.308, 2.9572]$	1	13
2	$x_I^0 = (-2, 2)$	$[-7.9105, -47.145]$	14	57
	$x_{II}^0 = (-2, 1)$	$[-12.9832, -42.983]$	11	40
	$x_{III}^0 = (-2, 2.5)$	$[-4.2531, -49.1227]$	2	22
3	$x_I^0 = (0, 1)$	$[-1.6317, 0.368]$	15	376
	$x_{II}^0 = (1, 0)$	$[-3.3475, 0.3886]$	16	161
	$x_{III}^0 = (-1, 0.5)$	$[-4867.6, 2.3]$	2	16
4	$x_I^0 = (-1, -1)$	$[-1E+06, -0.93548, -1.1249]$	1	9
	$x_{II}^0 = (0, -1)$	$[-1E+06, -0.9316, -2]$	1	9
	$x_{III}^0 = (1, 0.5)$	$[-1E+07, 0.5, -1.5000]$	2	26
5	$x_I^0 = (-3, -3)$	$[13.4537, -21.5783, 85.9058]$	4	22
	$x_{II}^0 = (-3, -2)$	$[-1.4875, -6.7309, 18.3793]$	5	23
	$x_{III}^0 = (3, -2.5)$	$[-7.5711, 14.6046, -35.6272]$	1	5
6	$x_I^0 = (5, 5)$	$[-217, -24490]$	1	3
	$x_{II}^0 = (5, 4)$	$[-199, -24714]$	8	28
	$x_{III}^0 = (5, -5)$	$[-144, -24803]$	7	14
7	$x_I^0 = (-3, 2, -1)$	$[-50.583, 25.8576, -1.0001]$	48	34
	$x_{II}^0 = (-2.5, 0, -3)$	$[-79.7454, 14.9891, -0.9745]$	2	17
	$x_{III}^0 = (-1.5, 2, -3)$	$[-64.757, 14.3809, -0.9466]$	3	20
8	$x_I^0 = (-1, 0, 1)$	$[1.8174, -3.7380, -1.1060$ $-0.6742, -0.4712, -0.3528$ $-0.2673, -0.1893, -0.4950$ $-0.6659, -1.0491, -1.6191$ $-2.5723, 4.4787, -10.1878]$	5	54
	$x_{II}^0 = (-2, -1, 0)$	$[1.6258, 1.7398, 1.8658$ $1.9967, 2.1570, 2.3328$ $2.5432, 2.8120, 3.0779$ $3.6691, 4.3528, 5.3833$ $7.0975, 10.5259, 20.8211]$	18	126
	$x_{III}^0 = (1, -1, 3)$	$[-1.3465, -1.2574, -1.1782$ $-1.1163, -1.0497, -0.9973$ $-0.9481, -0.8915, -0.8347$ $-0.5225, -0.2293, 0.2155$ $0.9535, 2.4295, 6.8675]$	11	103
9	$x_I^0 = (-2, 2, 1)$	$[-0.0009, -0.0044, -0.0176$ $-0.0552, -0.1414, -0.3207$ $-0.6912, -1.3475, -2.0753$ $-2.2746, -1.6864, -0.9284$ $-0.2676, -0.0569, -0.0080$	3	35
	$x_{II}^0 = (-1, 0, 3)$	$[-2.8624, -2.8656, -2.0789$ $-2.9162, -0.9931, -3.1076$ $-3.2202, -0.2700, -3.2268$		

	$x_{III}^0 = (-1, 2, -0.5)$	-3.1209, -3.0132, -2.9430 -2.9124, -2.9058, -2.9093] [-0.0009, -0.0044, -0.0180 -0.0601, -0.1804, -0.4999 1.1489, -1.9013, -2.0809, -1.4560, -0.6498, -0.1901 -0.0392, -0.0065, -0.0010]	7	24
10	$x_I^0 = (-4, 1, -3)$	[-3.4785, -2.8615, -2.3655 -1.9635, -1.6375, -1.3725 -1.1545, -0.9749, -0.8266 -0.7035, -0.6010, -0.5152 -0.4432, -0.3825, -0.3312 -0.2877]	8	42
	$x_{II}^0 = (1, 1, -1)$	[-3.4780, -2.8610, -2.3650 -1.9630, -1.6370, -1.3720 -1.1540, -0.9744, -0.8261 -0.7030, -0.6005, -0.5147 -0.4427, -0.3820, -0.3307 -0.2872]	2	13
	$x_{III}^0 = (-4, 0, -1)$	[-3.4785, -2.8615, -2.3655, -1.9635, -1.6375, -1.3725 -1.1545, -0.9749, -0.8266 -0.7035, -0.6010, -0.5152 -0.4432, -0.3825, -0.3312 -0.2877]	3	19
11	$x_I^0 = (3, 0.5, 4)$	[-0.01, -0.02, -0.03]	2	7
	$x_{II}^0 = (-4, 2, 0)$	[1.28, -1.27, 1.26]	2	9
	$x_{III}^0 = (4, -4, 0)$	[0.7669, -0.7569, 0.7469]	2	11
12	$x_I^0 = (-2, 3, 4)$	[-1.3216, -1.4969, -1.2716]	4	23
	$x_{II}^0 = (0, -3, 0.5)$	[-0.6183, -1.0638, -1.8051]	3	17
	$x_{III}^0 = (-3, 2, 4)$	[-1.6141, -1.579, -0.8515]	11	14
13	$x_I^0 = (-3, -1, 1, 1)$	[-15.3577, -2.2780, 3.1391, 48.3185]	10	130
	$x_{II}^0 = (-2, 2, 0, 1)$	[14.2052, -3.2466, 2.8908, 27.4823]	9	126
	$x_{III}^0 = (-1, -3, 2, -4)$	[-38.8400, 5.5818, 40.9959, 20.6372]	2	21
14	$x_I^0 = (-1, -3, 2, 0)$	[-45.0040, 2.2249, -84.5147 -1.1480, -0.7150]	3	44
	$x_{II}^0 = (-2, 0, -1, 1)$	[-40.9824, 3.0087, -1.0356, 2.0081 -3.6565, -0.3068]	11	209
	$x_{III}^0 = (1, -2, 0, 3)$	[-37.5199, -0.2277, 26.7976, 0.8220 -4.3900, -1.6131]	5	55

15	$x_I^0 = (-4, -2, -1, 1)$	[1.8406, -1.9919, -14.0991, -19.6675, -8.4125, -4.9303, -3.4354, -2.6322, -2.1246, -1.7869 - 1.5353]	5	80
	$x_{II}^0 = (-5, 1, 0, -3)$	[1.2679, 17.2940, -7.1334, -1.5709 -0.5730, -0.3401, -0.2431, -0.1909, -0.1520, -0.1326, -0.1108]	4	41
	$x_{III}^0 = (-2, 0, -4, -1)$	[33.1019, -2.8800, -0.9068, -0.3720, -0.2112, -0.1511, -0.1237, -0.1068, -0.0889, -0.0829, -0.0703]	7	91
16	$x_I^0 = (-1, 3, 0, 1)$	[1.7270, 0.8324, 0.3134, 0.1188]	3	42
	$x_{II}^0 = (-2, 0, 1, 2)$	[0.5215, 7.8838, 11.2741, 5.8818]	5	60
	$x_{III}^0 = (-3, -1, -1, -2)$	[4.6922, 2.9296, 0.1452, 3.3640]	7	28
17	$x_I^0 = (0, -1, 2, -2, -4)$	[-1.0465, -1.2105, -1.2345, -1.2385 -1.2275, -1.2105, -0.1835, -2.1525 -1.1205, -1.0805, -1.0535, -1.0205 -0.8875, -0.9605, -0.9680, -0.9055 -0.8025, -0.8605, -0.8405, -0.8245 -0.8085, -0.7925, -1.7805, -2.7695 -0.7595, -0.7505, -0.7405, -0.7335 -0.7265, -0.7305, -0.7465, -0.7135 -0.7085]	3	57
	$x_{II}^0 = (-3, 2, 1, 0, 3)$	[-2.1533, -2.2173, -2.2413, -2.2453 -2.2343, -2.2173, -2.1903, -1.1593, -0.1273, -2.8933, -2.0603, -2.0273, -1.9943, -1.9673, -1.9373, -1.9123 -1.8893, -1.8673, -1.8473, -1.8313, -3.8153, -1.7993, -1.0873, -0.7763, -1.3663, -1.7573, -1.7473, -1.7403, -1.7333, -2.7293, -1.7283, -1.7203, -1.7153]	2	44
	$x_{III}^0 = (2, -1, -4, 0, 4)$	[-0.7195, -4.7835, -4.8075, -4.82115, -1.8005, -4.7835, -3.7565, -4.7255, -2.6935, -4.6595, -4.6265, -4.5935, -4.5605, -4.8335, -4.5035, -4.1785, -4.7555, -0.4335, -1.4135, -4.3975, -4.3815, -4.3655, -4.3535, -4.3425, -4.3325, -3.3235, -4.3135, -4.3065, -4.2995, -4.2955, -4.2895, -4.2865, -4.2815]	2	19
18	$x_I^0 = (-2, -3, -1, 0, 0, 1)$	[-3.2108, -2.9202, -3.1737 -3.3717, -3.9466, -2.0400]	5	105
	$x_{II}^0 = (1, -1, 0, 0, -2, -1)$	[-3.4188, -2.1156, -3.3606, -4.5544, -5.8297, -4.2288]	2	17
	$x_{III}^0 = (-3, 1, 0, 0, -2, -1)$	[-2.9390, -4.5723, -3.7393, -5.8374, -3.9948, -2.2484]	2	21

19	x_1^0	_____	5	13
	x_{II}^0	_____	3	5
	x_{III}^0	_____	3	9
20	$x_I^0 = (-2, 3, -2.5, 0.5)$	$[-1.4713, -1.3069, -1.1750 - 1.0721,$ $-0.9954, -0.9419, -0.9093, -0.8951,$ $-0.8972, -0.9138 - 0.9429, -0.9832,$ $-1.0332, -1.0917, -1.1575, -1.2298,$ $-1.3076, -1.3904, -1.4774, -1.5683$ $-1.6626, -1.7600, -1.8602, -1.9632$ $-2.0688, -2.1769, -2.2876, -2.4009$ $-2.5169, -1.9044, -4.6267]$	2	21
	$x_{II}^0 = (-3.5, 1, -1, 2.5)$	$[-11.8353, -11.6556 - 11.4678, -11.2676$ $-11.0510, -10.8139 - 10.5526, -10.2635$ $-9.9431, -9.5883, -9.1959, -8.7635$ $-8.2885, -7.7691, -7.2037, -6.5915$ $-5.9320, -5.2256, -4.4734, -3.6775$ $-2.8408, -1.9675, -1.0629, -0.1338$ $0.8117, 1.7637, 2.7106, 3.6387$ $4.5323, -3.4676, -13.0241]$	4	54
	$x_{III}^0 = (3, -4, 0, 1)$	$[-12.7713, -12.0358, -11.3330, -10.6633$ $-10.0272, -9.4246, -8.8558, -8.3207,$ $-7.8189, -7.3502, -6.9141, -6.5100$ $-6.1372, -5.7947, -5.4817, -5.1970$ $-4.9396, -4.7080, -4.5010, -4.3170$ $-4.1547, -4.0122, -3.8880, -3.7803$ $-3.6873, -3.6072, -3.5382, -3.4782$ $-3.4256, 2.9453, -9.6750]$	2	15
21	$x_I^0 = (-3, 2, -1, 0)$	$[-82.7148, 4.1593, -15.4441, 2.2564]$	3	22
	$x_{II}^0 = (1, 4, -2, -4)$	$[27.6831, -0.4480, -82.4692, 3.1084]$	5	43
	$x_{III}^0 = (-3, -2, 4, 0)$	$[-114.5591, 4.0120, -161.3397, 2.9593]$	2	15
22	$x_I^0 = (1, 0, -2, -4)$	$[9.0267, -3.0092, 2.9617, 11.8983]$	3	32
	$x_{II}^0 = (-2, -3, 1, 1)$	$[-33.8402, -1.8382, 12.5545, 26.6246]$	3	31
	$x_{III}^0 = (-4, 1, 0, 0)$	$[2.0344, -0.5109, 0.1848, 50.0880]$	7	53
23	$x_I^0 = (0, -1, 2, 0)$	$[-0.0032, -0.0062, -0.000,$ $-0.0039, -0.2010]$	6	65
	$x_{II}^0 = (1, -2, 0, -0.5)$	$[-0.0004, -0.0094, -0.0032,$ $-0.0048, -0.0210]$	4	63
	$x_{III}^0 = (0.5, -1.5, 2, 0)$	$[-0.0019, -0.0093, 0.0022,$ $-0.0035, -0.1147]$	3	39
24	$x_I^0 = (-3, 0, 1, 0)$	$[-3.1818, -0.0012, -0.0010,$ $-0.0027, 0.0003, 0.0007,$ $0.0003 - 0.9995]$	7	99
	$x_{II}^0 = (2, -1, -2.5, 0)$	$[1.7998, -0.0016, -0.0032,$ $-0.0027, -0.0000, -0.0004,$ $0.0003, -0.9998]$	5	81
	$x_{III}^0 = (-1, 2, 0, -3)$	$[-1.608, 60.0003, -0.0018,$ $-0.0041, 0.0010, 0.0003,$ $-0.0005, 6.1773]$	6	58

25	$x_I^0 = (4, 2.5, -1, 3)$	[4.0000, 3.8458, 0.4757, -1.2968, -5.1873, 26.9081]	6	20
	$x_{II}^0 = (-2, 0, 0, 1)$	[-2.9914, -0.9220, -0.9220, -0.0846, -0.3384, 0.1145]	4	37
	$x_{III}^0 = (5, 1.5, 3.5, -1)$	[3.9999, 2.3377, 3.2896, -1.7770, -7.1081, 50.5255]	3	36
26	$x_I^0 = (2.5, -3, -1, 0)$	[2.8105, 7.8439, 2.8773, 9.9108]	4	60
	$x_{II}^0 = (-2, 3, -1.5, -3)$	[5.1211, 9.1111, 11.1011, 3.0911]	3	34
	$x_{III}^0 = (0, 1, -2, -1.5)$	[5.8508, 6.7775, 2.7043, 8.6310]	8	70
27	$x_I^0 = (-1, 0, -3, 2, 4)$	[-4.8056, -3.8248, -6.5962 -2.2379, 0.3961]	2	18
	$x_{II}^0 = (5, -2, 1, 2.5, 3)$	[1.6630 - 5.2914, -2.4121, -1.2977, 0.8511]	3	24
	$x_{III}^0 = (3, -3, 4, -1.5, 0)$	[-4.2485, -10.0772, -3.2497, -9.1547, 3.0774]	2	21
28	$x_I^0 = (-2, 0, 2, -1)$	[-3.9145, -1.7018, 4.7602, -1.7467]	5	34
	$x_{II}^0 = (0, 0.5, -2, -1.5)$	[-0.5747, 3.1569, -2.5111, -2.9796]	6	61
	$x_{III}^0 = (1, 2, -0.5, 1.5)$	[-0.2784, 4.8275, -2.8975, 3.3612]	7	79
29	$x_I^0 = (-3, 0, 5, -2.5, 1)$	[-2.4015, 0.0672, 10.8105, -1.9592, -1.2076]	10	59
	$x_{II}^0 = (-2, 5, -3, 3, 1)$	[-1.9723, 9.8080, -2.9107, 5.0850, 1.5341]	4	37
	$x_{III}^0 = (4, 2.5, -1.5, 3, -3)$	[5.6219, 3.4488, -1.5096, 4.8254, -3.0079]	4	49
30	$x_I^0 = (0, 2.5, -1.5, 1)$	[-2.6330, -0.7367, -12.1908, 2.8671]	3	19
	$x_{II}^0 = (-2, 1.5, -1, 3)$	[-15.4041, 3.6657, -4.8403, 0.8381]	6	41
	$x_{III}^0 = (2, -0.5, -1, 0)$	[0.4153, -1.1851, -6.6392, 2.4153]	2	13
31	$x_I^0 = (-2, 0, 2, 1.5)$	[-46.6869, -2.2139, 79.8433, 16.7264]	5	54
	$x_{II}^0 = (-2, 2.5, -1.5, 3)$	[-53.2211, 99.8729, -29.9576, 129.9962]	3	39
	$x_{III}^0 = (-1, 2, 0, -0.5)$	[-8.3346, 82.0791, -0.7975, -6.2396]	6	138
32	$x_I^0 = (-1, 0, 2, 4)$	[-3.5996, -2.6040, -0.7535, 1.3997, -2.5998]	6	55
	$x_{II}^0 = (1, 1.5, -1.5, -2)$	[0.6985, 1.1965, -1.8905 -2.2000, -0.2000]	3	51
	$x_{III}^0 = (3, -0.5, 4, -2)$	[2.5510, -0.9373, 3.3889 -2.1954, -0.2031]	4	32
33	$x_I^0 = (0, -1, 2, -3)$	[-12.9752, -24.9505, -16.9257, -18.9009, -4.9762]	2	19
	$x_{II}^0 = (2, -2, 0, -1.5)$	[-6.8974, -12.7949, -18.6923, -14.5897, -8.4872]	3	27
	$x_{III}^0 = (-2, -1.5, 0, -3)$	[-12.9116, -14.8231, -26.7347, -18.6462, -6.9578]	1	8

34	$x_I^0 = (-3, -2.5, 2.5, 1.5)$	$[-1.010, 5.1471, 13.2942,$ $17.4413, -1.0000]$	2	8
	$x_{II}^0 = (4, 1, -0.5, 0)$	$[-1.0000 - 2.5889, -4.1778$ $-5.7668, -1.010]$	2	12
	$x_{III}^0 = (2, 1, 3, -3.5)$	$[-1.0010, 6.4441, 11.881,$ $21.3322, -1.010]$	2	10
35	$x_I^0 = (-4.5, 0, -3)$	$[-2.5666, 18.7666, -150.0865]$	7	147
	$x_{II}^0 = (1, -1, -2)$	$[-0.8069, 2.9192, -8.8738]$	11	110
	$x_{III}^0 = (-3, -2, 1)$	$[-1.4867, 8.1715, -41.4987]$	6	94

Tabela 2 – Resultados do Algoritmo 4 quando o modelo é o polinômio de Taylor de ordem 3

NP	x^0	F	Outer	Inner
1	$x_I^0 = (1, -2)$	$[-29.309, -0.0075]$	2	28
	$x_{II}^0 = (-1.5, -2)$	$[-39.1692, 2.4517]$	1	10
	$x_{III}^0 = (-2, -0.8)$	$[-44.2517, 2.9102]$	1	11
2	$x_I^0 = (-2, 2)$	$[-7.9035, -47.146]$	11	48
	$x_{II}^0 = (-2, 1)$	$[-10.103, -46.382]$	8	37
	$x_{III}^0 = (-2, 2.5)$	$[-4.3019, -49.0744]$	1	14
3	$x_I^0 = (0, 1)$	$[-2.7066, 0.3735]$	3	48
	$x_{II}^0 = (1, 0)$	$[-1.6615, 0.368]$	15	207
	$x_{III}^0 = (-1, 0.5)$	$[-5000.2, 2.3001]$	1	10
4	$x_I^0 = (-1, -1)$	$[-1E+06, -0.8467, -1.2832]$	1	8
	$x_{II}^0 = (0, -1)$	$[-1E+06, -0.7452, -2]$	1	5
	$x_{III}^0 = (1, 0.5)$	$[-1E+07, 0.4599, -1.6147]$	1	24
5	$x_I^0 = (-3, -3)$	$[7.5501, -3.8606, -20.9574]$	1	1
	$x_{II}^0 = (-3, 2)$	$[-0.5229, -3.8177, -11.5301]$	3	6
	$x_{III}^0 = (3, -2.5)$	$[-5.5350, 12.8025, -30.7913]$	1	1
6	$x_I^0 = (5, 5)$	$[-281, -10678]$	1	2
	$x_{II}^0 = (5, 4)$	$[-198, -24784]$	6	21
	$x_{III}^0 = (5, -5)$	$[-124, -26443]$	1	1
7	$x_I^0 = (-3, 2, -1)$	$[-50.6405, 25.8372, -0.9935]$	4	23
	$x_{II}^0 = (-2.5, 0, -3)$	$[-30.3086, 18.2938, -1.9691]$	1	9
	$x_{III}^0 = (-1.5, 2, -3)$	$[-65.0035, 26, -2.9362]$	1	19
8	$x_I^0 = (-1, 0, 1)$	$[1.4070, 1.5695, 1.8364,$ $2.4307, 6.3063, -1.5397,$ $0.1321, 0.5771, 0.2704,$ $0.0982, -0.2869, -0.8595,$ $-1.8172, -3.7327, -9.4691]$	1	7
	$x_{II}^0 = (-2, -1, 0)$	$[1.5482, -1.6614, 1.7867,$ $-1.9172, 2.0773, 2.2537,$ $2.4657, 2.7382, 3.0050,$ $3.5973, 4.2826, 5.3155,$ $7.0337, 0.4700, 0.7891]$	16	144
	$x_{III}^0 = (1, -1, -3)$	$[-1.3646, -1.2734, -1.1944,$ $-1.1336, -1.0688, -1.0183,$ $-0.9711, -0.9165, -0.8598,$ $-0.5476, -0.2546, 0.1900,$ $0.9276, 12.4029, 6.8386]$	9	99
9	$x_I^0 = (-2, 2, -1)$	$[-0.0051, -0.0247, -0.0960,$ $-0.2983, -0.7417, -1.4774,$ $-2.3595, -3.0254, -3.1193,$ $-2.5896, -1.7332, -0.8362,$ $-0.0083, -0.14381 - 0.0409]$	1	4
	$x_{II}^0 = (-1, 0, 3)$	$[-1.5874, -1.5961, -1.6075,$ $-1.6354, -1.6955, -1.7861,$ $-1.8680, -1.0807, -1.7942,$		

	$x_{III}^0 = (-1, 2, -0.5)$	$[-1.6395, -1.4778, -1.3493,$ $-1.2564, -1.1843, -1.1197]$	6	41
		$[-0.0010, -0.0054, -0.0239,$ $-0.0848, -0.2401, -0.5379,$ $-0.9418, -1.2743, -1.3201,$ $-1.8394, -0.6188, -0.2776,$ $-0.0936, -0.8237, -0.0045]$	4	27
10	$x_I^0 = (-4, 1, -3)$	$[-3.4785, -2.8615, -2.3655,$ $-1.9635, -1.6375, -1.3725,$ $-1.1545, -0.9749, -0.8266,$ $-0.7035, -0.6010, -0.5152,$ $-0.4432, -0.3825, -0.3312,$ $-0.2877]$	1	5
	$x_{II}^0 = (1, 1, -1)$	$[-3.4780, -2.8610, -2.3650,$ $-1.9630, -1.6370, -1.3720,$ $-1.1540, -0.9744, -0.8261,$ $0.7030, -0.6005, -0.5147,$ $-0.4427, -0.3820, -0.3307,$ $-0.2872]$	1	11
	$x_{III}^0 = (-4, 0, -1)$	$[-3.4785, -2.8615, -2.3655,$ $-1.9635, -1.6375, -1.3725,$ $-1.1545, -0.9749, -0.8266,$ $-0.7035, -0.6010, -0.5152,$ $-0.4432, -0.3825, -0.3312,$ $-0.2877]$	1	12
11	$x_I^0 = (3, 0.5, 4)$	$[-0.01, -0.02, -0.03]$	1	3
	$x_{II}^0 = (-4, 2, 0)$	$[1.3857, -1.3757, 1.3657]$	1	4
	$x_{III}^0 = (4, -4, 0)$	$[0.7061, -0.6961, 0.6861]$	1	3
12	$x_I^0 = (-2, 3, 4)$	$[-1.6184, -1.4293, -1.2881]$	2	15
	$x_{II}^0 = (0, -3, 0.5)$	$[-0.5914, -1.0956, -1.6719]$	1	5
	$x_{III}^0 = (-3, 2, 4)$	$[-1.0809, -1.9014, -0.1685]$	4	14
13	$x_I^0 = (-3, -1, 1, 1)$	$[-15.1462, -2.0751, 3.8466, 48.1258]$	9	126
	$x_{II}^0 = (-2, 2, 0, 1)$	$[14.5927, -3.2715, 3.0120, 27.4908]$	6	90
	$x_{III}^0 = (-1, -3, 2, -4)$	$[-33.7175, 10.9947, 34.9804, 26.0461]$	1	9
14	$x_I^0 = (-1, -3, 2, 0)$	$[-36.5257, 1.8078, -84.9657, -1.8279,$ $-18.8437, -0.6454]$	2	19
	$x_{II}^0 = (-2, 0, -1, 1)$	$[-41.3982, 2.9596, -1.3711, 1.9315,$ $-4.9858, -0.3035]$	3	22
	$x_{III}^0 = (1, -2, 0, 3)$	$[-32.3648, -0.1030, 26.3831, 0.9386,$ $-3.9054, -1.5193]$	2	30

15	$x_I^0 = (-4, -2, -1, 1)$	[1.4534, -1.2143, -4.3300, -4.1163, -2.3926, -1.6197, -1.2168, -0.9755, -0.8080, -0.6960, -0.6046]	2	17
	$x_{II}^0 = (-5, 1, 0, -3)$	[1.3232, 14.8995, -10.9126, -2.1652, -0.7936, -0.4765, -0.3415, -0.2680, -0.2154, -0.1864, -0.1575]	1	9
	$x_{III}^0 = (-2, 0, -4, -1)$	[72.1449, -2.6414, -0.8351, -0.3256, -0.1754, -0.1805, -0.0968, -0.0829, -0.0673, -0.0632, -0.0523]	4	64
16	$x_I^0 = (-1, 3, 0, 1)$	[1.2376, 0.4975, 0.3356, 0.0882]	1	11
	$x_{II}^0 = (-2, 0, 1, 2)$	[0.7124, 6.0424, 11.4307, 6.0367]	3	33
	$x_{III}^0 = (-3, -1, -1, -2)$	[5.0502, 2.8391, 0.0564, 3.0365]	4	36
17	$x_I^0 = (0, -1, 2, -2, -4)$	[-1.1406, -1.2046, -1.2286, -1.2326, -1.2216, -1.2046, -2.1776, -1.1466, -1.1146, -0.0806, -1.0476, -1.0146, -0.9816, -0.8546, -1.9246, -0.8996, -0.8766, -0.0546, -0.8346, -0.8186, -0.8026, -1.7866, -0.7746, -0.7636, -0.7536, -0.7446, -0.7346, -0.7276, -0.7306, -0.0160, -0.7406, -0.7876, -0.7026]	1	16
	$x_{II}^0 = (-3, 2, 1, 0, 3)$	[-2.1544, -2.2184, -2.2424, -2.2464, -2.2354, -2.2184, -2.1914, -0.1604, -2.1284, -2.2944, -2.0614, -2.0284, -1.9954, -1.9684, -1.8384, -1.9134, -1.8904, -1.8684, -1.8484, -1.8324, -1.8164, -0.8004, -1.7884, -1.7774 -1.7674, -1.6584, -1.7484, -1.7414, -2.7344, -1.3304, -1.1244, -0.7214, -1.7164]	1	38
	$x_{III}^0 = (2, -1, -4, 0, 4)$	[-1.7634, -4.8274, -4.6510, -4.3554, -0.8444, -4.8274, -1.8004, -2.7694, -3.7374, -4.7034, -4.6704, -4.6374, -4.6044, -4.5774, -1.5474, -4.5224, -0.4994, -4.1774, -4.4574, -4.3414, -4.4254, -4.4094, -4.3974, -4.3864, -4.3764, -3.3674, -4.3574, -4.3504, -4.3434, -4.3394, -4.3334, -4.3304, -4.3254]	1	17
18	$x_I^0 = (-2, -3, -1, 0, 0, 1)$	[-2.1981, -3.9037, -2.1517, -2.3422, -4.6064, -2.9847]	1	23
	$x_{II}^0 = (1, -1, 0, 0, -2, -1)$	[-3.4076, -1.0992, -4.3397, -2.5302, -3.8037, -5.2031]	1	13
	$x_{III}^0 = (-3, 1, 0, 0, -2, -1)$	[-3.1091, -3.7846, -2.0129, -4.1962, -5.4685, -4.8741]	1	18

19	x_1^0	_____	1	7
	x_{II}^0	_____	1	2
	x_{III}^0	_____	2	5
20	$x_I^0 = (-2, 3, -2.5, 0.5)$	[−1.4456, −1.3041 − 1.1941, −1.1129 −1.0580, −1.0270, −1.0179, −1.0286, −1.0573, −1.1025, −1.1626, −1.2363, −1.3227, −1.4206, −1.5294, −1.6483 −1.7769, −1.9150, −2.0623, −2.2188 −2.3846, −2.5601, −2.7457, −2.9419 −3.1495, −3.3693, −3.6023, −3.8496 −4.1125, −1.8400, −4.3857]	1	15
	$x_{II}^0 = (-3.5, 1, -1, 2.5)$	[−11.3877, −11.1562, −10.9172, −10.6670 −10.4021, −10.1189, −9.8143, −9.4851, −9.1286, −8.7419, −8.3229, −7.8692, −7.3793, −6.8518, −6.2856, −5.6803, −5.0361, −4.3537, −3.6346, −2.8810, −2.0960, −1.2837, −0.4494, 0.4007, 1.2587, 2.1157, 2.9610, 3.7824, 4.5656, −3.4364, −12.8086]	1	12
	$x_{III}^0 = (3, -4, 0, 1)$	[−12.7742, −12.0381, −11.3345, −10.6581, −10.0272, −9.4248, −8.8560, −8.3209, −7.8189, −7.3470, −6.9130, −6.5103, −6.1395, −5.7947, −5.4822, −5.1993, −4.9398, −4.7080, −4.5010, −4.3192, −4.1547, −4.0148, −3.8901, −3.7783, −3.6881, −3.6082, −3.5411, −3.4785, −3.4312, 2.9522, −9.6834]	1	10
21	$x_I^0 = (-3, 2, -1, 0)$	[−75.5245, 4.0675, −17.4568, 2.2124]	1	14
	$x_{II}^0 = (1, 4, -2, -4)$	[19.64, −0.2153, −83.4710, 3.0231]	1	18
	$x_{III}^0 = (-3, -2, 4, 0)$	[−113.2920, 3.9770, −160.0948, 2.9829]	1	5
22	$x_I^0 = (1, 0, -2, -4)$	[−38.9998, 5.2604, 0.4971, 79.0567]	4	64
	$x_{II}^0 = (-2, -3, 1, 1)$	[−33.7206, −1.7233, 13.1843, 26.7433]	6	84
	$x_{III}^0 = (-4, 1, 0, 0)$	[−0.0061, −0.4235, 2.3649, 10.7043]	3	9
23	$x_I^0 = (0, -1, 2, 0)$	[−0.0032, −0.0071, −0.0031, −0.0031, −0.2500]	3	34
	$x_{II}^0 = (1, -2, 0, -0.5)$	[−0.0013, −0.0089, −0.0033, −0.0049, 0.0503]	1	11
	$x_{III}^0 = (0.5, -1.5, 2, 0)$	[−0.0015, −0.0085, 0.0031, −0.0030, −0.1187]	2	31
24	$x_I^0 = (-3, 0, 1, 0)$	[−3.2000, −0.0011, −0.0011, −0.0026, 0.0002, 0.0007, 0.0002, −0.9683]	7	99
	$x_{II}^0 = (2, -1, -2.5, 0)$	[1.7892, −0.0017, −0.0032, −0.0026, −0.0000, −0.0004, 0.0003, −0.9973]	2	26
	$x_{III}^0 = (-1, 2, 0, -3)$	[−1.4467, 0.0003, −0.0018, −0.0042, 0.0009, 0.0003, −0.0004, 7.1911]	5	49

25	$x_I^0 = (4, 2.5, -1, 3)$	$[2.3725, 0.2888, -2.9986,$ $-0.4790, -1.9160, 3.6712]$	2	12
	$x_{II}^0 = (-2, 0, 0, 1)$	$[-2.9565, -0.9567, -1.1371,$ $0.0532, 0.2138, 0.0453]$	3	19
	$x_{III}^0 = (5, 1.5, 3.5, -1)$	$[3.5414, -0.1993, 1.4621,$ $-2.9949, -11.9798, 143.5148]$	1	10
26	$x_I^0 = (2.5, -3, -1, 0)$	$[2.8213, 7.8492, 2.8770, 8.9049]$	1	5
	$x_{II}^0 = (-2, 3, -1.5, -3)$	$[4.1750, 8.9412, 10.7075, 2.4737]$	2	22
	$x_{III}^0 = (0, 1, -2, -1.5)$	$[4.2779, 4.5659, 2.8538, 5.1417]$	5	29
27	$x_I^0 = (-1, 0, -3, 2, 4)$	$[-4.8182, -3.8450, -6.6176,$ $-2.2056, 0.3822]$	1	13
	$x_{II}^0 = (5, -2, 1, 2.5, 3)$	$[1.4808, -5.1786, -2.3130,$ $-1.1648, 0.9176]$	1	11
	$x_{III}^0 = (3, -3, 4, -1.5, 0)$	$[-4.3407, -10.0792, -3.3377,$ $-9.1586, 3.0793]$	1	13
28	$x_I^0 = (-2, 0, 2, -1)$	$[-3.7773, -1.8861, 4.6971, -1.6900]$	1	6
	$x_{II}^0 = (0, 0.5, -2, -1.5)$	$[-0.3681, 3.2035, -2.5594, -2.9826]$	5	63
	$x_{III}^0 = (1, 2, -0.5, 1.5)$	$[-0.2665, 4.8120, -2.9425, 3.2507]$	4	48
29	$x_I^0 = (-3, 0, 5, -2.5, 1)$	$[-3.0514, 0.0883, -0.4591,$ $-2.6780, 0.4156]$	7	35
	$x_{II}^0 = (-2, 5, -3, 3, 1)$	$[-1.9693, 9.5448, -2.8153,$ $4.8880, 1.7274]$	2	25
	$x_{III}^0 = (4, 2.5, -1.5, 3, -3)$	$[5.4324, 3.3347, -1.4943,$ $4.7369, -2.8585]$	3	38
30	$x_I^0 = (0, 2.5, -1.5, 1)$	$[4.9757, -18.0872, -10.4263, 1.5751]$	2	15
	$x_{II}^0 = (-2, 1.5, -1, 3)$	$[-15.0008, 3.8812, -5.5911, 0.9406]$	1	9
	$x_{III}^0 = (2, -0.5, -1, 0)$	$[0.3739, -1.2408, -6.3902, 2.3846]$	1	7
31	$x_I^0 = (-2, 0, 2, 1.5)$	$[-44.0557, -11.3972, 39.9338, 125.4465]$	1	3
	$x_{II}^0 = (-2, 2.5, -1.5, 3)$	$[-50.7133, 119.9665, -30.1461, 127.3442]$	1	11
	$x_{III}^0 = (-1, 2, 0, -0.5)$	$[-6.9463, 56.0845, 0.8132, -1.3564]$	3	47
32	$x_I^0 = (-1, 0, 2, 4)$	$[-3.6227, -2.6487, -0.7050,$ $1.3705, -2.5804]$	3	42
	$x_{II}^0 = (1, 1.5, -1.5, -2)$	$[0.57721.0864 - 1.8637,$ $-2.20 - 0.2]$	1	16
	$x_{III}^0 = (-3, -0.5, 4, -2)$	$[-2.4157, -0.9025, 2.1807,$ $-1.8759, -0.4161]$	2	8
33	$x_I^0 = (0, -1, 2, -3)$	$[-12.9997, -24.9631, -16.9992,$ $-18.9989, -4.8886]$	1	8
	$x_{II}^0 = (2, -2, 0, -1.5)$	$[-6.8942, -12.7883, -18.6825,$ $-14.5767, -8.4709]$	2	17
	$x_{III}^0 = (-2, -1.5, 0, -3)$	$[-12.9816, -14.9994, -26.9447,$ $-18.9262, -6.5078]$	1	6

34	$x_I^0 = (-3, -2.5, 2.5, 1.5)$	$[-1.0011, 4.9520, 10.9040,$ $16.8561, -1.0000]$	1	6
	$x_{II}^0 = (4, 1, -0.5, 0)$	$[-1.0000, -3.2623, -5.5245$ $-7.7868, -1.0021]$	1	9
	$x_{III}^0 = (2, 1, 3, -3.5)$	$[-1.0015, 5.6627, 12.3253$ $18.9880, -1.0000]$	1	8
35	$x_I^0 = (-4.5, 0, -3)$	$[-2.5115, 18.7536, -149.9481]$	4	63
	$x_{II}^0 = (1, -1, -2)$	$[-0.8479, 2.9829, -8.9554]$	7	77
	$x_{III}^0 = (-3, -2, 1)$	$[-1.6824, 8.1842, -42.9398]$	5	85

Tabela 3 – Resultados do Algoritmo 4 para polinômios de ordem 2 e de ordem 3

NP	x^0	Taylor 2	Taylor 3
2	$x_I^0 = (-2, 2)$	14	11
10	$x_I^0 = (-4, 1, -3)$	2	1
	$x_{II}^0 = (1, 1, -1)$	3	1
	$x_{III}^0 = (-4, 0, -1)$	1	1
11	$x_I^0 = (3, 0.5, 4)$	2	1
20	$x_{III}^0 = (3, -4, 0, 1)$	2	1
23	$x_I^0 = (0, -1, 2, 0)$	6	3
	$x_{II}^0 = (1, -2, 0, -0.5)$	4	1
	$x_{III}^0 = (0.5, -1.5, 2, 0)$	3	2